Dynamics of Tuning in the Fourier Domain

Brian J. Malone and Dario L. Ringach
Department of Neurobiology and Psychology, David Geffen School of Medicine, University of California, Los Angeles, California

Submitted 15 February 2008; accepted in final form 8 May 2008

Malone BJ, Ringach DL. Dynamics of tuning in the Fourier domain. J Neurophysiol 100: 239–248, 2008. First published May 14, 2008; doi:10.1152/jn.90273.2008. Neurons in primary visual cortex (area V1) are jointly tuned to the orientation and spatial frequency of sinusoidal stimuli (the Fourier domain). The role that suppressive mechanisms play in shaping the tuning and dynamics of cortical responses remains the subject of debate. Here we used subspace reverse correlation to study the relationship between suppression by nonoptimal stimuli, the spectral-temporal separability of the responses, and their persistence in time. Two clear relationships emerged from our data. First, cells with inseparable responses were often accompanied by suppression to nonpreferred stimuli, while separable responses showed mostly enhancement by their preferred stimuli. Second, inseparable responses were characterized by a longer persistence in time compared with those with separable dynamics. A parametric model that assumes the additive combination of separable enhancement and suppression signals, with suppression constrained to be low-pass in spatial frequency and untuned for orientation, explained the data well. These new findings, in addition to an established correlation between selectivity and suppression for nonoptimal stimuli, clarify how the dynamics and selectivity of cortical responses are shaped by suppressive signals and how their interplay generates the large diversity of responses observed in primary visual cortex.

INTRODUCTION

Neurons in primary visual cortex are tuned to the orientation and spatial frequency of sinusoidal stimuli (K. K. De Valois et al. 1979; R. L. De Valois et al. 1982a; Jones et al. 1987; Movshon et al. 1978; Webster and De Valois 1985). How this selectivity arises in V1 from thalamic inputs that are untuned for orientation and low-pass in spatial frequency, and what role cortical inhibition plays in shaping the tuning of cortical cells, remains a central question in visual neuroscience (Ferster and Miller 2000; Priebe and Ferster 2008; Shapley et al. 2003; Sompolinsky and Shapley 1997).

It has long been known that nonoptimal stimuli can suppress firing rates in cortical neurons below their spontaneous levels (Bauman and Bonds 1991; Blakemore and Tobin 1972; De Valois and Tootell 1983; De Valois et al. 1982b; DeAngelis et al. 1992; Jones et al. 2001; Monier et al. 2003; Morrone et al. 1982; Nelson and Frost 1978; Sato et al. 1996; Sillito 1975). It has also been proposed that suppression by nonoptimal stimuli is important in establishing sharp neuronal selectivity (Bauman and Bonds 1991; Bonds 1989; Ringach et al. 2002; Shapley et al. 2003; Sillito et al. 1980). Detecting suppressive effects in extracellular recordings in response to single drifting sinusoidal gratings is difficult due to low spontaneous rates in V1, but can be readily unmasked by elevating the activity of cells pharmacologically or via a conditioning visual stimulus. Another strategy to tease apart suppressive effects has been to look at the dynamics of tuning with the idea that differential dynamics of enhancement and suppression could lead to inseparable behavior that could unveil the underlying components of the response. We adopted this strategy in previous studies that separately measured the dynamics of orientation and spatial frequency tuning (Bredfeldt and Ringach 2000; Ringach et al. 1997, 2003; Shapley et al. 2003). In the orientation domain, our results suggested that global and tuned suppression could generate a narrowing of the orientation tuning bandwidth, create Mexican-hat tuning profiles in the late parts of the response, and induce small but detectable shifts in the preferred orientation of neurons (Ringach et al. 1997, 2003). In the spatial frequency domain, suppression at low spatial frequencies correlated with a shift in the preferred spatial frequencies from low to high and an accompanying decrease of their bandwidth (Bredfeldt and Ringach 2000). In all these cases, suppression generates a temporal inseparability in the tuning curve of the cells.

To gain further insight about how suppression by nonoptimal stimuli affects V1 responses, we measured the dynamics of tuning in the Fourier domain, defined jointly by orientation and spatial frequency (Jones et al. 1987; Ringach et al. 2002; Webster and De Valois 1985). Our goals were to characterize the prevalence of suppression by nonoptimal stimuli in V1 neurons and the degree of spectro-temporal separability of their tuning curves and to test the notion that temporal inseparability and suppression may be correlated with each other.

We found a substantial diversity in the responses: some neurons have temporally separable spectral kernels (Mazer et al. 2002), but a sizeable fraction exhibit inseparable dynamics. We also find that the extent to which the kernel is dominated by a response of a single sign (i.e., net enhancement or net suppression relative to the baseline response) correlates with the degree to which its tuning in the Fourier domain is temporally separable, as well as its persistence in time (i.e., the length of the response tail). The observed dynamics in the Fourier domain are well accounted by a simple model that linearly combines two signals, enhancement and suppression, each of which is separable on its own.

These new findings add to an already established correlation between selectivity and suppression for nonoptimal stimuli (Ringach et al. 2002) and indicate that both the dynamics and the tuning selectivity of cortical responses are shaped by suppressive signals.
METHODS

Preparation, recording, and optics

All experiments were approved by the Chancellor’s Animal Research Committee at UCLA and were carried out following National Institutes of Health’s Guidelines for the Care and Use of Mammals in Neuroscience. Acute experiments are performed on anesthetized and paralyzed adult Old-World monkeys (Macaca fascicularis). Initially, the animal is sedated with acepromazine (30–60 mg/kg), anesthetized with ketamine (5–20 mg/kg im), and transported to the surgical suite. Initial surgery and preparation are performed under isoflurane (1.5–2.5%). Two intravenous lines are put in place. A urethral catheter is inserted to collect and monitor urine output, and an endotracheal tube is inserted to allow for artificial respiration. All surgical cut-down sites are infused with local anesthetic [lidocaine (Xylocaine), 2%, sc]. Pupils are dilated with ophthalmic atropine and custom-made gas-permeable contact lenses are fitted to protect the corneas. After this initial surgery the animal is transferred to a stereotaxic frame. At this point anesthesia is switched to a combination of sufentanil (0.15 μg·kg⁻¹·h⁻¹) and propofol (2–6 mg·kg⁻¹·h⁻¹). We then proceed to perform a craniotomy over primary visual cortex. The animal is maintained on physiological solution at a rate to maintain a fluid volume of 5–10 ml/h to ensure adequate hydration. Drugs are administered in balanced physiological solution at a rate to maintain a fluid volume of 5–10 ml·kg⁻¹·h⁻¹. Reticular pressure is maintained by a self-regulating heating pad at 37.5°C. Expired CO₂ is maintained within 4.5 and 5.5% by adjusting the stroke volume and ventilation rate. The maximal pressure developed during the respiratory cycle is monitored to ensure that there is no incremental blocking of the airway. A broad-spectrum antibiotic [penicillin G (Bicillin), 50,000 IU/kg] and anti-inflammatory steroid (dexamethasone, 0.5 mg/kg) are given at the beginning of the experiment and every other day.

In some experiments, a 10 × 10 electrode array (Cyberkinetics, Salt Lake City, UT) with 1- or 1.5-mm-long electrodes was implanted in primary visual cortex. The center of the array was aimed at 6 mm posterior to the lunate sulcus and 8 mm lateral to the midline. In other experiments, extracellular action potentials were recorded with an array of independently movable glass-coated tungsten microelectrodes with exposed tips of 5–15 μm (Alpha-Omega Engineering). Electrical signals were amplified and spikes discriminated using a Cerebus 128-channel system (Cyberkinetics). Spike sorting was performed off-line using principle-component analysis on the waveform shapes with software developed in our laboratory.

Stimuli were generated on a Silicon Graphics O2 and displayed on a monitor at refresh rate of 100 Hz and a typical screen distance of 80 cm. The mean luminance was 60 cd/m². A Photo Research Model 703-PC spectro-radiometer was used for calibration. The eyes were initially refracted by direct ophthalmoscopy to bring the retinal image into focus for a stimulus roughly 80 cm from the eyes. Once neural responses were isolated, we measured spatial frequency tuning curves and maximized the response at high spatial frequencies by changing external lenses in steps of 0.25 D. This procedure was performed independently for both eyes.

Kernel estimation, data selection, and model fitting

We measured the dynamics of tuning in the joint spatial frequency and orientation plane using the subspace reverse correlation technique employed by Ringach et al. (2002). The visual stimulus is a sequence of luminance modulated sine wave gratings varying in orientation, spatial frequency, and spatial phase, presented in pseudorandom order at an effective rate of 50 Hz (each frame is repeated twice with a monitor refresh rate of 100 Hz). Each image in the stimulus set consisted of a Hartley basis function of size $M \times M$ pixels, such that $H_{r,k}(l,m) = \text{cas}(2\pi (k/l + k/m))M$ for all $0 \leq l, m \leq M - 1$. Here, $\text{cas} \theta = \sin \theta + \cos \theta$ and $k_r$ and $k_m$ represent the spatial frequency of the gratings in units of cycles per stimulus side. The stimulus set consisted of all Hartley basis functions $S = \{H_{r,k}\}$ such that $|k_r| \leq k_{\text{max}}$ and $|k_m| \leq k_{\text{max}}$ (with the origin $H_{0,0}$ excluded). The maximum spatial frequency for the Hartley basis set was chosen to exceed the range of spatial frequencies (SFs) that elicited responses during initial characterizations with drifting gratings. In the Fourier plane, stimuli with the highest SFs lying along the perimeter of the stimulus domain (the edges bounding the kernel) can be used to estimate a baseline firing rate for the cell for noneffective stimuli, as they are outside the “window of visibility” for the neuron at hand (Ringach et al. 2002) (Fig. 1B). The maximum tested spatial frequency in these experiments (attained at the corners of the subspace tested, see Fig. 1B) varied from 2.3 to 19.6 cycles/°, with a median of 8.5. The number of unique images in the stimulus set varied from 160 to 5616 with a median of 1,248; with four spatial phases for each orientation/spatial frequency, this means that the typical experiment had roughly 312 unique combinations of spatial frequency and orientation. The stimulus contrast in all experiments was 99%. Stimuli were spatially extended and were about two to three times larger than the classical receptive field of the neurons. In general, the stimulus was $4 \times 4$° in size to allow the coverage of all receptive fields under measurement with the electrode array. These receptive fields scattered over an area of 1 deg² and had average linear size of $\sqrt{5}$°.

For each neuron, we generated an estimate of the spectral kernel by computing the probability that a grating with spatial frequency $o$, and orientation, $\theta$, preceded a spike by $\tau$ ms: $p(o;\theta;\tau)$ (Fig. 1, A and B). The responses to the different spatial phases were averaged for each cell to fire less than the baseline response.

We then proceed to estimate a baseline firing rate for the cell for noneffective stimuli, as they are outside the “window of visibility” for the neuron at hand (Ringach et al. 2002) (Fig. 1B). The maximum tested spatial frequency in these experiments (attained at the corners of the subspace tested, see Fig. 1B) varied from 2.3 to 19.6 cycles/°, with a median of 8.5. The number of unique images in the stimulus set varied from 160 to 5616 with a median of 1,248; with four spatial phases for each orientation/spatial frequency, this means that the typical experiment had roughly 312 unique combinations of spatial frequency and orientation. The stimulus contrast in all experiments was 99%. Stimuli were spatially extended and were about two to three times larger than the classical receptive field of the neurons. In general, the stimulus was $4 \times 4$° in size to allow the coverage of all receptive fields under measurement with the electrode array. These receptive fields scattered over an area of 1 deg² and had average linear size of $\sqrt{5}$°.

For each neuron, we generated an estimate of the spectral kernel by computing the probability that a grating with spatial frequency $o$, and orientation, $\theta$, preceded a spike by $\tau$ ms: $p(o;\theta;\tau)$ (Fig. 1, A and B). The responses to the different spatial phases were averaged for each cell to fire less than the baseline response.

We then proceed to estimate a baseline firing rate for the cell for noneffective stimuli, as they are outside the “window of visibility” for the neuron at hand (Ringach et al. 2002) (Fig. 1B). The maximum tested spatial frequency in these experiments (attained at the corners of the subspace tested, see Fig. 1B) varied from 2.3 to 19.6 cycles/°, with a median of 8.5. The number of unique images in the stimulus set varied from 160 to 5616 with a median of 1,248; with four spatial phases for each orientation/spatial frequency, this means that the typical experiment had roughly 312 unique combinations of spatial frequency and orientation. The stimulus contrast in all experiments was 99%. Stimuli were spatially extended and were about two to three times larger than the classical receptive field of the neurons. In general, the stimulus was $4 \times 4$° in size to allow the coverage of all receptive fields under measurement with the electrode array. These receptive fields scattered over an area of 1 deg² and had average linear size of $\sqrt{5}$°.

For each neuron, we generated an estimate of the spectral kernel by computing the probability that a grating with spatial frequency $o$, and orientation, $\theta$, preceded a spike by $\tau$ ms: $p(o;\theta;\tau)$ (Fig. 1, A and B). The responses to the different spatial phases were averaged for each cell to fire less than the baseline response.

The relative probabilities of occurrence were normalized by the median value of those in the perimeter of the region (indicated by “baseline” in Fig. 1B). The logarithm of this ratio was taken to provide a measure of discriminability of each stimulus from those in the baseline set (Ringach et al. 2002). The kernels, normalized in this way, are denoted by $R(o;\theta;\tau) = \log[p(o;\theta;\tau)/p(\text{baseline}; \tau)]$. Thus one can consider the kernels as representing the log-likelihood of a stimulus preceding a spike at different time delays. In the representation of Fig. 1B the radial distance from the origin (the center of the image) represents spatial frequency and the angle with respect to a line running horizontally through the origin represents orientation. Red regions indicate stimuli that induce the cell to fire more than the baseline, whereas blue regions correspond to stimuli that induce the cell to fire less than the baseline response.

The kernel norm at a fixed time delay is defined as the $L_2$ norm of $R(o;\theta;\tau)$ denoted by $\|R(o;\theta;\tau)\|$. We define the signal-to-noise ratio (SNR) for each kernel as $\text{SNR} = \|R(o;\theta;\tau)/\|E[\|R\|]\|$, where $E[\|R\|]$ is
onto the boundary, we first computed the sum of the absolute value of the frequency. To obtain a quantitative measure of the spillover of responses, it would mean that we had poorly chosen our maximal spatial frequency. For example, if the responses reached the boundaries of the subspace, it would mean that we had poorly chosen our maximal spatial frequency.

In addition to the foregoing SNR criterion, we measured the degree to which the responses were well contained within the subspace. For example, if the responses reached the boundaries of the subspace, it would mean that we had poorly chosen our maximal spatial frequency. To obtain a quantitative measure of the spillover of responses onto the boundary, we first computed the sum of the absolute value of the frequency. To obtain a quantitative measure of the spillover of responses, it would mean that we had poorly chosen our maximal spatial frequency. For example, if the responses reached the boundaries of the subspace, it would mean that we had poorly chosen our maximal spatial frequency.

The selection of the analysis interval was intended to ensure that our analyses spanned the duration of the response but avoided fitting the noise at long or short time delays.

The spectro-temporal separability of the kernels was assessed by singular value decomposition (SVD) of the kernel matrix within the analysis interval (Mazer et al. 2002). The three-dimensional spectral kernel \( R(\omega; \theta; \tau_{peak}) \) was rewritten as a matrix with the spectral coefficients at different time delays arranged in different columns. A singular value decomposition was performed on this matrix yielding a spectrum \( \lambda_i \) of decreasing eigenvalues. The relative variance accounted for by each of the individual components is given by \( \text{SVD}_i = \lambda_i^2 / \sum \lambda_i \). Thus for a kernel that is spectro-temporally separable, a single component is sufficient to account for the response and \( \text{SVD}_1 \sim 1 \). Thus we define the separability index as \( \text{SVI}_1 \). It should be noted that this estimate of separability applies to the entire spectral kernel bracketed by the analysis interval and that the choice of the analysis interval affects the resulting estimates of separability. Delimiting the analysis interval with respect to the noise floor was intended to ensure that changes in the spectral kernel over time reflect genuine response dynamics rather than fluctuations due to noise in the data. The results of the SVD analysis can also be used to estimate the best separable model of the spectral kernel in the least squares sense. As we will see, investigating the structure of the residuals of the optimal separable model provided some insights into the features of the responses that were left unexplained.

**Parametric analysis**

In addition to the optimal separable model derived from the SVD decomposition, we also used a parametric model to fit the data. The model assumes the responses are the result of a linear combination of two signals. Each signal is temporally separable as they are constructed by the product of a fixed spectral tuning and a temporal response function. One of the signals is positive (enhancement) and the other is negative (suppression). The spectral tuning of enhancement...
ment is given by a two-dimensional Gaussian function centered at a given preferred location on the Fourier domain (as would be expected from a 2-dimensional Gabor receptive field; Fig. 1C, left). The suppressive component is constrained to be isotropic and centered at the origin (Fig. 1C, middle). The model also included a baseline constant which was always much smaller than the two other components and is not discussed further here. The sum of these two components provided an adequate fit to most of our data.

In mathematical form, the model is expressed as

\[ R(\omega) = A(\tau) \left\{ \exp\left[ -\frac{1}{2} (\omega - \omega_0)^T \Sigma_T^{-1} (\omega - \omega_0) \right] + \exp\left[ -\frac{1}{2} (\omega + \omega_0)^T \Sigma_T^{-1} (\omega + \omega_0) \right] \right\} + B(\tau) \exp(-||\omega||^2/2\sigma_e^2) + C(\tau) \]

Here \( \omega = (\omega_x, \omega_y) \) is the spatial frequency vector. The first term represents the enhancement component centered around \( \pm \omega_0 \) with a covariance matrix of \( \Sigma_T \). The covariance matrix is unconstrained, allowing for the elongation of the enhancement signal along different directions. In a simple linear Gabor receptive field, this could occur if the carrier and the envelope do not share a common axis (Jones and Palmer 1987). The amplitude of the enhancement at any one time is given by \( A(\tau) \). The second term represents suppression centered at the origin, with a SD given by \( \sigma_e \). The amplitude of the suppressive component at any one time is given by \( B(\tau) \). Finally, the third term, \( C(\tau) \) represents a baseline.

To fit the model, we first found the optimal parameters of the spectral tuning at the optimal time delay. We then fixed the shape of the two components and found the weights for the enhancement and suppressive components that best accounted for the response at each time slice. This results in a nonparametric estimate of the temporal profiles for each component. We then iterated the procedure by fixing the temporal profiles and finding the best shape for the components that explained the entire kernel. It turned out that, in practice, this step did not usually modify the results of the first iteration significantly, so the first estimates were taken as the best possible fit. The quality of the fit was assessed by computing the fraction of the total kernel variance explained by each model throughout the analysis interval.

**RESULTS**

We observed a diversity of dynamical responses in the Fourier domain, as conveyed by the examples in Fig. 2. The kernels are presented at time lags for which their norms are a fixed fraction of the one attained at the optimal delay time. The selected fraction levels appear at the top of Fig. 2, and time proceeds from left to right. The frames occurring after the optimal delay are shown at finer temporal separations because, as we will see in the analysis in the following text, the responses exhibit more complex dynamics in their tails.

Kernels that were dominated by net enhancement in the entire Fourier domain over the full duration of the response were commonly observed (Fig. 2A). Less frequently we also observed responses that were dominated by net suppression (Fig. 2B). It was normal to find areas of net suppression developing later in the response. Figure 2C shows an example where net enhancement dominates early while net suppression dominates in the late phase of the response. The most prevalent shape of the dynamical response is shown in Fig. 2D. Here

---

**FIG. 2.** Examples of spectro-temporal dynamics. Each row displays a series of time slices of the spectro-temporal kernel at the delays indicated by the inset. The numbers at the top of each column indicate the values of the kernel norm with respect to the peak value at the optimal delay, normalized to 1. The separability indices for each response are indicated on the right side.

*J Neurophysiol* • VOL 100 • JULY 2008 • www.jn.org
there is an initial net enhancement in a restricted region of the Fourier domain close to the origin (the center of the image). As the response progresses, one observes net suppression developing at orientations orthogonal to the preferred one and at the origin. At the optimal delay time, both enhancement and suppression have a characteristic bow-tie shape. Later in time, the response is dominated by net suppression. It can also be seen that the peak spatial frequency of enhancement shifts from lower to higher spatial frequencies as if being “pushed away” by the development of suppression at lower spatial frequencies. A similar example is shown in Fig. 2E. From these examples, it appears that situations where the relative dominance of enhancement and suppression shifts during the response (as in these 2 examples) might be related to the inseparability of the responses. Finally, the example in Fig. 2F was included to demonstrate that regions of enhancement can sometimes be surrounded by suppression at higher spatial frequencies.

We computed the separability of V1 spectral kernels via singular value decomposition (SVD) (see METHODS). The method allowed us to compute the percentage of variance in the spectral dynamics that could be accounted for by an optimal separable model (by including only the 1st component) as well as the best possible combination of two separable components (by including the 1st 2 components). We note, however, that the latter method does not necessarily yield components with opposite signs, meaning that they cannot be readily interpreted as separate enhancement/suppression signals. The separability indices for the kernels in Fig. 2 are indicated to the right of each sequence.

Across the population, the mean fraction of variance accounted for the optimal separable model was 81% (Fig. 3A, ■). The second component on its own accounted for an additional 11% of the variance on average (Fig. 3A, □). Both components together, therefore, accounted for 92% of the dynamic structure of the kernels (Fig. 3A, ○). There was no correlation between the relative variance accounted for by the first component and the modulation index of the cell, meaning that the degree of separability was not statistically correlated with cell class (simple or complex; $r = -0.07; P > 0.05$).

Our discussion of the examples in Fig. 2 suggested that a possible reason for temporal inseparability is the different tuning and timing of enhancement and suppression. For example, the two kernels that are most separable are either entirely enhanced (Fig. 2A) or entirely suppressed (Fig. 2B), whereas those that show early enhancement and late suppression are more inseparable. We thus asked if spectro-temporal inseparability, as measured by the amount of variance explained by the first principal component, correlates with a measure of overall dominance by one response sign. To do so we defined the response valence index by $RVI = (v_e - v_s)/(v_e + v_s)$, where $v_e = \text{var}[R(\omega; \theta; \tau)]^+$, $v_s = \text{var}[R(\omega; \theta; \tau)^+]$, and $[x]^+$ represents half-way rectification. The variance is computed over the entire spectro-temporal analysis window, so these represent measures of total net suppression and enhancement. An RVI near +1 implies that the response is dominated by enhancement; an RVI near −1 implies a response dominated by suppression; and an RVI near 0 implies net enhancement and suppression are well balanced over the entire response period (though one or the other may dominate at different time delays).

When the response valence index is plotted against the separability index (the amount of variance accounted for by the 1st component alone), the points are distributed in the form of a “V” (Fig. 3B), indicating that their kernels are least separable when enhancement and suppression have roughly equivalent magnitudes. In other words, the separability of V1 spectral kernels is correlated with the balance of net enhancement and suppression. This is verified by the fact that the absolute value of the response valence index
was significantly correlated with the separability index ($r = 0.58; P < 10^{-10}$). Also, note that the distribution of the response valance index makes it evident that while there is a wide range of dynamical responses, there is a preponderance of cells dominated by enhancement.

Having shown that the net balance of enhancement and suppression across the entire response is correlated with the separability of the responses, we attempted to characterize the time courses of these components. To do so, we defined a measure of net enhancement in the Fourier domain at a given time delay as $\sqrt{v_e}$ and, similarly a measure of net suppression as $\sqrt{v_s}$. Here the quantities were computed across the Fourier domain for each time delay. We independently normalized these values by subtracting a baseline estimated at time delays where the response and the stimulus were uncorrelated, and then divided by the norm of the full kernel at the optimal time. To average across cells, we first aligned the curves at the optimal delay time.

We performed the analysis for our population by splitting responses into those that were dominated by enhancement (RVI > 0; Fig. 4A) and those dominated by suppression (RVI < 0; Fig. 4B). In both groups of cells, it is evident that suppression peaked after enhancement (Fig. 4, A and B). Therefore even when suppression dominates, the response enhancement is seen as preceding.

The differential timing of enhancement and suppression, of course, cannot be captured by a separable model. This is made clear by repeating these calculations for the optimal separable fits obtained in the SVD analysis (Fig. 4, C and D). As these curves are based on kernels that are by construction separable, the time courses of enhancement and suppression differ only in their magnitude and not their shape. There is, furthermore, an interesting difference between the two cell groups that can be seen in these fits. The optimal separable fits to the kernels dominated by suppression have long tails that extend $\leq 40$ ms after the response peak. In contrast, responses that are dominated by enhancement have a more symmetric temporal profile and tails decay within 25 ms of the peak of the responses. This suggested to us that an increase presence of suppression may cause longer tails and therefore a more asymmetric profile of the temporal time course of the response.

To investigate if the temporal shape of the kernel was indeed more asymmetric and had longer tails for the responses dominated by suppression, we performed an additional analysis. We first found, for each cell, the best Gaussian fit to the kernel norm as function of time delay. The quality of this fit was assessed by the percentage of variance it explained. We observed that the average residuals from these fits showed structure that concentrated in the late part of the response, around 20–60 ms after the optimal time delay (Fig. 5A). This implies

![Figure 4](http://jn.physiology.org/)

**FIG. 4.** Average time course of kernel norms. A and B: average (normalized) time courses of the kernel norm (blue), net enhancement (red), and suppression (green), are shown for cells that were dominated by enhancement (A) or suppression (B). The average behavior shows a delayed suppression for both cell groups. A similar analysis for the optimal separable model is shown in C and D. Error bars are twice SE.
there was also a significant but weaker correlation between the separability and the response valance ratio (Fig. 3B; \( r = 0.22; P < 2 \times 10^{-6} \)). These results indicate that responses with balanced suppression and enhancement tended to have longer tails and be less separable than responses dominated by one response sign.

It should be noted that regions of enhancement and suppression in our data represent areas of net enhancement and suppression. The tuning of these signals in the Fourier domain can certainly overlap to generate the measured kernels. We wanted to investigate if a parametric model, where enhancement and suppression had simple shapes, could explain some of the dynamics of responses in the Fourier domain. The simplest model is one where enhancement has a spectral tuning consistent with that of a Gabor filter. In this case, the spectral tuning is the sum of two Gaussian functions symmetric across the origin. Suppressive influences from within the classical receptive field have often been described as being broadly tuned in spatial frequency and untuned for orientation (Bauman and Bonds 1991; Bonds 1989; Deangelis et al. 1992). Thus we opted to model suppression as a single Gaussian centered at the origin of the Fourier plane. The model assumes that both enhancement and suppression are separable in time and that they add linearly. This model was adequate, on average accounting for 71% of the variance of the responses (Fig. 6).

The model fits were used to estimate the relative magnitude and time course of the underlying enhancement and suppression components. To do this, we computed the population average of their magnitudes after normalizing to the maximum absolute value between the two. Cells were split into enhancement or suppression dominated as above. This analysis showed that the magnitudes of the two components inferred from the model were well balanced with each other for both groups of cells (Fig. 7). Consistent with the results of the nonparametric analysis, peak enhancement precedes suppression in both cases by a few milliseconds. Similar results were also obtained when the population averages were computed for cells with the highest and lowest separability indices (SI > 0.9 and SI < 0.7; data not shown). This indicates that balance of the magnitudes of suppression and enhancement was rather typical of the responses. The variability in the degree of net suppression and inseparability seen in the data, therefore must be generated by a difference in the tuning and timing of the components in the Fourier domain and not by a strong imbalance in the underlying magnitude of the two signals.

**DISCUSSION**

Here we showed two novel relationships between suppressive signals and the dynamics of V1 neurons. First, there is a correlation between temporal separability in the Fourier domain and the dominance of net suppression or enhancement in the responses (Fig. 3B). A different way to express this fact is that when responses are dominated by net enhancement (or, in a minority of cells, by net suppression), they tend to be separable. Inseparable responses, on the other hand, exhibit both net enhancement and suppression in different regions of the Fourier domain. Second, inseparable responses tend to have longer tails than those that are separable (Fig. 5). These new data add to an already established correlation between selectivity in the Fourier domain and the presence of net suppression: cells that are suppressed by nonoptimal stimuli are better
tuned than cells with only an enhanced response (Ringach et al. 2002). Altogether these findings show that the balance among enhancement/suppression, tuning selectivity, temporal separability, and persistence in time are factors that are closely linked in the dynamics of V1 neurons.

The observed relationships can be understood within the context of the conceptual two-component model where a suppressive component, delayed in time, shapes the initial tuning of the enhancement signal. In the process, suppression not only increases the selectivity of the cell (by restricting the region in the Fourier plane that induces the cell to increase its firing rate) but also introduces, due to its different tuning in the Fourier domain, an associated temporal inseparability. We note that while the majority of the kernels could be adequately accounted for using a broadly tuned suppressive signal, there is no denying that there are clear instances where tuned suppression is necessary to explain the responses. This is particularly the case for extended stimuli that cover both the classical receptive field and its surround (Blakemore and Tobin 1972; Ringach et al. 2003; Xing et al. 2005) and is most prevalent in well-tuned neurons (Ringach et al. 2003).

Modeling the responses as the sum of two separable enhancement/suppression components allowed us to tease apart their magnitude and temporal dynamics. This simple model is meant to be interpreted as a convenient and succinct summary of the dynamical responses and cannot be directly mapped onto the different components of the underlying circuitry. In particular, there might be a number of different mechanisms contributing to the generation of a broadly tuned suppression signal. It may originate from saturation in the LGN (Li et al. 2006; Priebe and Ferster 2006), from short-term synaptic depression in thalamo-cortical synapses (Carandini et al. 2002; Freeman et al. 2002) [although recent data suggest that synaptic depression is nearly saturated in vivo, limiting its potential role (Boudreau and Ferster 2005)] and from intra-cortical inhibition (Bauman and Bonds 1991; Bonds 1989; Deangelis

---

FIG. 6. Examples of spectral kernels and their separable and parametric fits for 3 neurons. The examples shown in A–C correspond to Fig. 2, C–E, respectively. In each set of panels, the actual kernel is shown in the top row, the separable model in the second, the parametric model as the bottom row. As in Fig. 2, the columns indicate the magnitude of the kernel norm, relative to the peak, for the kernel frame that is shown.
the dynamics in the Fourier domain may depend on the laminar location and the tuning properties of neurons in the neighborhood could help specify the contribution of feed-forward inputs and intra-cortical feedback to the dynamics of cortical responses.

Prior studies have confirmed the presence of some of the inseparable features we observed in the Fourier domain, such as the sharpening in orientation bandwidth, shifts in preferred spatial frequency, and shifts in preferred orientation (Chen et al. 2005; Mazer et al. 2002; Schummers et al. 2007), but some conflicting data exist (Gillespie et al. 2001; Mazer et al. 2002; Sharon and Grinvald 2002). One previous study explicitly quantified separability in the Fourier domain, reporting the kernels to be largely separable (Mazer et al. 2002). One key difference between the studies is the way the data are normalized (see a detailed discussion in Ringach et al. 2003). Indeed when we repeat the SVD analyses on the raw data, $p(o;\theta;r)$, instead of the normalized log-likelihood responses, $R(o;\theta;r)$, we obtain separability indices with a mean of 94%, which is very similar to what they reported. We note that despite such high separability indices, these authors also found a clear shift in the optimal spatial frequency of neurons (an inseparable feature of the dynamics). Mazer et al. (2002) did not report signs of suppression for nonpreferred stimuli. This prior study was carried out in awake animals, whereas ours was done in anesthetized animals. One possibility is that the presence of fixation noise would make suppression more difficult to detect in the awake preparation. Another possibility is that suppression is enhanced in the anesthetized preparation. A recent study in (anesthetized) cat visual cortex used a dynamic noise pattern to stimulate the cells and computed the average local spectrum of the stimulus that correlated with the spiking of the neuron (Nishimoto et al. 2006). While some of the measured responses resemble the ones reported here (see their Fig. 10), they comprised a substantially smaller proportion of the cases. It is likely the differences here are due to differences in the class of stimuli used. The finely grained dynamic noise used in Nishimoto et al. (2006) could fail to efficiently drive suppressive mechanisms that, in our hands, are centered at low spatial frequencies.

To summarize, the observed relationships among tuning selectivity (Ringach et al. 2002), separability, and persistence in time of V1 responses can be explained by the combination of two signals that are differentially tuned in the Fourier domain and have different temporal dynamics. Further studies are needed to explain the circuitry underlying these two components and to investigate if the substantial variability in the dynamics could be partly explained by the laminar locations of cells or their position within the functional maps of the cortex.

**Acknowledgments**

We thank I. Nauhaus for helpful comments regarding an earlier version of the manuscript.

**Grants**

This work was supported by National Eye Institute National Research Service Award EY015365-01 and National Eye Institute Grants EY-12816, and EY-18322 and Defence Advanced Research Projects Agency FA 8650-06-C-7633.
REFERENCES


