# Intracellular Measurements of Spatial Integration and the MAX Operation in Complex Cells of the Cat Primary Visual Cortex

**Ilan Lampl,[1,3] David Ferster,[3] Tomaso Poggio,[2] and Maximilian Riesenhuber[2,4]**

[1]*The Weizmann Institute of Science, Department of Neurobiology, Rehovot, 76100 Israel;* [2]*Massachusetts Institute of Technology, Department of Brain and Cognitive Sciences, Center for Biological and Computational Learning, McGovern Institute for Brain Research, and Artificial Intelligence Laboratory, Cambridge, Massachusetts 02142;* [3]*Northwestern University, Department of Neurobiology and Physiology, Evanston, Illinois 60208; and* [4]*Georgetown University Medical Center, Department of Neuroscience, Washington, DC 20007*

**Lampl, Ilan, David Ferster, Tomaso Poggio, and Maximilian Riesenhuber.** Intracellular measurements of spatial integration and the MAX operation in complex cells of the cat primary visual cortex. *J Neurophysiol* 92: 2704–2713, 2004; 10.1152/jn.00060.2004. We have examined the spatial integration properties of complex cells to determine whether some of their responses can be described by a maximum operation (MAX)-like computation, as suggested by Riesenhuber and Poggio's model of object recognition. Membrane potential was recorded from anesthetized cats while optimally oriented bars were presented, either alone or in pairs, in different parts of the cells' receptive field. In most cells, the membrane potential response to two bars presented simultaneously could not be predicted by the sum of the responses to individual bars. In many cells, however, the responses closely approximated a MAX-like model. That is, the response of the cell to two bars was similar to the larger of the two individual responses ("soft-MAX"). The degree of nonlinear summation varied from cell to cell and varied within single cells from one stimulus configuration to another but on average fit most closely to the MAX model. The firing response of the cells was also well predicted by the MAX-like model. The MAX-like behavior was independent of the distance between the bars (orthogonal to the preferred orientation), independent of the relative amplitude of the responses, and slightly less pronounced at low levels of contrast. This MAX-like behavior of a subset of complex cells may play an important role in invariant object recognition in clutter.

## INTRODUCTION

A model of object recognition in cortex proposed by Riesenhuber and Poggio (1999b) consists of a hierarchy of neurons performing one of two operations on their afferent inputs: a weighted combination leading to multidimensional tuning and a maximum operation (MAX). According to the model, the former operation serves to increase selectivity by building more complex feature detectors from simpler ones, whereas the MAX operation serves to increase response invariance to translation and scaling by pooling over afferents tuned to the same feature at different locations and sizes. The MAX function is also crucial in generating robustness to clutter (Riesenhuber and Poggio 1999a). Like Hubel and Wiesel's (1962) hierarchical model of simple and complex cells, this model of object recognition is based on feedforward processing.

The MAX-like operation in the model can be formalized mathematically as an operation that returns the largest of its inputs. Thus a neuron that performs a MAX operation on its pooled inputs will respond to the strongest of its inputs only and will not be affected by other weaker inputs (hence the model's robustness to clutter or distracting stimuli). Note that while the model was developed using this ideal form of the MAX operation, invariance properties of model units have also been shown to be robust for the case of a more graded nonlinearity ("soft-MAX") (Riesenhuber and Poggio 1999b). The presence of MAX-like behavior is tested experimentally by comparing the responses to two stimuli presented alone and simultaneously. A neuron that exhibits a MAX-like behavior will give a response to the pair of stimuli that is similar to the larger of the two individual responses. Such a protocol was used to demonstrate MAX behavior in a subset of neurons in inferotemporal cortex (Sato 1989) and more recently in some V4 neurons (Gawne and Martin 2002). The model predicts, however, that a MAX-like operation is already performed in the earliest stage of the cortical visual pathway by a subset of complex cells receiving input from simple cells in V1 as a first step to build invariance to stimulus translation and possibly to stimulus scaling. Previous extracellular recording studies have shown that spatial summation of complex cells is sublinear (Henry et al. 1978; Movshon et al. 1978). These studies were not motivated by a computational theory of object recognition, however, and did not specifically test the MAX hypothesis.

To test whether the response of some complex cells can be described by a MAX-like operation, we recorded the membrane potential of complex cells in anesthetized cats while presenting single bars or pairs of bars. We found that the subthreshold behavior of most of the complex cells we tested could be well described by a MAX-like response function. Our results also have implications for the mechanism underlying MAX-like behavior in cortical cells in that the graded synaptic potentials by themselves, prior to filtering by the spike-generating mechanism, can perform the MAX operation.

## METHODS

Experiments were performed on young adult female cats (2–3 kg). All procedures were approved by the Northwestern University Animal Care and Use Committee. Anesthesia was induced with ketamine (5–15 mg/kg) and acepromazine (0.7 mg/kg) and maintained intravenously with sodium thiopental (20–30 mg/kg initial dose; 1–2 mg · kg$^{-1}$ · h$^{-1}$ maintenance). Animals were paralyzed with gallamine triethiodide (10 mg/kg initial; 10 mg · kg$^{-1}$ · h$^{-1}$ maintenance) or

vecuronium bromide (Norcuron, 0.21 mg · kg$^{-1}$ · h$^{-1}$) and artificially respirated. To reduce mechanical artifacts, the animal's body was suspended by the stereotaxic frame using a clamp attached to the spine. Body temperature was kept at 38.2°C with a thermostatically controlled heat lamp. The electrocardiogram, electroencephalogram, end-tidal CO$_2$, autonomic signs and rectal temperature were continuously monitored to ensure the anesthetic and physiological state of the animal. Intracellular recordings from area 17 (lateral 1 to 2, AP −9 to −5) were obtained using sharp electrodes (40–100 MΩ) filled with 2 M potassium acetate as previously described (Lampl et al. 2001). Membrane potential was amplified using an Axoclamp-2A amplifier (Axon Instruments, Foster City, CA) and low-pass filtered before being digitized at 4 kHz and stored.

Cells were hyperpolarized by current injection to counterbalance the inward leak of current that occurs as a result of the electrode penetration. The current level was chosen to set the firing rate of the cells within or below the range of values normally observed in extracellular recording. Using this approach, the average number of spikes per stimulus trial was approximately two or less. Larger amounts of current could have been used to suppress spiking completely, but this would have put the cells at an unnatural resting potential, which in turn might have distorted the way in which synaptic current summated. Because the nature of this summation was the subject of the study, we wanted to keep the resting potential at more physiological levels.

Spikes were detected using a threshold of 10 mV after applying a high-pass filter (with a cutoff of 200 Hz) to the digitized membrane potential. Firing rate was determined by counting the number of spikes in each response. Prior to measuring the trajectory of the graded membrane potentials, spikes were removed from the traces by interpolation. The potential for a period of 4 ms, starting 2 ms prior to each spike, was replaced with the value measured at the start of the period. The results of this interpolation procedure were nearly identical to the results of using a median filter to remove spikes (Ferster and Jagadeesh 1991).

Visual stimuli were generated on a Macintosh computer using the Psychophysics Toolbox (Pelli 1997) running under Matlab (The Mathworks, Natick, MA) environment and presented on a 17-in Viewsonic monitor (100-Hz refresh rate, mean luminance: 20 cd/m$^2$, 1,024 × 768 resolution) positioned 40 cm away from the cat's eyes. The eyes were focused on the screen with a combination of contact lenses (with 3 mm artificial pupils) and auxiliary lenses. Focus was measured by projecting the image on the retina onto the monitor screen with a fiber optic illuminator. Orientation tuning curves were obtained on-line using drifting gratings of 12 different orientations presented several times in pseudorandom order and were used to find the preferred orientation of the cell. Receptive fields were located within 10° of the area centralis. Receptive fields in the dominant eye were monocularly mapped using bright and dark bars of 50-ms duration, flashed repeatedly in pseudorandom order at the cell's preferred orientation (see Fig. 1). Following the criteria of Hubel and Wiesel (1962), the degree of overlap of ON and OFF responses was used to distinguish simple and complex cells. Additional method of classification was obtained by measuring the membrane potential response of the cells to drifting grating. Cells were classified as complex cells if the mean elevation was greater than half of the modulation (potential modulation index <0.5) (Carandini and Ferster 2000). Most of the cells that were classified as complex cells, based on receptive field geometry, also passed the second classification method [12 of 16 cells in which the orientation tuning curve was successfully fitted with a double Gaussian function (Carandini and Ferster 2000)]. Contrast is given by Eq. 1 where $L_s$ and $L_b$ are the luminance of the bar and the background, respectively. The contrast, C, used to map the receptive field and to test spatial summation was +90% and −90% unless specified otherwise, where

$$C = 100*(L_s - L_b)/L_b \qquad (1)$$

Data are presented as means ± SE.

## RESULTS

To study the computation that complex cells perform on their inputs, we stimulated each cell with bars of its preferred orientation positioned at different points within the receptive field. Bars were flashed alone or in pairs. From the responses to the presentation of each bar alone, two different predictions were made for response to the paired stimuli: Linear summation, and MAX-like pooling—where the larger of the two responses was used to predict the response (Riesenhuber and Poggio 1999b). These predictions were then compared with the response of the cell to the two bars flashed simultaneously.

After encountering a cell, the receptive field was first mapped with a sparse noise stimulus (Lampl et al. 2001). The cell in Fig. 1 was mapped with light and dark bar stimuli made from a 12 × 3 grid that covered an area of 5 × 5°. The responses to each bar are shown in Fig. 1A, and a map of the receptive field derived from these responses in Fig. 1B, where color represents response amplitude at a given point in the visual field. Because the ON and OFF responses overlapped nearly completely, we classified this cell as complex (see METHODS). To study how the cell responded to the simultaneous presentation of two stimuli, a subset of stimuli was selected from the map to span a range of response amplitudes and a range of interstimulus spacings. Only a subset was used to test summation (rather than the entire set of mapping stimuli) so that each stimulus and stimulus pair could be presented repeatedly in the time allowed by intracellular recordings. In Fig. 1, four bright bars and one dark bar were selected from the responsive area of the map (black outlines, Fig. 1B). The selected stimuli were then flashed alone and in all possible pairwise combinations. Thus for five selected bars, 15 possible configurations (5 singles and 10 pair combinations) were presented. The entire sequence of stimuli was presented repeatedly in pseudorandom order (100-ms stimulus duration and 150-ms interstimulus interval), with different randomization each time. Membrane potential responses were averaged for each stimulus after removing action potentials from the traces using an interpolation procedure (see METHODS). The resulting averaged responses are presented in Fig. 1C. The *left column* and *top row* show the responses to individual stimuli ($R_a$, $R_b$, thin black traces), together with the SD of the responses (surrounding gray areas). The responses to two stimuli presented simultaneously are presented in the inner panels ($R_{a+b}$, gray traces), together with the individual response to the stimuli that made up the pair (thin black traces), and the arithmetic sum of the individual responses ($R_a + R_b$, thick black traces). For this cell, most of the responses to the paired stimuli were smaller than the sum of the two individual responses. Furthermore, in most cases, the response to the paired stimulus was similar in its amplitude and trajectory to the larger of the two individual responses. Thus the cell's response could be characterized as approximating the MAX operation.

To quantify the responses and compare their amplitudes to the linear and MAX predictions, we measured the average potential between 50 and 150 ms after stimulus onset, relative
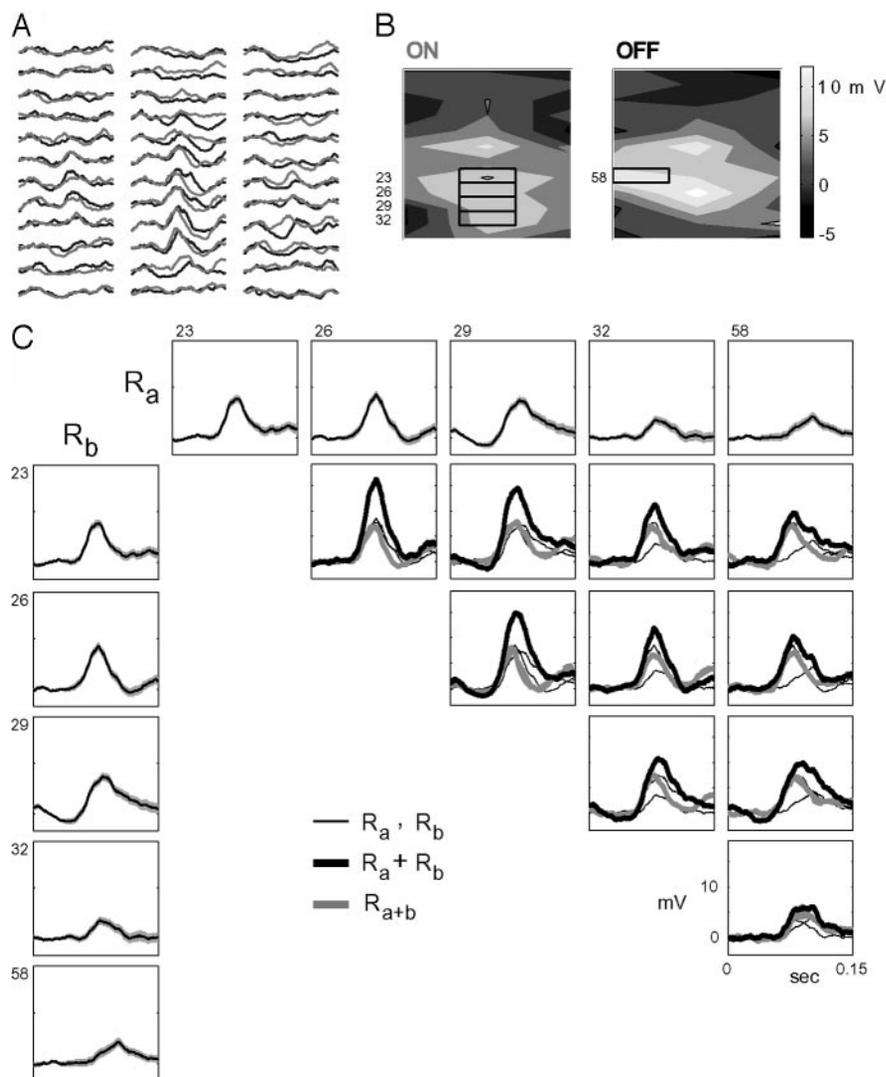
FIG. 1. The response of a complex cell to the simultaneous presentation of 2 bars. *A*: average membrane potential measured from the response of the cell to bars of the optimal orientation each flashed for 100 ms aligned to the onset of stimulation. The mapped area was $5 \times 5°$, sampled with a $12 \times 3$ grid. Black traces are the responses to dark bars (OFF responses) and gray traces are the responses to bright bars (ON responses). *B*: intensity plots obtained from the mean potentials between 50 and 100 ms post stimulus onset. *C*: the responses of the cell to each of the selected bars shown in *B* by thick lines around the rectangles. Lines in the 1st row and 1st column panels are the averaged responses to the presentation of a single bar, and the shaded area shows the mean $\pm$ SE. The inner panels present the response of the cell to the simultaneous presentation of the 2 bars whose positions are given by the corresponding column and row (gray traces), the responses to the 2 stimuli presented individually (thin black traces) and the linear sum of the 2 individual responses (thick black traces).

to the baseline or resting voltage, taken as the average potential during the 30 ms just prior to stimulus onset. Amplitude measurements from six representative cells are compared with the linear and MAX predictions for each of the tested stimulus pairs in Fig. 2; *A* shows the measurements from the cell presented in Fig. 1; *B–F* show five other cells. Responses are ordered from left to right according to their MAX index (see following text). In most cases, the linear prediction (*) exceeded the actual response of the cells (△). Close examination of most of the data points indicates that many of the responses to the paired stimuli were similar to the larger of the two individual responses (○) as expected if the cells pooled their inputs in a MAX-like operation. Not all the combined stimuli evoked responses similar to the larger of the two responses, however; some variability in the behavior of the cells was observed. In a few cases, the combined response was closer in amplitude to the smaller of the two responses (pairs 1 and 2 in *C, D,* and *E*), and in others, it was closer to the arithmetic sum of the two (pair 10 in *D,* pair 6 in *F*).

The entire data set of 215 responses to two stimuli flashed together obtained from 21 cells is shown in Fig. 3. The predictions from a linear model (*A*) and from a MAX-like model (*B*) were plotted against the actual responses to two bars

flashed simultaneously. The figure shows only responses for which reliable amplitude measurements could be obtained; nonsignificant responses ($P > 0.05$) were excluded using a Student's *t*-test. Note that for the linear model (Fig. 3*A*), most of the points are located under the diagonal line, indicating that the model overestimates the response of the cells to combined stimuli. The MAX model, on the other hand (Fig. 3*B*), shows much better prediction: many of the response of the cells to two stimuli flashed together are close to the diagonal, indicating a much better fit. Yet, as expected from the variability of the individual cells (Fig. 2), some scattering can be observed.

To measure more quantitatively how well the models predicted the actual response of the cells to the paired stimuli, we performed a normalized residual analysis, calculating *se(i),* the normalized squared error of the measured responses relative to the each model for each tested stimulus pair, i.e.

$$se(i) = [(R_p(i) - R_m(i))/R_m(i)]^2 \qquad (2)$$

where $R_p$ is the predicted response for one of the models and $R_m$ the measured response. The error for the MAX model is plotted against the error for the linear model in Fig. 3*C*. For 80% of the data points, the prediction of the MAX model yielded lower residuals compared with the linear model. Only
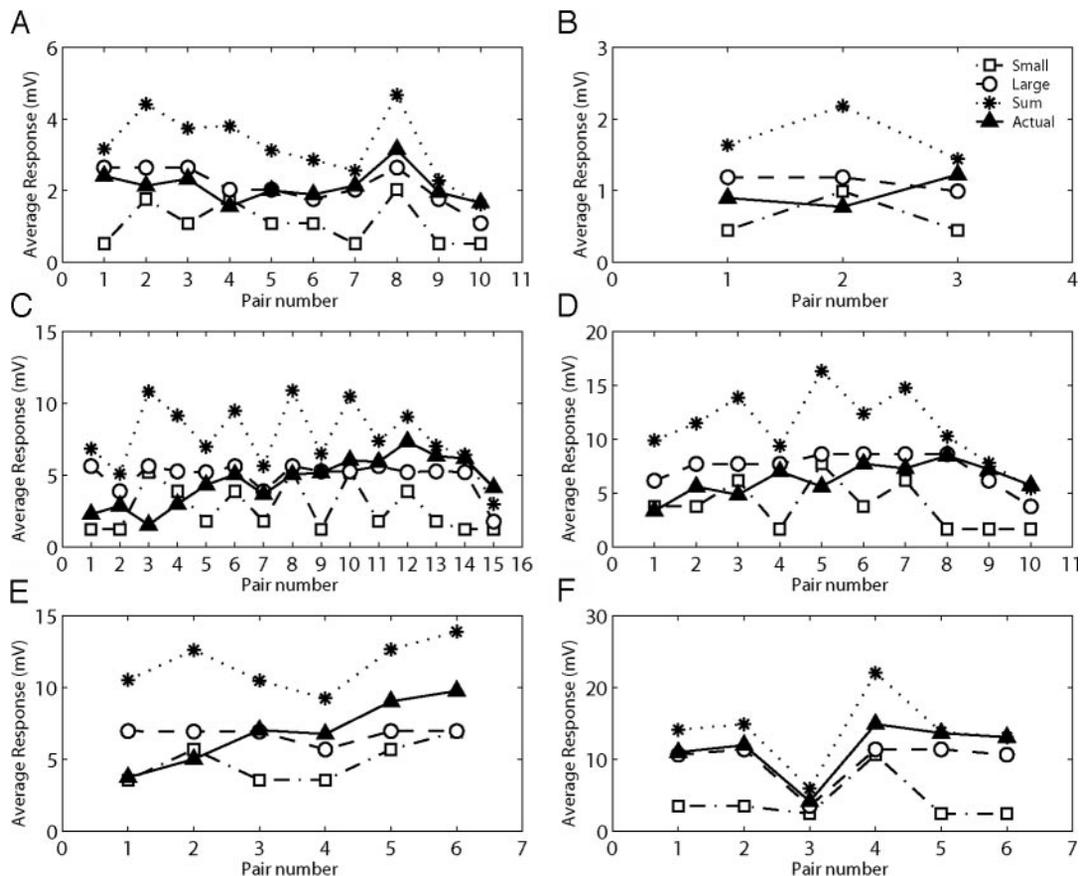
FIG. 2. Mean responses of 6 cells to single bars and bar pairs. Responses shown are the mean potential between 50 and 100 ms after stimulus onset. *A*: for this cell (shown in Fig. 1), 5 bars were presented alone or in combination with another bar, yielding 10 possible bar pairs. In addition to the actual response of the cell to each 1 of the bars (□ - · - □, smaller of the 2 individual responses; ○ - - - ○ larger response) and the combination (actual; △ — △), we plotted the predicted response from a linear summation (sum, * · · · *, see legend in *B*). *B—F*: result from 5 other cells. For each cell, different numbers of bars were used. For each cell, the pairs are ordered in the plot by increasing MAX index (from left to right, see *Eq. 3* in text). The average MAX indices for these 6 cells are 0.07, −0.18, −0.11, −0.13, −0.02, 0.46.

when the error values for both models were very small did the linear prediction show a slightly better fit than the MAX function (Fig. 3*C, inset*).

One possible mechanism that might underlie sublinear summation of synaptic potentials could be saturation of the membrane potential at threshold. That is, the membrane potential could be clamped at threshold by the occurrence of spikes and the opening of voltage-gated ion channels. Saturation would tend to generate MAX-like summation of inputs, especially in the case where one stimulus triggered a subthreshold response and the other triggered a suprathreshold response. To test whether saturation was a contributing factor in generating MAX-like behavior, we calculated the MAX index on the responses in which no spikes were present (Fig. 3, *D–F*). Note that selecting traces without spikes slightly lowers the average size of the responses, but it does not eliminate all large responses because the triggering of a spike depends on d*V*/d*t* as well as on *V* itself. As for the complete data set, the traces without spikes also showed much lower residual error for the MAX model than for the linear model.

These calculations of which operation (linear or MAX) was performed by the cells are derived entirely from subthreshold modulations of the membrane potential. In most of our recordings, we hyperpolarized the resting potential using injection of

a steady current to minimize firing (METHODS). In 14 cells with sufficient firing (on average ≥1 spikes/flash), we also compared the MAX and linear models based on the firing rates evoked by the stimuli. For each stimulus, we counted the number of spikes in the same window that was used for the membrane potential analysis (from 50 to 150 ms). As for membrane potential, the average firing rate of the cells was better predicted by the MAX-like model than by the linear model (Fig. 3, *G–I*). However, there was larger scatter in the spike-rate data than the membrane potential data, possibly due to the low firing rate of the cells in our experiments, together with the short time window for measuring spike rate. More precise measurements for spikes will require more stimulus trials.

MAX-like behavior can be quantified using the following index (Sato 1989)

$$I = (R_{a+b} - \text{MAX}(R_a, R_b))/\text{MIN}(R_a, R_b) \qquad (3)$$

where $R_{a+b}$ is the response to the two bars flashed simultaneously, and $R_a$ and $R_b$ are the responses of the cell to each one of the two stimuli presented alone (all positive). A perfect MAX operation produces an index of 0, whereas perfect linear summation will yield an index of 1. Negative index values indicate suppressive interaction (that is, the response to the

## Membrane potential (all trials)



## Membrane potential (excluding trials with spikes)
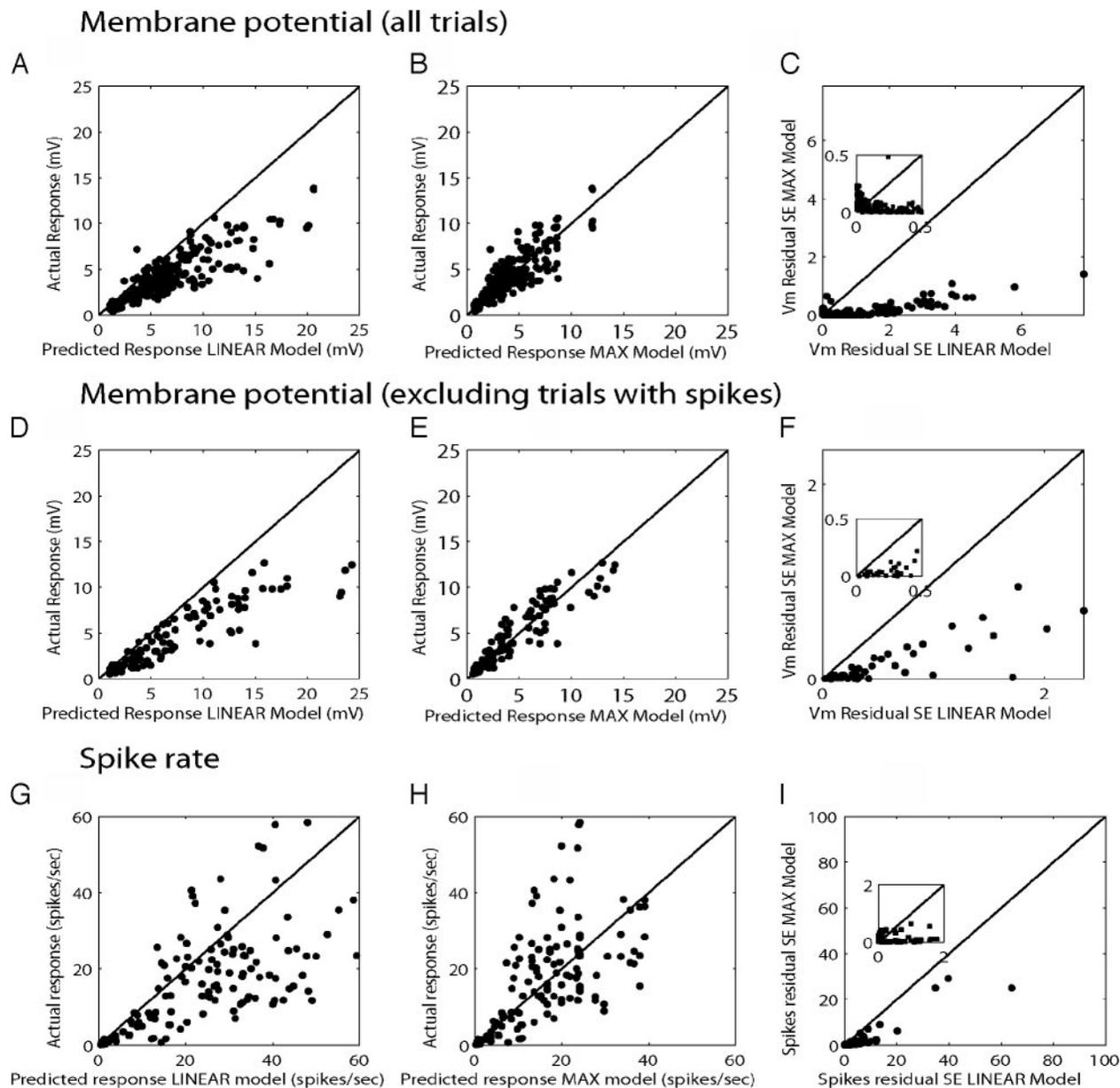


## Spike rate



FIG. 3. Failure of the linear model and success of the MAX model in predicting the responses of complex cells for the membrane potential and for firing rate. Predictions made from linear (*A*) and maximum operation (MAX; *B*) models. Plotted are the actual amplitudes of 215 averaged responses to double bar presentations pooled from 21 cells. *C*: normalized residuals for the MAX model plotted on the vertical axis and residuals of the linear model on the horizontal axis. *Inset*: a blow-up of the very small values in *C*. Similar results were obtained when we calculated the mean firing responses. *D–F*: similar measurements done after removing any trials in which spikes were evoked. Analyzed from 107 combinations that were pooled from 13 cells of the cells shown in *A–C*. *G–I*: firing rate was analyzed from 138 combinations, pooled from 14 cells of all 21 cells shown in *A–C*.

stimulus pair is smaller than the response to either stimulus presented alone). Figure 4*A* shows the distribution of MAX integration indices for membrane potential for all stimulus pairs tested. Indices calculated from the full data set (134 stimulus pairs) are shown as gray bars; indices calculated from those traces in which no spikes occurred (sufficient data were available for 107 stimulus pairs) are shown as black bars. Figure 4*B* shows the distribution of the indices for each recorded cell obtained by averaging the indices obtained from all pairs presented to a given cell. The mean index value was close to 0 in each case: 0.11 ± 0.55 for full data set (Fig. 4*A*, gray bars); 0.10 ± 0.51 for traces without spikes (Fig. 4*A*,

black bars); 0.11 ± 0.23 after averaging the data for each cell (Fig. 4*B*, numbers are means ± SD). These results suggest that at least some of the complex cells integrate spatially distributed inputs in a MAX-like fashion. The distribution of the MAX index for firing rate is presented in Fig. 4*C*. The mean value (−0.01 ± 0.82) indicates that a MAX operation was performed not only at the subthreshold level but also at the level of spike output of the cells. To compare the subthreshold performance to the suprathreshold performance, we plotted the measured index for spikes against the index of the membrane potential (for all the significant subthreshold responses in which firing was induced, Fig. 4*D*). The relation between the MAX indices
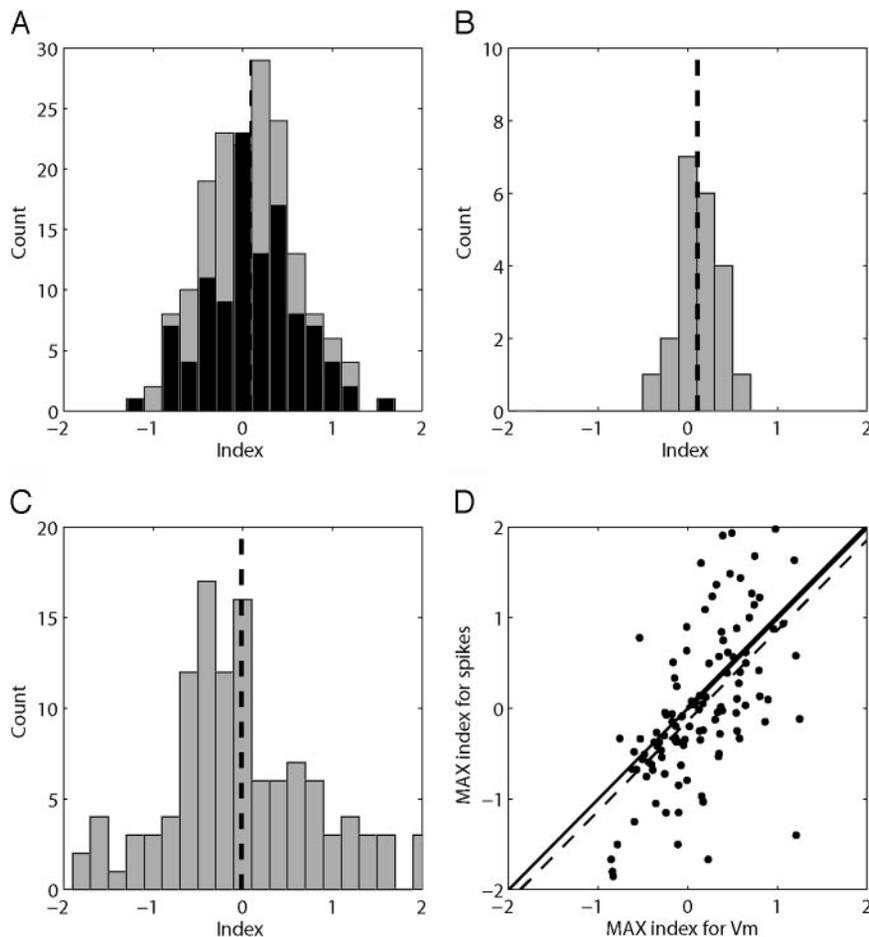
FIG. 4. The distribution of the MAX index is centered close to 0, suggesting that cells respond in a MAX-like fashion. *A*: distribution of the MAX index (gray bars, see text) of 210 responses pooled over 21 cells. The average value ($0.11 \pm 0.55$, mean $\pm$ SD, $n = 215$) is shown by - - -. (Five of the total of all 215 significant responses are not shown because of their extreme values.) The distribution of responses for trials in which no spikes were evoked (black bars) is similar to the entire set of data ($0.10 \pm 0.51$, 107 responses, 13 cells). *B*: the distribution of index values averaged separately for each cell. (Average $0.11 \pm 0.23$, $n = 21$). *C*: the distribution of MAX index calculated from mean firing rate for each combination of pair of stimuli, calculated from 114 combinations measured in 14 cells ($-0.01 \pm 0.82$). *D*: MAX index for firing rate is plotted as function of MAX index for membrane potential for the same data shown in *C*. - - -, the linear fit ($y = 99x - 0.14$, $R^2 = 0.33$). Significant difference was found between the 2 indices ($P = 0.02$, paired Student's *t*-test). Means $\pm$ SE are $0.13 \pm 0.05$ and $-0.01 \pm 0.08$ for the potential index and spikes.

was close to the diagonal (the slope was 1.01 and the intercept was $-0.14$) even though there was significant scatter due to the high variability of the spike rate caused by the low number of spikes emitted by the hyperpolarized cells. The index for firing rate was slightly smaller than for the membrane potential. However, the positive and close to linear correlation between the two indices suggests that the output operation of the cells reflects the subthreshold operation.

We explored several factors that could have an effect on the integration properties of the cells including the effect of response amplitude, spatial separation between stimuli, and contrast. The response magnitude, for example, could have an effect on the way in which the cell integrates its inputs: Small responses could be integrated more linearly than larger responses because they are farther from the reversal potential of the excitatory synapses and might activate fewer voltage-sensitive currents. We therefore plotted the normalized residuals, as shown in Fig. 3*C* and *Eq. 2,* against the amplitude of the response to the presentation of stimulus pairs (Fig. 5). To compensate for differences in absolute amplitude across cells (caused, for example, by differences in input resistance), we normalized the response amplitudes of each cell by the largest response of the cell to any two-bar combination. As shown in Fig. 3, the residuals from the linear model were larger than those obtained from the MAX model. The trend toward larger errors in the linear model for larger normalized responses (Fig. 5*A*) is consistent with the sublinear integration of inputs as observed in Fig. 3*A*. The MAX model, on other hand, fits

consistently well across all normalized responses (Fig. 5*B*). Amplitude-independent MAX-like operation is also evident in the individual cells presented in Fig. 2. The lack of correlation
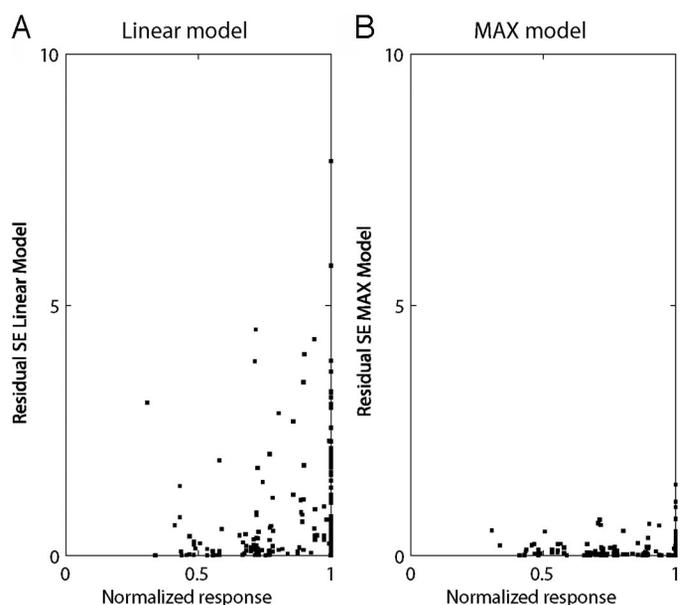


FIG. 5. The errors of the MAX model were not dependent on the amplitude of the response. Residual error of the linear model (*A*) and the MAX model (*B*) against the normalized amplitudes (normalized by the largest response for each cell) of the response to 2 bars flashed together.

between the errors of the MAX model to the amplitude suggests that the MAX operation is independent of amplitude.

Response amplitude does not correlate with the degree of MAX-like behavior across the population of cells. We also tested whether response amplitude within a single cell affected behavior as well by measuring the responses of nine complex cells at two different contrast levels. The lower-contrast stimuli evoke smaller responses, which are therefore assured to be nonsaturating (whether the larger responses saturate or not). An example cell is shown in Fig. 6. We first obtained the response of the cell to different combinations of two bars at 90% contrast (Fig. 6A). We then repeated the measurements with lower stimulus contrast (30%), which reduced the response of the cell by ~50% (Fig. 6B). At 90% contrast, the cell summed its inputs sublinearly, similar to what is predicted for MAX-like integration. At low contrast, with the reduced response amplitudes, the summation was still sublinear, and some of the combined stimuli responses could be predicted from a MAX-like model. When averaged across all stimulus pairs, for this cell, the MAX index at high contrast was −0.38,



FIG. 7. Comparison of MAX indices at different contrast levels. The MAX index (*Eq. 3*) is slightly lower at high contrast compared with low contrast. The MAX index at high contrast of 100% was plotted against the indices at a lower contrast (20–30%) for which the response was about half the size of the higher contrast.

whereas at the lower contrast, it was −0.33. The negative index values indicate that the actual response of this cell to the combined stimulation was slightly smaller than expected from a MAX operation. We tested the effect of reducing contrast in 33 pairs of stimuli taken from nine cells (Fig. 7). The mean index at high contrast ($0.02 \pm 0.05$) was different from the mean index at low contrast ($0.20 \pm 0.10$). However, the effect that was only borderline significant ($P = 0.049$, Student's *t*-test). The slightly higher index at low contrast suggests that the cells became more linear when the responses were smaller as would be predicted by a saturation mechanism. Even at the low contrast, however, the average index was much closer to 0 (pure MAX) than to 1 (linear integration), indicating that saturation alone could not account for the MAX-like behavior.

Finally, we tested the effect of spatial separation of the paired stimuli on the operation performed by the cells. We observed a large variability of receptive field sizes ($7.3 \pm 7.7°^2$, mean $\pm$ SD). Because in most cases we kept the ratio of bar width to receptive field width similar (usually between 1:4 and 1:7), we measured the separation between stimuli in units of number of stimulus widths (in the direction perpendicular to the cell's preferred orientation). We found no clear correlation between bar separation and MAX index, for either all stimulus pairs (Fig. 8A) or pairs of stimuli of opposite polarity, that is, with one dark stimulus and one bright stimulus (Fig. 8B).

### DISCUSSION

A growing body of evidence suggests that some neurons of the visual system perform a MAX-like operation on the visual image. That is, when two visual stimuli are simultaneously presented in different parts of the receptive fields, the cell's response is equal to the larger of the responses to the individual stimuli. (Sato 1989) has shown that when neurons of the inferior temporal cortex are stimulated by two bars, their responses could be approximated by a MAX operation. Recently it was found that a significant number of neurons in V4 perform a MAX-like operation when presented with complex stimuli flashed simultaneously in different parts of their receptive fields (Gawne and Martin 2002). In this study, we show that a MAX-like operation is also performed by a subset of



A

Contrast of 90%

Ra , Rb
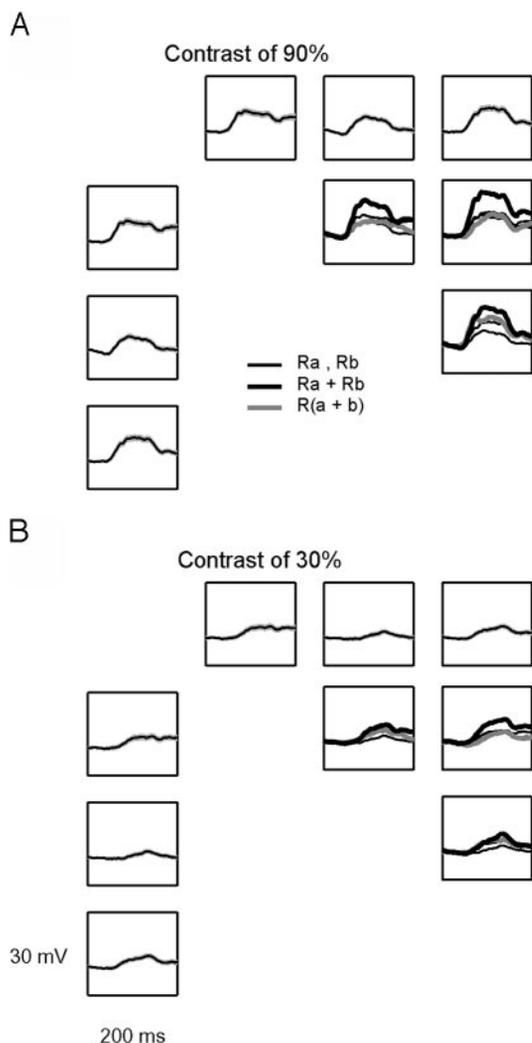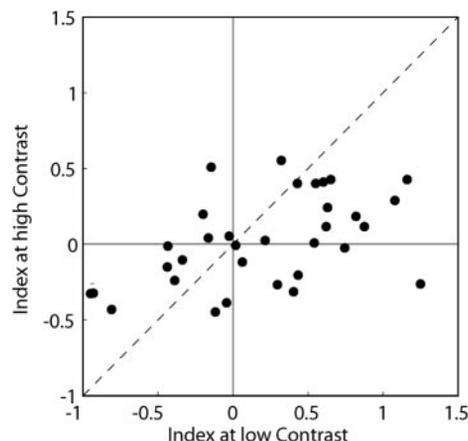Ra + Rb
R(a + b)

B

Contrast of 30%

30 mV

200 ms

FIG. 6. The effect of contrast reduction on the response of a complex cell. Reducing the contrast from 90 to 30% effectively reduced the response of the cell. The response of the cell, however, at low contrast remained sublinear and the responses closely matched the expected response from the MAX model. The average response of a complex cell to 2 bars at 100% contrast (*A*) and at 30% contrast (*B*). For more details about the presentation, see Fig. 1.
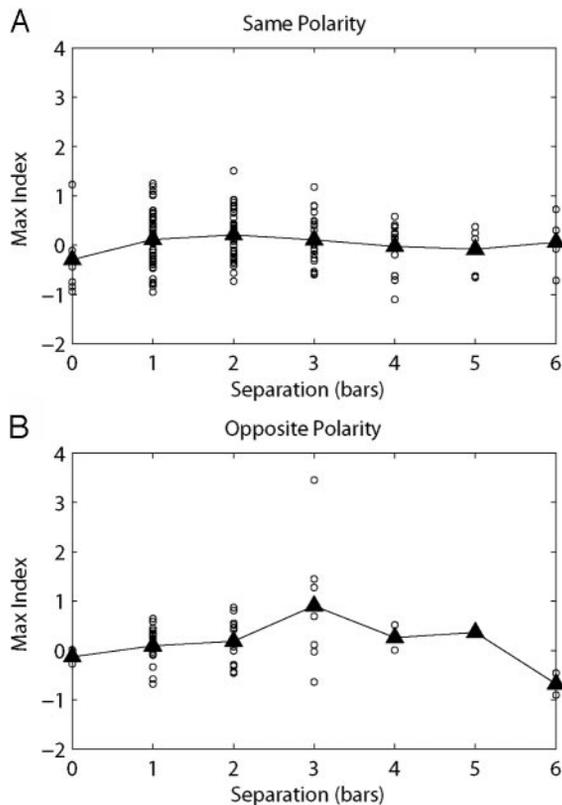
FIG. 8. MAX index plotted against separation between the 2 stimuli in a pair (measured in bar widths orthogonal to the preferred orientation of the cell) for bar pairs with the same polarity (*A, n* = 214) and for pairs in which 1 bar was dark and 1 bright (*B, n* = 50).

complex cells in the primary visual cortex. We recorded the membrane potential of the cells and found that the membrane potential and the firing rate responses to the presentation of two bars were better predicted in most complex cells by a MAX-like pooling model compared with a linear summation model. Together, these studies are consistent with Riesenhuber and Poggio's model, which uses MAX-like pooling of afferents in some neurons in the ventral visual processing stream to increase response invariance and robustness to clutter.

Spatial pooling or summation has been studied intensively in neurons of the primary visual cortex. In the feedforward model of Hubel and Wiesel (Hubel and Wiesel 1962), the superimposed ON and OFF receptive fields of complex cells are created by spatial summation of inputs from different simple cells (Alonso and Martinez 1998; Anzai et al. 1999; Martinez and Alonso 2001; Movshon et al. 1978; Szulborski and Palmer 1990). Hubel and Wiesel, however, did not quantify the properties of spatial summation in complex cells or in simple cells. Several subsequent studies (Baker 2001; Emerson et al. 1987; Szulborski and Palmer 1990) have used random noise stimulus ensembles to estimate the second-order Wiener kernel of complex cells, finding kernels consisting of several elongated subregions of opposite polarity, similar to the receptive fields of simple cells. The MAX pooling operation is consistent with these kinds of kernels, as Sakai and Tanaka (2000) have shown. While a number of other complex cell models such as the Energy Model (Adelson and Bergen 1985) are also consistent with the measured kernels, unlike the MAX model they

do not predict the predominantly MAX-like interaction or the narrow range of MAX indices found in our study (Fig. 9).

Other studies have used two-bar displays in an attempt to distinguish between simple and complex cells with respect to their spatial summation properties (Henry et al. 1978) or to study the interactions between the subunit inputs to the cells (Movshon et al. 1978). Both studies used approaches similar to ours, and in both studies, the linear prediction exceeded the actual response of the cell to two simultaneously presented bars. Furthermore, similar to our study, Henry et al. (1978) recorded the firing activity of complex cells and found that the responses to two bars closely resembled the more vigorous of the responses to either one of the bars presented alone.

Movshon et al. (1978) observed two types of interactions between bars flashed simultaneously in complex receptive fields. When two bars of the same polarity were presented nearly adjacent to one another, the response was greater than either of the individual responses but less than their sum (MAX index >0 but significantly <1). When bars of opposite polarity
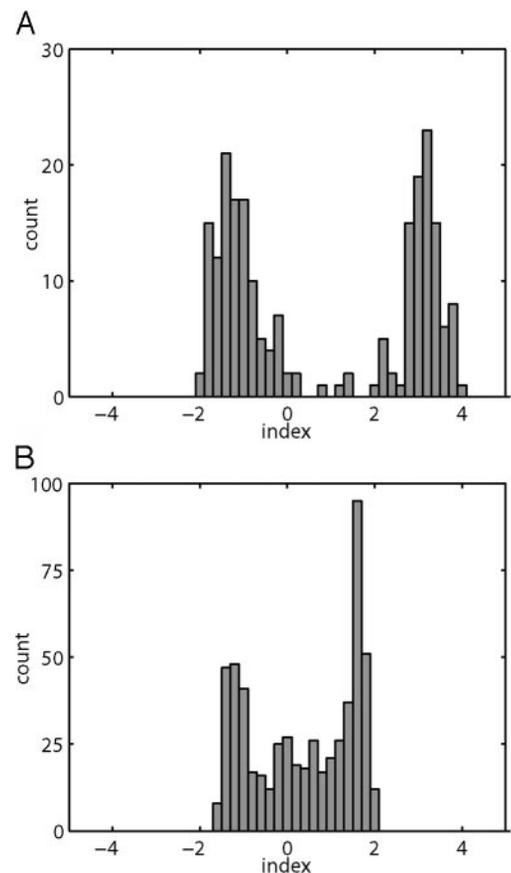


FIG. 9. Prediction of the MAX index distribution by the Energy Model of complex cells (Adelson and Bergen 1985). *A:* the index distribution obtained with a standard implementation of the Energy Model in which the outputs of 2 simple cells (modeled as Gabor functions, $\exp(-x^2/2*172) \cdot \cos(2\pi \cdot 3x/100)$ and $\exp(-x^2/2*172) \cdot \sin(2\pi \cdot 3x/100)$, on a 100 pixel-wide model retina) are squared individually and then summed by the model complex cell, along with additive uniform noise of amplitude 0.05. Stimuli were individual bars (1 pixel wide) or pairs of bars separated by ≥15 pixels. Responses were averaged over 4,000 trials. Polarities of bars were chosen randomly for each trial. Only trials in which each bar evoked a response of ≥0.4 were included. *B:* same as *A* but with the output of the simple cells raised to a power of 1.35, which increases the agreement of simulation results with the experimental data. Compare with the experimental index distribution in Fig. 4*A*. The Energy Model captures neither the unimodal shape of the experimental distribution nor its range.

were presented nearly adjacent to one another, the response was less than either of the individual responses (MAX index <0). When the bars were presented farther apart from one another, the opposite behavior was observed with MAX indices <0 for the same polarity and >0 for the opposite polarity. Effects like these are expected from the feedforward model of Hubel and Wiesel in which complex cells receive input from multiple simple cells. Closely spaced pairs of bars will fall into the same subfield of a presynaptic simple cell. When of the same polarity, they will facilitate one another; when of the opposite polarity, they will antagonize one another, independent of the pooling mechanism used by the complex cell. More widely space bars will fall into different subfields of a simple cell afferent and so will facilitate one another when of opposite polarity and antagonize one another when of the same polarity. Finally, for even bigger separations, if the two stimuli fall into the receptive fields of different simple cells that both provide input to the same complex cell performing a MAX pooling operation over its afferents, a pure MAX interaction might be observed. Given that each stimulus will activate multiple presynaptic simple cells, the exact details of bar interactions at the level of the complex cell thus depend not just on the pooling mechanism but also on the interactions among simple cells. In the absence of more precise information about an individual complex cell's afferents, one would thus expect based on the model to find a range of MAX indices.

The full range of suppressive and facilitatory effects that were described by Movshon et al. (1978) were also present in our study—in many cells both suppression (MAX integration indices smaller than 0) and enhancement (indices larger then 0) were found as apparent from the scatter of points above and below the diagonal in Fig. 3B and in the histograms in Fig. 4. And we, like Movshon et al., found that the enhancement of the response to the conditioning bar evoked by the test bar is most often much smaller than the response to the test bar itself (MAX index significantly smaller than 1). In contrast to Movshon et al., we did not find a clear correlation between the degree of enhancement or suppression and the separation of the bars or bar polarity (compare our Fig. 8 with Fig. 8 of Movshon et al. 1978). Figure 8, however, pools data from many different cells with different preferred spatial frequencies, each of which was tested at only a few separation distances. These and other differences in method might blur any underlying relationship to bar polarity and separation and account for some of the difference in observed behavior. Further experiments are needed, however, to resolve this issue.

*Mechanisms underlying the MAX operation in complex cells*

A number of mechanisms could contribute to the MAX-like behavior of complex cells. It could be, for example, that their synaptic inputs might already sum their inputs in a MAX-like manner such that the complex cells simply inherit this property from those inputs. It appears unlikely, however, that the MAX-like operation is performed at the level of simple cells. While Henry et al. (1978) have shown that width summation of simple cells is sublinear (which can be explained probably by antagonistic interactions between the ON and OFF subfields), the average receptive field size of complex cells in our sampled population of cells ($7.3 \pm 1.7°^2$, mean $\pm$ SE, $n = 19$) was substantially larger ($P < 0.05$, Student's $t$-test) than the aver-

age receptive field area of simple cells sampled from our database ($2.5 \pm 0.5°^2$, $n = 13$, unpublished). Therefore it is likely that bars positioned far apart from each other were likely to stimulate different simple cell inputs. In addition, our preliminary experiments on simple cells (using pairs of spot stimuli) suggest that simple cells sum their inputs in a much more linear fashion than do complex cells.

A simple way that MAX behavior could be achieved is if under some circumstances active conductances related to spiking would clamp the membrane potential at or near threshold. In that case, as long as the first stimulus in a pair brought the membrane potential close to threshold, the additional excitation generated by the second stimulus in the pair would not be able to raise the potential any further. Our results argue against such a mechanism. First, the largest responses for most of the cells were not necessarily the ones that behaved in the most MAX-like fashion. That is, they were not the ones for which the MAX index was closest to zero (pure MAX, see Fig. 2), suggesting that the cell could be depolarized further. Second, although input integration was slightly less MAX-like at low contrast compared with high contrast, it was still very nonlinear and the MAX index was still much lower than would be expected for linear summation. Third, MAX behavior was observed even in cells that had been hyperpolarized to the point that firing was low or absent. Fourth, MAX-like behavior was observed even when the analysis was limited to those traces in which no spikes were present Figs. 3, *D–F* and 4*A*, black bars).

Previous work has shown that threshold does not always clamp the membrane potential. In simple cells, for example, high-contrast drifting gratings evoke sinusoidal modulations of the membrane potential, the peaks of which rise significantly above threshold, even as spikes at a frequency of ≥50 Hz rise off the suprathreshold portion (M. Carandini and D. Ferster, unpublished data). It appears that as soon as each spike is over, the membrane potential immediately takes up a potential dictated by the balance of visually evoked synaptic excitation and inhibition, even if that potential is above threshold. Such behavior could occur if the point of spike initiation is some distance from the soma (in the 1st node of Ranvier, for example) and therefore at a different potential than the soma (Stuart et al. 1997). Whatever the case, it seems unlikely that simple voltage-dependent saturation can underlie the MAX-like behavior of complex cells.

A third way in which the MAX operation could be implemented is by arranging for each bar stimulus to evoke a simultaneous increase in excitatory and inhibitory input. Imagine, for example, that bar 1 of a pair activated excitatory and inhibitory conductances $g_e^1$ and $g_i^1$ in a proportion that gave a combined reversal potential 10 mV above rest; similarly bar 2 would activate combined excitatory and inhibitory conductances $g_e^2$ and $g_i^2$ with a combined reversal potential of 5 mV above rest. The reversal potential of the response to both bars presented together will depend on the relative amplitudes of two sets of conductances. If $g_e^1$ and $g_i^1$ are far larger than $g_e^2$ and $g_i^2$, then they will dominate the response to the pair of bars and the response will resemble the response to bar 1 (10 mV). Conversely, if $g_e^2$ and $g_i^2$ are far larger than $g_e^1$ and $g_i^1$, then they will dominate and the response to the pair will resemble the response to bar 2 (5 mV). Given that the larger of the two responses most often dominate in complex cells, then for this scenario to work, the larger of the two responses will most

often have to be associated with the larger conductances. Whether this is true or not in cortex can be tested by directly measuring conductances associated with the responses to flashed bars and determining their relative sizes and whether they are strong enough to shunt the membrane as effectively as do responses to moving stimuli (Borg-Graham et al. 1998).

Finally, a MAX operation could be implemented by more complex network connections, very similar to microcircuits implementing a gain control function. In fact, appropriate nonlinearities in circuits proposed for gain control (Carandini et al. 1997) would make them perform a MAX-like operation. Physiologically plausible network models to approximate a MAX-like behavior have been suggested by Yu et al. (2002). Among the models that were suggested by Yu et al. are feedforward and feedback models that use shunting inhibition and a feedforward model in which inhibition is used to linearly reduce the responses from other inputs. It is likely, however, that any network operation generating the MAX operation does not depend primarily on feedback from the recorded cell itself since in most of our recording we prevented the firing of the cells by current injection.

It is possible that under low-contrast conditions the integration mechanisms and the input sources of the cell are different compared with those at high-contrast conditions. In V1 neurons of the monkey, it was demonstrated that at low-contrast conditions, length and width summation are different compared with at high contrast (Sceniak et al. 1999). It is also possible that differences in integration properties at low compared with high contrasts are related to the well-known stimulus-dependent level of saturation of V1 neurons that was suggested to arise from a contrast normalization mechanism (Carandini and Heeger 1994; Heeger 1992). Some neurons showed little change in MAX index with changing contrast, whereas others exhibited greater variability. Future studies will have to investigate whether these play different functional roles in perception. If, for example, the nonlinearity of the pooling operation is affected by contrast in a significant number of cells, it would predict an effect of stimulus contrast on invariant object recognition and its robustness to clutter (Riesenhuber and Poggio 1999a).

The mechanisms that underlie receptive field properties of complex cells and their possible role in higher perceptual functions are poorly understood. Complex cells likely perform a variety of functions in vision. The evidence that some of them may perform a MAX-like operation on their inputs supports the model that predicted a key role for such a transfer function in translation- and scale-invariant object recognition, thereby linking a cellular mechanism to cognitive behavior.

REFERENCES

**Adelson EH and Bergen JR.** Spatiotemporal energy models for the perception of motion. *J Opt Soc Am* 2: 284–299, 1985.

**Alonso JM and Martinez LM.** Functional connectivity between simple cells and complex cells in cat striate cortex. *Nat Neurosci* 1: 395–403, 1998.

**Anderson JS, Lampl I, Gillespie DC, and Ferster D.** Membrane potential and conductance changes underlying length tuning of cells in cat primary visual cortex. *J Neurosci* 21: 2104–2112, 2001.

**Anzai A, Ohzawa I, and Freeman RD.** Neural mechanisms for processing binocular information. II. Complex cells. *J Neurophysiol* 82: 909–924, 1999.

**Baker CL Jr.** Linear filtering and nonlinear interactions in direction-selective visual cortex neurons: a noise correlation analysis. *Vis Neurosci* 18: 465–485, 2001.

**Borg-Graham LJ, Monier C, and Fregnac Y.** Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature* 393: 369–373, 1998.

**Carandini M and Ferster D.** Membrane potential and firing rate in cat primary visual cortex. *J Neurosci* 20: 470–484, 2000.

**Carandini M and Heeger DJ.** Summation and division by neurons in primate visual cortex. *Science* 264: 1333–1336, 1994.

**Carandini M, Heeger DJ, and Movshon JA.** Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci* 17: 8621–8644, 1997.

**Dreher B.** Hypercomplex cells in the cat's striate cortex. *Invest Ophthalmol* 11: 355–356, 1972.

**Emerson RC, Citron MC, Vaughn WJ, and Klein SA.** Nonlinear directionally selective subunits in complex cells of cat striate cortex. *J Neurophysiol* 58: 33–65, 1987.

**Ferster D and Jagadeesh B.** An *in vivo* whole cell patch study of the linearity of IPSP-EPSP interactions in cat visual cortex. *Soc Neurosci Abstr* 17: 176, 1991.

**Gawne TJ and Martin JM.** Responses of primate visual cortical V4 neurons to simultaneously presented stimuli. *J Neurophysiol* 88: 1128–1135, 2002.

**Gilbert CD.** Laminar differences in receptive field properties of cells in cat primary visual cortex. *J Physiol* 268: 391–421, 1977.

**Heeger DJ.** Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9: 181–197, 1992.

**Henry GH, Goodwin AW, and Bishop PO.** Spatial summation of responses in receptive fields of single cells in cat striate cortex. *Exp Brain Res* 32: 245–266, 1978.

**Hubel DH and Wiesel TN.** Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160: 106–154, 1962.

**Kato H, Bishop PO, and Orban GA.** Hypercomplex and simple/complex cell classifications in cat striate cortex. *J Neurophysiol* 41: 1071–1095, 1978.

**Lampl I, Anderson JS, Gillespie DC, and Ferster D.** Prediction of orientation selectivity from receptive field architecture in simple cells of cat visual cortex. *Neuron* 30: 263–274, 2001.

**Martinez LM and Alonso JM.** Construction of complex receptive fields in cat primary visual cortex. *Neuron* 32: 515–525, 2001.

**Movshon JA, Thompson ID, and Tolhurst DJ.** Receptive field organization of complex cells in the cat's striate cortex. *J Physiol* 283: 79–99, 1978.

**Pelli DG.** The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10: 437–442, 1997.

**Riesenhuber M and Poggio T.** Are cortical models really bound by the "binding problem"? *Neuron* 24: 87–93, 111–125, 1999a.

**Riesenhuber M and Poggio T.** Hierarchical models of object recognition in cortex. *Nat Neurosci* 2: 1019–1025, 1999b.

**Rose D.** Responses of single units in cat visual cortex to moving bars of light as a function of bar length. *J Physiol* 271: 1–23, 1977.

**Sakai K and Tanaka S.** Spatial pooling in the second-order spatial structure of cortical complex cells. *Vision Res* 40: 855–871, 2000.

**Sato T.** Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. *Exp Brain Res* 77: 23–30, 1989.

**Sceniak MP, Ringach DL, Hawken MJ, and Shapley R.** Contrast's effect on spatial summation by macaque V1 neurons. *Nat Neurosci* 2: 733–739, 1999.

**Stuart G, Spruston N, Sakmann B, and Hausser M.** Action potential initiation and backpropagation in neurons of the mammalian CNS. *Trends Neurosci* 20: 125–131, 1997.

**Szulborski RG and Palmer LA.** The two-dimensional spatial structure of nonlinear subunits in the receptive fields of complex cells. *Vision Res* 30: 249–254, 1990.

**Yu AJ, Giese MA, and Poggio TA.** Biophysically plausible implementations of the maximum operation. *Neurol Comput* 14: 2857–2881, 2002.