Spike synchrony generated by modulatory common input through NMDA-type synapses

Nobuhiko Wagatsuma,1,2 Rüdiger von der Heydt,2 and Ernst Niebur2
1School of Science and Engineering, Tokyo Denki University, Saitama, Japan; and 2Zanvyl Krieger Mind/Brain Institute, Johns Hopkins University, Baltimore, Maryland

Submitted 22 December 2015; accepted in final form 30 June 2016

When a neuron activates other neurons through excitatory NMDA-type synapses

Recent neurophysiological results show that in V1 and V2, presumed modulatory top-down input due to contour grouping increases synchrony but that additional modulatory input due to top-down attention does not change tight synchrony (correlations on the order of milliseconds) and decreases loose synchrony (tens of milliseconds). These findings are understood from our model of integrate-and-fire neurons under the assumption that contour grouping as well as attention lead to additive modulatory input through NMDA-type synapses. In contrast, circuits with common projections through model α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors did not exhibit the paradoxical decrease of synchrony with increased input. Our results suggest that NMDA receptors play a critical role in top-down response modulation in the visual cortex.

NEW & NOTEWORTHY

We compared our model results with recently published neurophysiological data on spike train correlations between neurons in cortical areas V1 and V2 in behaving macaques (Martin and von der Heydt 2015). Zhou et al. (2000) found that orientation-selective neurons in these areas, especially those in V2, are often border ownership selective (BOS). That is, a neuron responds to an edge in its receptive field with a higher firing rate when the edge is part of a figure on one side (the neuron’s preferred side) of the receptive field compared with when the same edge is part of a figure on the other side. This differential firing has been attributed to the modulatory influence of feedback from hypothetical grouping cells (G-cells) at higher levels (Craft et al. 2007). Martin and von der Heydt (2015) recorded from pairs of neurons responding either to contours of the same object or to contours of different objects (Fig. 1A). For the individual neuron, the object could be either on its preferred side of border ownership or on its nonpreferred side. Accordingly, if the preferred sides of both members of a neuron pair were consistent with the common object, the pair was called “consistent” for this object; all other pairs were called “inconsistent.” According to the grouping cell model (Craft et al. 2007), neurons that are members of a consistent
pair receive feedback from the same G-cells, whereas neurons of inconsistent pairs receive feedback from different G-cells (Fig. 1B). The model thus predicts that stimulation by a common object produces common input to both neurons and, therefore, enhanced spiking synchrony of consistent but not inconsistent pairs. This was, indeed, observed (Martin and von der Heydt 2015). As mentioned above, in the Craft et al. (2007) model, the feedback from G-cells is assumed to be modulatory. While it is fairly obvious that fast, driving input (like from glutamatergic AMPA-type) synapses enhances synchrony between postsynaptic neurons, it is not a priori clear that this applies to slow, modulatory (NMDA) synapses as well. Thus, the first main question of the present study was whether NMDA-type synapses can produce increased correlation between spike trains, as observed in the Martin and von der Heydt (2015) experiments.

These experiments also required the animals to attend to specific visual stimuli, and the effect of selective attention on neural responses was recorded. The other critical question of the present study was, therefore, whether our model predicts the observed effects of top-down attention on firing rates and spiking synchrony, both of which were modulated by attentional selection. Craft et al. (2007) proposed that attention to objects selectively excites the G-cells that represent the target (attended) object. Because feedback from G-cells modulates the responses of BOS neurons representing local features (Fig. 1C), the feature signals of the target object can be “gated” by activating the corresponding G-cell (as shown in the computational model by Mihalas et al. 2011). Martin and von der Heydt (2015) found that attention to an object increased the firing rates, as predicted in this model, but did not change the level of tight synchrony (defined below) and actually decreased the level of loose synchrony (ditto). This seems difficult to reconcile with the proposed mechanism of attention since increased common input from G-cells should increase, not decrease, synchrony between BOS cells. This intuition, however, is based on the assumption that BOS cells receive driving input from G-cells, like that provided by glutamatergic synapses of the AMPA type. There is, however, substantial evidence that cortical back projections use slow, modulatory NMDA receptors rather than fast, driving AMPA receptors. We therefore studied the effect of attention in model BOS neurons that receive feedforward input from edge-selective cells through synapses with AMPA-like kinetics and modulatory feedback from G-cells through synapses with NMDA characteristics. We also considered combinations of both. Here, we show that this model is in agreement with the observed changes in both firing rates and synchrony levels between neural pairs.

**MATERIALS AND METHODS**

**Network architecture.** Figure 2A shows the architecture of our network model, which consists of two BOS populations (BOS_L and BOS_R) and one grouping cell population (G). The function of G-cells requires very little specificity in their responses; they are fully characterized by center location and size of receptive field (Craft et al. 2007). While BOS neurons are observed in cortical area V2 and neighboring areas V1 and V4, it is unknown where G-cells reside.

We include in our model only the minimum number of neurons and synaptic connections necessary to understand the fundamental mechanism of top-down modulation of BOS neurons in binding and selective attention. The arrows in Fig. 2 indicate the synaptic connections of our network model. The BOS neuron whose receptive field is shown by the left oval has right side-of-figure preference and is therefore named BOS_L; the BOS neuron whose receptive field is shown by the right oval has left side-of-figure preference and is therefore named BOS_R. BOS_L and BOS_R neurons receive input from two sources. Bottom-up input represents visual stimuli, whereas top-down input from G-cells mediates grouping structure and the influence of selective attention. For simplicity, all inputs are modeled as stochastic random processes with Poisson statistics, where each event stands for an incoming action potential. As shown by the arrows in Fig. 2, bottom-up inputs are independent processes, whereas the signals from G-cells are common to BOS_L and BOS_R neurons representing the same object (consistent cell pair). The
firing rate ($v_G$) of a G-neuron in the “bound” condition (Fig. 2A) is higher than in the “unbound” condition (Fig. 2B). In the latter situation, two objects (gray shapes) are located left and right in the scene and two different G-cell populations (not shown) are activated, one for each object. In addition to this G-cell activity corresponding to the geometry of the scene, we assume that attention to an object increases the firing rate of the corresponding G-cell further, as in the “bound-attend” condition (Fig. 2C).

The BOS neurons in our model are simulated as integrate-and-fire neurons. We used generic but biologically realistic parameters, as proposed by Buehlmann and Deco (2008) and Deco and Thiele (2011) and described in detail as follows. We chose a firing threshold ($\theta$) = −50 mV and reset potential ($V_{reset}$) = −60 mV. Membrane capacitance ($C_m$) was 0.5 nF, and membrane conductance ($g_m$) = $C_m/\tau_m$ = 25 nS, resulting in membrane time constant ($\tau_m$) = 20 ms. The dynamics of the subthreshold membrane potential ($V$) of a BOS neuron are given by the following:

$$\frac{dV(t)}{dt} = -\frac{V(t)}{\tau_m} + \frac{I_{syn}(t)}{C_m} \quad (1)$$

where $I_{syn}(t)$ comprises the synaptic currents that flow into BOS neurons. They are the sum of two different types of inputs: $I_{vis}$ from bottom-up visual stimuli and $I_G$ from top-down modulatory input from G-cells, as follows:

$$I_{syn}(t) = I_{vis}(t) + I_G(t) \quad (2)$$

Bottom-up excitatory postsynaptic currents ($I_{vis}$) are mediated by glutamatergic receptors of the AMPA type (Buehlmann and Deco 2008; Deco and Thiele 2011). G-cells provide common modulatory inputs for BOSL and BOSR neurons (Figs. 1C and 2). A fundamental hypothesis of our model is that modulatory feedback $I_G$ takes the form of currents through NMDA receptors. All NMDA receptors have a voltage dependence that is controlled by the extracellular $\text{Mg}^{2+}$ concentration ($[\text{Mg}^{2+}]$) (Jahr and Stevens 1990), which we assume as $[\text{Mg}^{2+}] = 1 \text{mM}$. A large variety of NMDA receptors are expressed in the mammalian brain, with different physiological properties that depend on the combination of their subunits (for a review, see Cull-Candy et al. 2001). Few details are known about the properties of NMDA receptors in the visual cortical circuits we are concerned with; therefore, we will use a standard computational model for generic NMDA receptors (Wang 1999). The details of these synaptic currents are given in the APPENDIX.

**Numeric experiments.** In our model, BOS neurons integrate bottom-up inputs, representing object borders, with top-down influences mediating the perceptual grouping structure and attention to objects (Figs. 1 and 2). Since the contents of the classical receptive fields are identical for all visual inputs considered (compare the three configurations of Figs. 1A and configurations A, B, and C in Fig. 2), the bottom-up input has the same statistics in all three situations, always modeled as Poisson spike trains with a mean rate of 200 Hz except for in Fig. 9, where we study how results change when we vary this rate. This input should be understood as originating from a population of visually responsive neurons rather than from a single neuron. This rate as well as the rates for the top-down input for each stimulus condition (bound, unbound, attended, and unattended) are chosen to approximately reproduce the observed firing rates of BOS neurons (see also the discussion below).

The role of G-cells is to integrate the responses of BOS neurons to generate a representation of the visual scene in terms of protoobjects, thus providing a fast sketch of the location and rough shapes of objects in the scene (Fig. 1, B and C) (Craft et al. 2007; Russell et al. 2014). Since we focused on the interaction of modulatory top-down influences with the driving bottom-up input, we did not implement the activation of G-cells by BOS cells in response to an object but, instead, simply increased the G-cell activity in the bound condition and observed the influence of its activity on BOS cells. Likewise, we increased the activity of a G-cell to mimic attentional selection of the target object without being concerned about the source of attentional input. In the unbound condition (Figs. 1A, top, and 2B), G-cell activity was simulated as Poisson-distributed spike trains. We assumed that they fired spontaneously with an average frequency of 3 Hz. In the presence of an object in its receptive field, a G-cell increases its activity by about an order of magnitude. If the object is present but not attended (bound-ignore; Figs. 1A, middle, and 2A), the firing frequency increases to 25 Hz. If the object is attended (bound-attend; Figs. 1A, bottom, and 2C), the G-cell receives both bottom-up input from BOS cells and top-down input from attentional control areas and its firing frequency increases to 45 Hz, about doubling the rate for the bound-ignore condition. Although we will, for simplicity, refer to the activity of a (single) G-cell, the top-down spike train should be understood as activity originating from a population of G-cells, just as the bottom-up input is the combined activity from many visually responsive neurons. Note, again, that the contents of the classical receptive fields of both BOS neurons is identical in all experimental conditions (except for the results shown in Fig. 9).
We integrated the differential equations using a fourth-order Runge-Kutta algorithm with a time step of 0.1 ms. We simulated 50 trials of a length of 200 s each, for a total of 10,000 simulated biological seconds per condition. The first 750 ms of simulated results was always discarded to minimize the effect of transients [in analogy to the onset transients that were removed by Martin and von der Heydt (2015)], and the simulation was extended beyond 200 s by the length of the correlation window to allow computation of the correlation functions (see below). The code for the simulations was written in the C programming language.

Computation of neuronal synchrony. We quantified spike synchrony by first dividing time into bins of 1 ms width, each containing either 0 or 1 spike. A spike train was thus transformed into a stochastic binary vector in which each component takes on the value of 0 if no spike (BOSL or BOSR), and n is the bin index. $S_i(n)$ can be represented as a binary vector in which each component takes on the value of 0 if no spike is present in the interval $[n, n + 1)$ ms in the BOS neuron during trial $i$ or 1 if there is such a spike.

The correlogram between two spike trains $S_L(n)$ and $S_R(n)$ of BOSL and BOSR neurons is the mean of the cross-correlations over all trials $i$. The cross-correlation operator $(\otimes)$ is defined as follows:

$$C^i(\tau) = S_L^i \otimes S_R^i = \sum_{\mu=1000}^{201000} S_L^i(\mu + \tau)S_R^i(\mu)$$

where $wd$ = 250 ms is the maximal window of the cross-correlation function, and $\tau$ is the time lag between spike trains ($-wd \leq \tau \leq wd$). Note that the lower bound of the sum results in the removal of the onset activity of 750 ms duration, as discussed above.

Changes of firing rate, e.g., those produced by attention, of course change the correlations. We followed the experimental literature in compensating for any such effects by subtracting the average spike frequency from the neuron spike train for each trial (Roelfsema et al. 2004). The mean spike count per bin of spike train of BOS in trial $i$ was as follows:

$$f_j^i = \frac{1}{\theta} \sum_{n=1000}^{201000} S_j(n)$$

where $\theta = 200$ s is the length of the simulated trials. The cross-correlation function $[CC(\tau)]$ was computed as follows:

$$CC^i(\tau) = \frac{1}{\theta} (S_L^i - f_L^i) \otimes (S_R^i - f_R^i)$$

and the correlogram $[CCG(\tau)]$ was as follows:

$$CCG(\tau) = <CC^i(\tau)>_i$$

where $< >_i$ denotes the average over trial $i$. Since the spike trains are time-density functions (e.g., counts/ms), the cross-correlation and correlogram have the dimension of coincidences/s². Correlograms were smoothed using a Gaussian kernel of $\sigma = 4$ ms to facilitate comparison with the neurophysiological data (Martin and von der Heydt 2015).

The magnitude of synchrony between BOSL and BOSR neurons ($M^i$) is the integral of the correlogram (Eq. 5) in the range of $\pm T$:

$$M^i = \sum_{\tau=-T}^{T} CC^i(\tau) \times \text{bin size}$$

where bin size = 1 ms. The average magnitude of synchrony over trials (AM) is

$$AM = <M^i>_i$$

Loose synchrony (correlations on the order of tens of milliseconds) was computed using $T = 40$ ms. Tight synchrony (order of milliseconds) was computed from the “jitter-reduced” correlogram with $T = 5$ ms (see the appendix for details).

**RESULTS**

We performed numeric simulations of the experiments performed by Martin and von der Heydt (2015) and Dong et al. (2008). Figure 3A shows spike raster plots of BOSL and BOSR neurons for 50 simulated trials where G-cells were activated between 1000 and 2000 ms and inactive otherwise. As expected, activity of both neurons is increased by the feedback from G-cells. Note that the spike trains shown in Fig. 3 only serve to illustrate the network dynamics; there was no step in the mean firing rates of the spike trains used for the correlation analyses anywhere in this report.

Responses of BOS neurons to visual stimuli and selective attention. The three conditions shown in Fig. 1A were simulated by setting the mean firing rates of G-cells during the activation period ($\nu_G$) to 3 Hz (unbound-ignore), 25 Hz (bound-ignore), and 45 Hz (bound-attend). Figure 3B shows a summary of the average firing rates of G-cells and BOS neurons. The firing rates of BOS neurons ($\nu_{BOS}$) were significantly higher in the bound condition than in the unbound condition, and they were significantly higher in the bound-attend condition than in the bound-ignore condition, all in agreement with experimental results (Qiu et al. 2007; Martin and von der Heydt 2015).

An important result of Martin and von der Heydt (2015) was the observation that binding increased correlations between members of consistent pairs but not between members of inconsistent pairs, supporting the grouping-by-feedback hypothesis. This result is “built into” our model: since we assumed, for simplicity, that there was no interaction between G-cells and that visual input to all BOS neurons was independent, members of inconsistent pairs can, by design, have no correlations beyond chance in our model (result not shown). Of interest here are Martin and von der Heydt’s findings for consistent BOS pairs. Their results show that binding increases loose synchrony in the absence of attention compared with the unbound case and that, furthermore, selective attention to an object decreases synchrony between BOS neurons responding to that object (Fig. 3, C and D). We simulated this experiment by computing spike train correlations between BOSL and BOSR populations in the unbound-ignore, bound-ignore, and bound-attend situations. Figure 3E shows the correlations in the three simulated conditions. In the ignore conditions (unbound-ignore and bound-ignore), the correlation was much stronger for the bound configuration (Fig. 3E, black line) than for the unbound configuration (gray). In contrast, correlation for the bound-attend condition (red) was lower than for the bound-ignore condition, again as in the experimental data.

To compare the results between conditions, we computed the magnitude of the loose synchrony by integrating the correlogram at a $\pm 40$-ms interval around lag zero (see MATERIALS AND METHODS). Since the mean firing rates were subtracted before computing the cross-correlograms, this integral gives the increase in frequency of coincidences beyond the mean level (the frequency of coincidences within that interval ex-
range of cally varied the mean rate of the modulatory feedback in the activity of BOS neurons and their correlations, we systemati-

while increasing their firing rates.

between pairs of BOS neurons responding to bound features (2015) (see Fig. 3

D

with the physiological findings of Martin and von der Heydt

1422 SPIKE SYNCHRONY GENERATED BY MODULATORY COMMON INPUT

expected by chance). Figure 3F shows that there was a significant increase of synchrony from the unbound-ignore condition to the bound-ignore condition ($t$-test, $P < 0.01$). In contrast, loose synchrony decreased from the bound-ignore condition to the bound-attend condition ($P < 0.01$). Again, these results agreed with the physiological findings of Martin and von der Heydt (2015) (see Fig. 3D). Top-down visual attention mediated by NMDA synaptic receptors thus decreases loose synchrony between pairs of BOS neurons responding to bound features while increasing their firing rates.

To understand the influence of NMDA input on the mean activity of BOS neurons and their correlations, we systemati-

cally varied the mean rate of the modulatory feedback in the range of $\nu_G = 0$–100 Hz. Figure 4A shows that the firing rates of BOS neurons monotonically increased with increasing feed-

back frequency. In contrast, loose synchrony between BOS$_L$ and BOS$_R$ neurons showed nonmonotonic behavior, rising to a peak at $\sim 15$ Hz and then decreasing (Fig. 4B).

We also studied an alternative model in which feedback from G-cells to BOS cells was imparted by fast, AMPA-type glutamatergic synapses rather than NMDA-type synapses. In this model, increased feedback from G-cells resulted in an increase of mean firing rates (Fig. 5B) for BOS neurons. This is inconsistent with the physiological data (Martin and von der Heydt 2015), which led us to conclude that the feedback must have a strong component with slow (NMDA-type) kinetics.

**Tight synchrony.** The peaks in the correlograms of both experimental and simulated data are several tens of milliseconds wide. To determine whether these broad peaks “hide” evidence of tighter synchrony, Martin and von der Heydt (2015) extracted tight synchrony using an analysis in which the original spike trains were transformed by distributing the spikes randomly within intervals of width $\Delta$ (chosen as 20 ms; results did not sensitively depend on the exact choice). A large number of correlograms between such “jittered” spike trains
Fig. 4. Firing rates (A) and loose synchrony (B) for BOS neurons as a function of the mean rates of G-cells ($v_G = 0–100$ Hz). Data were obtained from 50 simulation trials for each mean $v_G$. Gray dots show the rate and synchrony for each simulated trial. A: average firing rates of BOS neurons are shown as white triangles in steps of 5 Hz of mean $v_G$. B: white squares show the average magnitude of the loose synchrony of 50 simulated trials, in steps of 5 Hz of mean $v_G$.

were computed, and their mean was subtracted from the original correlogram, resulting in a “jitter-reduced” correlogram. This procedure removes all correlations at times scales larger than the jitter interval, revealing any underlying tight synchrony. In the language of hypothesis testing, any significant remaining synchrony is support for rejecting the null hypothesis that the exact spike timing on a time scale smaller than $\Delta$ is irrelevant (Amarasingham et al. 2012; Smith and Kohn 2008; see the appendix).

Martin and von der Heydt (2015) found significant peaks at zero lag for consistent pairs in the bound conditions (ignore as well as attend) but not in the unbound condition, as shown in Fig. 6A. We applied the same analysis to our simulated data. Figure 6B shows jitter-reduced correlograms between model BOSL and BOSR neurons in the three stimulus conditions. Perhaps surprisingly, NMDA-based input did produce tight synchrony despite the long decay time constant of the NMDA currents. This can be understood from the fact that while the decay of these currents is slow, their onset is fast (a few milliseconds; Hestrin et al. 1990). As we did for mean rates and loose synchrony above, we again parametrically varied the frequency of feedback activity. The results shown in Fig. 6C indicate a trend for a maximum at intermediate frequencies, but it was much weaker than that shown in Fig. 4B. Interestingly, tight synchrony did not drop off beyond G-cell firing rates of 15 Hz, as loose synchrony did (Fig. 4B), but remained high up to much higher firing rates. These results suggest that selective attention mediated by NMDA synaptic receptors induces tight synchrony between BOS neurons over a broad range of feedback firing rates.

Synchrony between BOS and grouping cells. We have shown that the common modulatory inputs from G-cells induce spike synchrony between BOS neurons. To study the role of G-cells for the synchrony between BOS neurons in more detail, we computed the spike train correlation between G-cells and BOS neurons and investigated its dependency on the frequency of the feedback projections. Loose synchrony calculated by integrating the correlogram between G-cells and BOS neurons at a ±40-ms interval around lag zero showed nonmonotonic behavior (Fig. 7) similar to that between BOSL and BOSR neurons (Fig. 4B). However, correlation for the bound-attend condition (red line in the inset in Fig. 7) was higher and sharper than that for the bound-ignore condition (black line in inset in Fig. 7). The correlation between G-cells and BOS neurons showed a similar dependence on the mean firing frequency of G-cells as the correlation between BOS cells. These results suggest that the nonmonotonic relationship between G-cells and BOS cells is the cause of the behavior of synchrony between pairs of BOS neurons.

Influence of top-down synaptic strength. We were curious about how our results depend on the strength of the feedback modulation. We therefore simulated the model with various combinations of synaptic weights ($w_{G\text{BOS}}$) and NMDA conductances ($g_{\text{NMDA}}$).

We ran the model, varying the NMDA synaptic impact $w_{G\text{BOS}} \times g_{\text{NMDA}}$ between 0 and 100 nS, bracketing the value used so far...
Fig. 6. Tight synchrony between model BOS neurons. Reduced cross-correlation after subtraction of 20-ms interval jitter cross-correlation is shown. See MATERIALS AND METHODS and the APPENDIX for details. A: experimentally observed tight synchrony between consistent BOS neurons, replotted from Martin and von der Heydt 2015 for the unbound-ignore (gray), bound-ignore (black), and bound-attend (red) conditions. B: simulated tight synchrony for the unbound-ignore, bound-ignore, and bound-attend conditions. The same color scheme as in A was used. C: magnitudes of tight synchrony in the range ±5 ms around lag zero as function of G-cell frequency \(v_G\). Triangles correspond to data in B with the same colors. Tight synchrony clearly increased from the unbound to bound conditions (gray to black triangle), but no clear increase occurred from the bound-ignore to bound-attend conditions (black to red triangle).

Fig. 7. Synchrony between G- and BOS cells as a function of feedback frequency (\(v_G = 0–100\) Hz). Symbols are the same those as in Figs. 4 and 5. The inset shows their cross-correlations for the unbound-ignore (gray), bound-ignore (black), and bound-attend (red) conditions.

Fig. 8. Influence of NMDA feedback strength (\(w_{\text{BOS}}^G \times g_{\text{NMDA}}\)) on simulated BOS neuron responses. A: firing rates. B: loose synchrony integrated ±40 ms around lag zero. C: tight synchrony integrated ±5 ms around lag zero. Symbols are the same as in Fig. 4.

(~45 nS, see Eq. A3) about symmetrically on both sides. Plots summarizing the responses of BOS neurons are shown in Fig. 8. Firing rates increased slightly steeper than linearly with increasing NMDA synaptic impact (Fig. 8A) while synchrony increased clearly supralinearly (Fig. 8, B and C). Note that the results reported above for values of \(w_{\text{BOS}}^G \times g_{\text{NMDA}}\approx 45\) nS are at the lower end of the parameter range where synchrony increases substantially. While we used generic values for synaptic strengths, we do not know how closely they match those found in the feedback connections in the extrastriate cortex. Our results in Fig. 8 show that the generic values we used place the system just at the transition between very low and increasingly greater synchrony. Obtaining experimental data on the functional synaptic strength in these circuits would be of great interest.

Influence of bottom-up input frequency. We next investigated the influence of the strength of bottom-up inputs representing visual stimuli by carrying out simulations with a variety of input frequencies \((v_{\text{vis}})\) for three values of G-cell firing rates, \(v_G = 3, 25, \) and 45 Hz (the same as shown in Fig. 3). Simulated BOS firing rates and loose synchrony (range of ±40 ms around lag zero) are shown in Fig. 9, A and B, respectively. As already shown in Fig. 3B, the firing rates of BOS neurons increased with \(v_G\) (Fig. 9A). For all levels of \(v_G\), the firing rates of BOS neurons increased monotonically and approximately linearly with the input rate, \(v_{\text{vis}}\). In contrast, synchrony between BOS cells varied nonmonotonically with \(v_{\text{vis}}\) (Fig. 9B), increasing from zero to a maximum and then slowly decaying for large \(v_{\text{vis}}\) (over 300 Hz). It would be interesting to see if future experimental manipulations of different firing rates bear out these predictions. This should be possible by systematically varying the orientation of stimulus...
edges in the receptive fields relative to the cells’ preferred orientations [data in Dong et al. (2008) and in Martin and von der Heydt (2015) were always obtained at optimal orientation].

While the main point of Fig. 9B is to show how synchrony between BOS cells varies with the bottom-up input frequency ($\nu_{vis}$), we can also see how synchrony depends on $\nu_G$, at least for the three values of the latter shown in this figure, $\nu_G = 3, 25, \text{and} 45 \text{Hz}$. Nonmonotonic modulation of synchrony with $\nu_G$ can be seen only in the range of $\nu_{vis}$ values of about 125–200 Hz (as shown by the vertical order of symbols for the different values of $\nu_G$). For larger visual input frequencies ($\nu_{vis} \geq 225 \text{Hz}$), loose synchrony between BOS neurons increased monotonically with $\nu_G$. This result would not allow our interpretation of the observed higher synchrony for bound versus unbound and the decreased synchrony with attention shown in Fig. 3. To test the robustness of our findings of spike synchrony induced by modulatory common input through NMDA-type synapses, we again systematically varied the mean rate of the feedback in the range of $\nu_G = 0–100 \text{Hz}$ and computed the level of loose synchrony for values of $\nu_G = 150, 300, \text{and} 400 \text{Hz}$. Figure 9C shows that loose synchrony as a function of $\nu_G$ is a nonmonotonic function not only for the value $\nu_{vis} = 200 \text{Hz}$ shown in Fig. 4B but also for the larger range of input frequencies in this figure, even though the location of the peak varied somewhat for different values of $\nu_{vis}$. These results suggest that common modulatory feedback induces nonmonotonic modulations of synchrony in a wide range of firing rates of feedforward inputs.

**Diversity of glutamatergic synaptic currents.** There is strong evidence that feedback connections target synapses especially rich in NMDA receptors (Self et al. 2012), and, so far, we have assumed that the feedback is exclusively mediated by NMDA currents. This is a simplification, however, since it is unlikely that feedback connections are exclusively targeting NMDA-type synapses and that feedforward connections are exclusively AMPA driven. To study whether the results shown in Figs. 3, 4, and 6 are robust in the presence of such mixed projections, we extended our model to include modest admixtures of the other glutamatergic channels: AMPA in top-down projections and NMDA in bottom-up projections. In Fig. 10, we assumed that the strength of the AMPA receptor in feedback connections corresponded to 0.2$I_G$ ($\rho = 0.2$, where $\rho$ is the index for the strength of NMDA in the bottom-up projection and AMPA in the top-down projection; see Eq. A7 in the APPENDIX) and that of NMDA in the bottom-up postsynaptic currents corresponded to 0.2$I_G$ ($\rho = 0.2$; see also Eq. A6 in the APPENDIX). The average firing rates, spike train cross-correlation, and tight synchrony of BOS neurons for the extended model are shown in Fig. 10, A, B, and C, respectively. We found that this modification did not change our previous results substantially, as shown by comparison with Figs. 3, B and E, and B. The behavior changed when $\rho$ in feedback connections exceeded −0.4, at which point the correlation at lag zero increased monotonically with $\nu_G$ rather than showing a maximum at intermediate values (data not shown).

We also wondered about the robustness of our results with respect to assumptions made about the dynamics of AMPA receptors. Various decay dynamics of AMPA synaptic receptors have been reported (data summarized by Santhakumar et al. 2005 and Benke et al. 2001). To test the robustness of our findings, we reran the simulations of the model with AMPA receptors in the top-down projections $I_G$, but assuming four types of AMPA synapses, with postsynaptic decay time constants ($\tau_{\text{AMPA}}$) of 0.5, 1.0, 2.0, and 4.0 ms (see also Eq. A2). Equal proportions of these four components were assumed, and, as before, the total strength of these AMPA receptors in feedback inputs corresponded to 0.2$I_{vis}$. Average firing rates, spike train cross-correlation, and tight synchrony of BOS neurons for this model are shown in Fig. 10, D, E, and F, respectively. The results are similar to those of our original model (Fig. 3) and the model with the mixtures of NMDA and AMPA considered above.

As we did for our original model above, we again parametrically varied the frequency of feedback activity $\nu_{vis}$ with respect to these extended models. For both models, Fig. 10G shows that firing rates of BOS neurons increased monotonically with increasing feedback frequency. Loose synchrony between BOS neurons exhibited nonmonotonic modulation patterns (Fig. 10H). Furthermore, tight synchrony moderately and nonmonotonically increased with increasing frequency of feedback activity $\nu_{vis}$ (Fig. 10J). All these behaviors were similar to that of our original model (Figs. 4 and 6C), demonstrating the robustness of the latter.

**Effect of NMDA receptor decay time.** A large number of NMDA-type glutamate receptors have been observed, with different functional properties. Although some of these receptors have been quite well characterized in various neuronal structures, the details of the distribution of the different sub-
types in extrastriate cortex are not well known [Self et al. (2012) is a rare exception of a study that partially addressed this question in the awake behaving primate]. The specific parameters of NMDA-based receptor projections are not known, and we used a generic model of NMDA receptors (Eqs. A3–A5). To obtain at least a rough idea of the robustness of our results, we decided to vary, in addition to the synaptic strength, the time constant of conductance decay ($\tau_{\text{NMDA,decay}}$), covering the approximate range that can be expected for this type of receptor (~10–400 ms). The results are shown in Fig. 11, A and B. As in our previous simulations, visual input and G-cell input were simulated as independent Poisson spike trains of $\nu_{\text{vis}} = 200$ Hz and $\nu_{G} = 3, 25$, and 45 Hz. For all levels of $\nu_{G}$, the firing rates of BOS neurons increased with larger values of $\tau_{\text{NMDA,decay}}$ (Fig. 11A). In contrast, loose synchrony showed nonmonotonic behavior when $\tau_{\text{NMDA,decay}}$ was increased (Fig. 11B), somewhat similar to the behavior when $\nu_{G}$ was increased (Fig. 4). Increasing either the frequency of synaptic events or their postsynaptic decay time increased the postsynaptic impact of a spike train, which may explain that both manipulations showed a similar influence on the correlation functions. Interestingly, under bound conditions ($\nu_{G} = 25$ and 45 Hz), maximal synchrony was observed for $\tau_{\text{NMDA,decay}}$ in the range beyond 40 ms, well within the range where many physiological studies place typical NMDA receptor decay times.

We again plot loose synchrony as a function of $\nu_{G}$, as shown in Fig. 4B, for three values of $\tau_{\text{NMDA,decay}}$ (40, 120, and 160 ms; Fig. 11C). Again, we found nonmonotonic behavior. Thus, the nonmonotonic behavior induced by modulatory common inputs is robust for a wide range of $\tau_{\text{NMDA,decay}}$.

DISCUSSION

Neuronal synchrony from common input. It is a well-established fact that common input to a cell population generates correlations between its members (Rosenbaum and Josic 2011). Usually it is assumed in such studies that the synaptic input is driving the cells, i.e., generating action potentials on its own. An exception is a study by Yim et al. (2014), who considered nonspike-driving input, which was defined as a mixture of inhibitory and excitatory inputs that by itself was not strong enough to bring the membrane voltage over the threshold and generate a spike. The subject of the present
report is, however, different: Our interest is on modulatory input, which, in and by itself, does not change the membrane voltage at all, at least in an idealized version of the mechanism. Instead, activation of the modulatory input increases the impact that driving input has on the membrane voltage. Perhaps the best-known example of such a modulatory mechanism is mediated by NMDA-type glutamatergic receptors. Since close to the resting potential the channel is blocked by Mg\(^{2+}\), its activation does not modify the membrane voltage. Activation of AMPA receptors increases the voltage, which removes the Mg\(^{2+}\) block, which results in substantially higher net inward currents than if NMDA receptors were not activated.

The main goal of our study was to quantify to what extent common input using this modulatory mechanism leads to correlated firing in postsynaptic populations. This is of interest for understanding the nature of the neural codes by which neurons communicate and compute (e.g., “rate coding vs. temporal coding”). Another question that has been discussed for decades (e.g., Aertsen et al. 1989) is whether the structure of neuronal circuitry can be deduced from analyses of correlations between neurons (although it is by no means certain to what extent the question can be answered in a rigorous way even for driving inputs; see Jeck and Niebur 2015).

In both cases, it is necessary to understand the relationship between correlations in the input and correlations in the output.

We formulated our analysis of synchrony generated by top-down modulatory input in terms of a specific model of perceptual organization for three reasons. One reason was that we felt that a concrete example, based on circuitry likely to exist in a specific part of primate visual cortex, is more useful and intuitive than an abstract theoretical model. The second reason was that this question is of considerable interest in itself. The third reason was that neurophysiological results in this system allowed us to immediately test predictions of our model. We will therefore start with a brief discussion of the problem and its context followed by an examination of the evidence for NMDA receptors in cortical feedback circuitry before discussing results of our modeling study proper.

**Perceptual organization and selective attention.** An important step in visual perception is the structuring of the raw input, consisting of spatiotemporal activities of photo receptors and their low-level combinations into perceptual objects (“proto-objects,” Rensink 2000). Over the last decade and a half, neurophysiological work has shown that the organization of a complex scene into perceptual objects is implemented at least partially by the border ownership selectivity of individual neurons in the striate and (mainly) extrastriate cortex. As shown first by Zhou et al. (2000), BOS neurons are driven by the stimulus in their classical receptive fields, but this activity is modulated by the visual context. Specifically, a visual object border activates two different groups of BOS neurons, one of which fires at a higher rate when the object is located to one side of the border, whereas the other group fires more when the object is on the opposite side, even when the local border provides no information about location of the object. From the population activity of these neurons, widely separated pieces of border can then be assigned to a common object. According to a model by Craft et al. (2007), BOS neurons receive their selectivity through modulatory input from populations of G-cells that respond to object-like configurations and modulate the activity of BOS cells in agreement with the figure context.

In addition to providing this bias to BOS cells, the activity of G-cells then also represents the location of protoobjects in the visual scene (Russell et al. 2014).

Another important perceptual mechanism is the choice of a subset of all available information in the visual scene for more detailed processing, i.e., selective attention (Nobre and Kastner 2014). Visual attention enables the brain to reduce its computational load by allocating its resources to the most behaviorally relevant parts (Posner 1980; Carrasco 2011), a process that can be understood at least in part with quantitative computational models (Itti et al. 1998; Niebur and Koch 1998). Attention can be directed to entities at different functional levels, e.g., to different spatially defined areas of the visual field, to different features (like colors), and to different objects. In all cases, attention has been found to modulate the activity of neurons in the visual cortex (Martinez-Trujillo and Treue 2004; Motter 1993; Ogawa and Komatsu 2009; Reynolds et al. 2000; Treue and Maunsell 1999). Of particular interest to us is attention to objects, a capability clearly shown in behavioral work (e.g., Egly et al. 1994). An important advantage of the organization of a scene in perceptual objects is that it allows to select relevant parts of the scene to be attentively selected for more detailed scrutiny. The selection of a set of inputs for closer scrutiny is much more efficient if it can be made from a
relatively small number of perceptual objects than from an arbitrary subset of raw input signals.

Qiu et al. (2007) demonstrated interactions between attentional selection and border ownership selectivity in the responses of individual neurons in the early visual cortex. Perceptual organization and selective attention thus seem related functionally. Craft et al. (2007) suggested that this functional connection is implemented by specific anatomical circuitry. Specifically, they proposed that G-cells not only structure the visual scene into protoobjects but also are used to control attention to the corresponding protoobjects.

Understanding cortical circuitry from observing spike-spike correlations. To understand the mechanisms that group features into coherent objects and their interaction with selective attention, Martin and von der Heydt (2015) recorded simultaneously from pairs of neurons in the striate and extrastriate cortex of awake behaving macaques. In their experimental paradigm, a given neuron can represent a part of the border of one of several simultaneously presented objects depending on the stimulus configuration presented. An object can be located on the preferred side of a BOS neuron or on its nonpreferred side. If an object is on the preferred side of both members of a recorded neuron pair, then this is called a consistent pair (at least for this visual scene) and the neurons are assumed to be grouped together to a common representation of this object. We assume that the grouping is implemented in the activity of G-cells that provide common feedback to BOS neurons that are associated with this object (Craft et al. 2007; Fig. 1, B and C). Common feedback from the G-cell should then lead to enhanced spike-spike synchrony between consistent neurons but not between inconsistent neurons. This is precisely what Martin and von der Heydt (2015) found.

These authors also controlled the attentional state of the animals, known to modulate the mean firing rates of BOS cells (Qiu et al. 2007), and again focused on spike-spike correlations between BOS neurons. Attention to an object did not change the level of tight synchrony (Fig. 6A), and it actually decreased the level of loose synchrony (Fig. 3, C and D). This was an unexpected result since, according to the grouping hypothesis, top-down attention increases the firing rates of G-neurons and one might have expected that common feedback to BOS neurons increases synchrony between them rather than decrease it. This naive expectation fails to take into account, however, that the feedback from G-cells, as implemented in the models by Craft et al. (2007) and Mihalas et al. (2011), is modulating rather than driving, making an interpretation in terms of spike-spike correlations more difficult. In fact, since these models are rate based, they do not generate quantitative predictions for spike-spike correlations. To obtain such predictions requires the development of a spike-based model, which is one of the main contributions of the present study. NMDA projections likely play an important role in modulatory recurrent (feedback) connections in the visual cortex (experimental evidence: Johnson and Burkhalter 1994; model: Lumer et al. 1997). Based on the time course of the observed correlation functions, we follow Martin and von der Heydt (2015) in suggesting that the modulatory feedback is implemented in terms of glutamatergic projections from the NMDA receptor family.

Top-down modulatory mechanisms. Although we are not aware of neurotransmitter-specific studies of circuitry controlling border ownership selectivity, a recent report (Self et al. 2012) addressed the question of subtypes of glutamatergic projections involved in feedforward and feedback interactions resulting in the modulation of figure-ground segregation. Figure-ground segregation is distinct from border ownership selectivity, but these two mechanisms, both serving perceptual organization, share functional properties and may share neuronal circuitry. Self et al. (2012) found that feedforward activity in the primary visual cortex is driven by glutamatergic receptors of the AMPA type, whereas feedback activity responsible for figure-ground modulation relies on NMDA receptors. Another recent study (Herrero et al. 2013) showed that NMDA-based synapses are also involved in feedback projections needed for top-down attentional modulation. While other neurotransmitters are likely involved (e.g., Bentley et al. 2003; Herrero et al. 2008) and while subtypes of glutamatergic receptors may play distinct roles (Self et al. 2012), no details are known. We therefore used a somewhat simplified model in which afferent visual information is transmitted by AMPA-type synapses while modulatory feedback, serving both perceptual grouping and attentional selection, relies on NMDA-type projections.

Model assumptions. Previous grouping models (e.g., Craft et al. 2007; Mihalas et al. 2011; Russell et al. 2014) have agreed that feedback to the extrastriate cortex cells needs to be modulatory, rather than generating activity in downstream neurons by itself, to avoid spurious activity that is generated in the absence of sensory input. The same applies to attentional control signals. Those previous rate models implemented modulatory feedback through functional firing rate-enhancing mechanisms (similar to multiplying firing rates), but such a functional implementation is at odds with the level of biological realism inherent in a spiking model, which calls for specification of the mechanism in more biophysical terms. In our model network, bottom-up input drives the responses of BOS neurons through glutamatergic synapses of the AMPA type. Responses are modulated by feedback from G-cells through glutamatergic synapses of the NMDA type. The strength of this feedback, characterized entirely by the G-cell mean firing frequency, is determined both by the visual context and attentional state.

A G-cell that represents an object in the visual scene has higher activity than other G-cells that fire at low (spontaneous) rates. In existing grouping models (Craft et al. 2007; Mihalas et al. 2011), this is achieved by determining G-cell activity from the weighted sum of BOS cell inputs, and G-cell activity, in turn, provides the bias that imparts border ownership selectivity on BOS cells. Since our interest is the influence of top-down modulation on spike train correlations, we focused on the second part of this process and therefore simplified the connectivity of the network: we assumed that the feedforward input to the G-cell layer is suitably structured to activate those G-cells that correspond to a visual object rather than explicitly simulating this feedforward process. As before (Mihalas et al. 2011), we also assumed that top-down attention increases the activity of the G-cell population associated with the attended object without being concerned with the origin of the attentional control signal (see MATERIALS AND METHODS for details).

An important constraint on the model is that the system cannot rely on highly specific feedback connections. Although we made no specific assumptions about the anatomic location...
of the G-cell population, it is likely that they are not colocalized with BOS cells. They exert their influence, therefore, over white matter connections of a length of, at a minimum, several centimeters. We therefore assume that these feedback connections are diffuse and nonspecific, each effectively targeting a large area in V2. An illustration of the lack of specificity is shown in Figs. 4 and 5 of the Craft et al. (2007) study, which described the model in more detail. The projective fields of G-cells are spatially diffuse annuli, whose weight is spread around the mean distance (r) by a Gaussian with a SD of r/2.5 (see Eq. A5 in that report). As that report shows (in particular, their Fig. 5B), this spreads the synaptic weights very broadly.

Given this very broad feedback anatomical pattern, how is it possible that only neurons that show consistent border ownership preference will be grouped and driven to fire synchronously? The reason is that the feedback is not driving the cells but only modulating their activity, through NMDA channel-mediated currents. Therefore, even though a very large population of V2 neurons receives nonspecific top-down modulatory input, it will influence the responses only of those neurons that, in addition, receive (bottom-up) driving input. The latter originates in the sensory input and in our model is mediated by AMPA-type currents. As a consequence, only those neurons that respond to sensory stimulation will be observed to fire synchronously since only those neurons fire at all.

Members of consistent neuronal pairs (defined by Martin and von der Heydt 2015 and discussed above) receive top-down input from the same G-cell population, whereas neurons in inconsistent pairs receive top-down input from different G-cells. Since we assume, for simplicity, no interaction between G-cells and, furthermore, that visual input to all BOS neurons is independent, members of inconsistent pairs have no correlations beyond chance in our model. They are therefore not explicitly modeled; all results presented are for consistent pairs.

Robustness of nonmonotonic correlation functions. Loosely synchronous between BOS neurons showed nonmonotonic behavior as a function of the mean rates of G-cells (Fig. 4B). Such nonmonotonic modulation pattern of loose synchrony was also observed under a wide range of bottom-up input frequencies (νvis) and τNMDA,decay, as shown in Figs. 9C and 11C. For 5τNMDA,decay, we used the physiologically realistic range of 40−160 ms. The mean rates of afferent inputs νvis was varied over a range from 150 to 400 Hz (results for 200 Hz are shown in Fig. 4B). We also investigated the robustness of our results when we relaxed the oversimplified assumptions that feedback connections are purely NMDA driven and feedforward connections are purely AMPA driven. We found that the nonmonotonic dependence of correlation functions on the strength of the feedback is preserved when the feedback connections are dominated by NMDA but also contain modest amounts of AMPA (~20% of synaptic weights) and, likewise, when a similar amount of NMDA is mixed into the AMPA-dominated feedforward input (Fig. 10).

Dependence of synchrony on common input. We found that synchrony between spiking activity of neurons is generated not only when common input to these neurons drives them to fire but also by common input that is modulatory. Furthermore, and perhaps surprisingly, we found that the strength of synchrony at the timescale of a few tens of milliseconds (“loose synchrony”) is nonmonotonically related with the strength of the modulatory input, more specifically, that it has a marked maximum at intermediate values. In this section, we discuss the origins of this behavior.

It is easy to see that synchrony is low for small common input; in the extreme case of zero common input, synchrony between postsynaptic neurons cannot be higher than expected by chance. Increasing common input then leads naturally to increasing levels of synchrony. What remains to be explained is why synchrony decreases again for larger values of common input. There are at least two mechanisms responsible for this behavior.

The first mechanism is present in principle for all synapses but, for the (biologically realistic) firing rates in our model, it is particularly effective for synapses with slow temporal dynamics, as for NMDA-type currents. The basic observation is that, as the firing rate of G-cells becomes large, and given the slow dynamics of NMDA channels, the input to BOS cells becomes essentially a constant barrage with very little internal time structure. Such nearly constant modulatory input increases the mean firing rate, as shown in Fig. 4A, but it generates only little spike-spike correlations beyond chance. Therefore, for high νG, the correlation between BOS cells is low. This mechanism also explains the strong influence of the NMDA synaptic time constant on synchrony. Figure 11B shows this effect very clearly. Synchrony decreases once the modulation becomes nearly constant, which is the case when the product of decay time and event rate is on the order of unity. This is the case for two of the three values of νG shown: for νG = 25 Hz, the peak occurs at ~80 ms, so the product is 2; for νG = 45 Hz, the peak is at ~40 ms and the product of time constant and rate is 1.8. For the third value of νG = 3 Hz, the peak is not yet reached but the curve is clearly flattening at the highest considered time constant (400 ms) and the product will likely be in the same range as for the other two frequencies.

The second mechanism that contributes to nonmonotonic synchrony was identified by de la Rocha et al. (2007) and by Shea-Brown et al. (2008). These authors argued that spike train covariability (which is directly related to cross-correlation and synchrony) between neurons receiving common input is approximately proportional to the product of the covariance of the inputs and their gain in the target neurons. The nonmonotonic behavior is understood because, as discussed above, for low νG the presynaptic covariance is low due to the low level of common input, resulting in low postsynaptic correlation. The correlation is also low for high values of νG, since there the gain is low, as seen by the flattening of the νBOS versus νG relationship shown in Fig. 4A. The product of these two functions is therefore low for both low and high values of νG, and it has one maximum at intermediate values.

In our model, input covariability arises from the common components of I syn impinging onto two BOS neurons (see Eq. 2 in MATERIALS AND METHODS), and the gain is the change of νBOS caused by a change of νG. We therefore obtained the gain by computing numerically the derivative of the curve shown in Fig. 4A, and we multiplied it with the covariance between I syn obtained from Eq. 2. The result of this multiplication is shown in Fig. 12. The shape of this curve is similar to that showing loose synchrony in Fig. 4B.

We emphasize that, although each of the two discussed mechanisms by itself can lead to nonmonotonic synchrony as
function of common input levels, in our model both mechanisms contribute.

Limitations of the present model. Our model is certainly oversimplified. Self et al. (2012) showed that, at least in area V1, there is not one homogeneous population of NMDA receptors but, instead, at least two pharmacologically dissociable types with different neurophysiological properties that modulate firing rates during figure-ground modulation. NMDA-dependent rate modulation with changes in attention was not found by Herrero et al. (2013). Instead, those authors found that what was modulated by NMDA-type projections were temporally varying quantities, like rate variation and noise correlations (fluctuations of firing rates between trials that are reduced by attention; also observed by Cohen and Maunsell 2009; Gomez-Ramírez et al. 2014; Martin and von der Heydt 2015; Mitchell et al. 2009). It is unclear whether the discrepancy between these two studies is due to different mechanisms underlying attentional modulation and figure-ground modulation or other factors. In any case, the experimental paradigms of these studies differed from each other as well as from those in Dong et al. (2008) and Martin and von der Heydt (2015), which may explain the different results.

It is also well known that glutamate is not the only neurotransmitter that plays an important role in attentional control. Herrero et al. (2008) demonstrated the importance of acetylcholine on V1 firing rates in top-down attention. An influential model of cholinergic influences on attentional feedback was published by Deco and Thiele (2011). Our model is, in some sense, complementary to that work. While we focused on the influence on different glutamatergic pathways and used mainly spike synchrony to compare model results with experimental data, the Deco and Thiele model was concerned with cholinergic interactions and only considered firing rates.

Cortical distance and intra-areal connections. As an alternative to the proposed grouping mechanism, which requires an interaction between BOS neurons and grouping cells interacting via white matter connections, some computational models (discussed in the next section) have been proposed, which assume that BOS cells communicate with each other by intraareal “horizontal” connections. We believe that such a mechanism is very unlikely to be in use, due to the large distances in the extrastriate cortex between neurons that represent different components of stimuli. In physiological experiments by Martin and von der Heydt (2015), cortical distances between the recorded BOS neurons of consistent pairs extended 3–14 mm. Angelucci et al. (2002) reported that the mean length of horizontal fibers in monkey V1 (we are not aware of any such measurements for V2) is 6 ± 0.7 mm, with a maximum length of 9 mm. In fact, 9 mm is the largest extent ever reported in the literature for cortical horizontal fibers in macaque V1 (Angelucci and Bullier 2003). The longest distance between the consistent pairs of BOS neurons in the physiological study (Martin and von der Heydt 2015) is thus more than twice that of the mean length of horizontal fibers in V1, and it is also substantially higher than the longest horizontal fibers ever reported. It is therefore impossible that selectivity of border ownership is implemented by direct connections mediated through horizontal fibers.

The only way horizontal connections could influence synchrony would be by indirect connections in several “hops,” but that would be a completely different and more complicated model. These connections would have to follow along the contours of the stimuli since shorter connections over neurons that represent the interior of the visual stimuli (where no stimulus is presented) would lead to neurons firing in response to stimuli located far away from their receptive fields, which is, by definition, impossible. An immediate consequence of this contour-following mechanism is that the time between responses to the first stimulus should be proportional to the linear size of the stimulus. However, Sugihara et al. (2011) found no corresponding increase in the latency of border ownership selectivity even when nearly tripling the stimulus size (from 3° to 8°). We therefore conclude that horizontal connections in early cortical areas cannot explain the observed synchrony structure between BOS neurons and are unlikely to form the substrate of figure-ground segregation.

Comparison with previous models. Several models have been proposed to understand attentional modulation in BOS neurons. Wagatsuma et al. (2008, 2013) assumed that attentional modulation of BOS neurons in the extrastriate cortex as well as behaviorally observed switching of figure perception by attention is caused by mechanisms located in early visual areas. In addition, border ownership has been proposed to be implemented entirely by mechanisms within V1 or V2, as in models by Sakai and Nishimura (2006) and Zhaoping (2005). Based on the fast emergence of border ownership selectivity, which would be difficult to explain by propagation along slow horizontal connections (see also the previous section), another class of models assumes that border ownership selectivity arises from interactions through fast white matter projects. Craft et al. (2007) and Qiu et al. (2007) introduced the concept that grouping by specific neuronal populations (G-cells) is responsible for generating a bias that leads to border ownership selectivity in the extrastriate cortex and that the same circuitry might also be used for attentional selection. Craft et al. (2007) implemented grouping in a computational model, whereas Mihalas et al. (2011) implemented the top-down attention mechanism. Russell et al. (2014) focused on the representation of object-based saliency in the population activity of G-cells. All of these models (as well as ours) assume that grouping is implemented by modulating activity of BOS neurons rather than directly driving it. One question we address in this report is which biophysical mechanism generates this modulation. We propose that this mechanism is a projection based on
NMRE-type glutamatergic synapses, whereas the bottom-up purely visually driven input is glutamatergic with AMPA-type synapses.

Our proposed model is only concerned with correlations between consistent pairs of BOS neurons, i.e., those that receive projections from the same G-cell population (Figs. 1 and 2). In the model introduced by Craft et al. (2007), feedback from G-cells generated border ownership selectivity through disinhibition of BOS neurons. In their model, consistent pairs of BOS neurons receive common input from G-cells opposing the cell we modeled (see their Fig. 3), and all synchrony has to be explained by the inhibitory inputs. In our model, however, G-cells modulate BOS neurons by direct synaptic projections, not through disinhibition (Fig. 2). Furthermore, and again different from the Craft et al. model, we do not implement mutual inhibition between opposing BOS neurons. In fact, even if mutual inhibition from other BOS neurons (as in the Craft et al. model) were added to our model, it would not make a systematic contribution to the described correlation functions because the activity of these inhibitory neurons is not correlated with that of the studied neurons (they receive their input ultimately from other G-cells, corresponding to opposite border ownership relationship). Therefore, mutual inhibition would not contribute to synchrony in our model.

There are, of course, many ways to model the same data, and we do not claim that NMDA-based feedback is the only possible way. We showed that simply implementing the previously postulated modulatory feedback (Craft et al. 2007) with NMDA synapses produces spike time cross-correlations that closely resemble the experimentally observed ones and also produces the observed paradoxical decrease in correlation with attention (Martin and von der Heydt 2015). Whether other models can achieve the same remains to be seen.

Applicability to other systems. We applied a model of interacting modulatory and driving inputs to neuronal populations developed in this report to understanding how a structured representation of the environment in terms of perceptual objects is formed from raw visual input and how the mechanisms that accomplish this feat are interfaced with the top-down influences of selective attention. However, the structure of our model is much more general and not limited to the early visual cortex. The motif of afferent driving input, which is shaped by efferent modulatory influences, is found in many perceptual systems, and similar schemes may exist also in cognitive and motor structures. The architecture that we study in this report (Fig. 2) should thus be applicable far beyond the specific question that we have considered here, and our conclusions should apply mutatis mutandi to a much larger class of neuronal systems. In particular, it was an a priori unexpected observation that common input can decrease, rather than increase, neuronal correlations. It will be interesting to see if this phenomenon, which is well understood in the framework of our model, is present in brain areas other than the visual cortex.

APPENDIX

Synaptic currents flowing into BOS neurons. BOS neurons receive input from two external sources: bottom-up input representing visual stimuli and top-down input from G-cells (Eqs. 1 and 2). Bottom-up excitatory postsynaptic currents \( I_{\text{vis}} \) are mediated by glutamatergic receptors of the AMPA type (Buehlmann and Deco 2008; Deco and Thiele 2011) and can be defined as follows:

\[
I_{\text{vis}}(t) = g_{\text{AMP}}{{V(t) - V_E}}_{\text{BOS}}w_{\text{BOS}}s_{\text{AMP}}(t) \quad (A1)
\]

where \( V_E = 0 \text{ mV} \) is the reversal potential and \( w_{\text{BOS}} \) represents the excitatory synaptic weight from visual stimuli to BOS neurons, chosen as \( w_{\text{BOS}} = 140 \). \( V \) is the subthreshold membrane potential of a BOS neuron (see also Eq. 1). The conductance of the fully activated synapse is \( g_{\text{AMP}} = 0.104 \text{ nS} \), and the fraction of open channels of AMPA receptors \( (s_{\text{AMP}}) \) can determined as follows:

\[
\frac{ds_{\text{AMP}}(t)}{dt} = \frac{s_{\text{AMP}}(t)}{\tau_{\text{AMP}}} + \sum k \delta(t - t_{i_{\text{vis}}}) \quad (A2)
\]

where the postsynaptic decay time constant is \( \tau_{\text{AMP}} = 2.0 \text{ ms} \). The sum over \( k \) runs over all spikes originating from edge-selective neurons responding to the visual stimuli. Each spike is entered as a Dirac delta function, \( \delta(t) \), which assumes a nonzero value at the spike times of the visually driven input neurons \( (t_{i_{\text{vis}}}) \), zero elsewhere, and has an integral of unity over any interval that includes \( t_{i_{\text{vis}}} \).

G-cells provide common modulatory inputs for BOSL and BOSR neurons (Figs. 1C and 2). A fundamental hypothesis of our model is that the modulatory feedback takes the form of currents through NMDA receptors. All NMDA receptors have a voltage dependence that is controlled by \( [Mg^{2+}] \) (Jahr and Stevens 1990), which we assume as \( [Mg^{2+}] = 1 \text{ mM} \). Here, we will use a standard computational model for generic NMDA receptors (Wang 1999) in which the NMDA receptor mediated synaptic current \( I_{\text{G}} \) is defined as follows:

\[
I_{\text{G}}(t) = \frac{g_{\text{NMDA}}[V(t) - V_E]}{1 + [Mg^{2+}]^3}w_{\text{BOS}}^{\text{vis}}s_{\text{NMDA}}(t) \quad (A3)
\]

where \( V_0 = 16.13 \text{ mV} \) (all voltages are in mV) and \( w_{\text{BOS}}^{\text{vis}} = 140 \) is a parameter symbolizing the synaptic weight from G- to BOS neurons. The synaptic conductance of a fully open NMDA synapse is \( g_{\text{NMDA}} = 0.327 \text{ nS} \); however, at resting voltage \((-70 \text{ mV})\), voltage-dependent \( Mg^{2+}\)
The tightened, jitter-corrected correlogram (CCG*) was found by subtracting the mean of the \( r \) jittered correlograms, \( \langle J^* \rangle_r \), for the amount of overlap, as follows:

\[
\text{CCG}^*(\tau) = \langle C^*(\tau) \rangle_r - \langle J^*(\tau) \rangle_r
\]

We also computed the integral of the tight synchrony between BOSL and BOSR neurons (Eq. A9) in the range of \( \pm 5 \) ms, as follows:

\[
M^* = \sum_{r=5}^{\infty} \text{CCG}^*(\tau) \times \text{bin size}
\]

M*, defined by Eq. A9, implied the index as the magnitude of the tight synchrony. In a manner similar to regular loose synchrony, the spike trains have bin size of 1 ms.

Experimental studies often subtract a shuffle predictor from the cross-correlogram to compensate for systematic rate changes during each trial (e.g., due to the stimulus onset). Since the input spike trains to the BOS neurons were stationary, we did not use shuffle correction. We also show that doing so does not change our results (compare Fig. 3E with Fig. A1).

ACKNOWLEDGMENTS

The authors thank Anne Martin and Brian Hu for valuable discussions.

GRANTS

This work was partly supported by Office of Naval Resources-Multidisciplinary University Research Initiative Grant N000141010278 and National Institute on Drug Abuse Grant R01-DA-040990, and KAKENHI Grant 26888019.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

N.W. and E.N. conception and design of research; N.W. performed experiments; N.W. analyzed data; N.W., R.v.d.H., and E.N. interpreted results of experiments; N.W. and R.v.d.H. prepared figures; N.W. drafted manuscript; N.W., R.v.d.H., and E.N. edited and revised manuscript; N.W., R.v.d.H., and E.N. approved final version of manuscript.

REFERENCES


Qiu FT, Sugihara T, von der Heydt R. Figure-ground mechanisms provide structure for selective attention. Nat Neurosci 10: 1492–1499, 2007.


