



Temporal receptive field estimation using wavelets

Alan B. Saul*

Department of Ophthalmology, Medical College of Georgia, Augusta GA 30912, USA

Received 14 June 2007; received in revised form 13 November 2007; accepted 15 November 2007

Abstract

A standard goal of many neurophysiological investigations is to obtain enough insight into a neuron's behavior that it becomes possible to predict responses to arbitrary stimuli. Techniques have been developed to solve this system identification problem, and the numerical method presented here adds to this toolbox. Stimuli and responses, beginning as functions of time, are transformed to complex-valued functions of both time and temporal frequency, giving amplitude and phase at each frequency and time point. The transformation is implemented by wavelets. The kernel describing the system is then derived by simply dividing the response wavelet by the stimulus wavelet. The results are averaged over time, incorporating median filtering to remove artifacts. Estimated kernels match well to the actual kernels, with little data needed. Noise tolerance is excellent, and the method works on a wide range of kernels and stimulus types. The algorithm is easy to implement and understand, but can be applied broadly. © 2007 Elsevier B.V. All rights reserved.

Keywords: Receptive field; Impulse response; Reverse correlation; Wavelet; Temporal frequency; System identification; Kernel estimation

1. Introduction

Numerous methods for extracting information about how a neuron might respond to arbitrary stimuli have been developed. Marmarelis and Marmarelis (1978) inspired much of this by introducing Wiener kernel analysis to neuroscience. The field has been dominated by methods that work in the time or space (or the tonotopic axis in the auditory system, the spatial dimension along the cochlear receptor surface) domains. Frequency domain approaches have nonetheless provided important insights (French, 1976; Victor, 1979; Victor and Shapley, 1979; Christakos et al., 2004; Luczak et al., 2004; Nishimoto et al., 2006).

Wavelets (or alternatives such as Wigner transforms) provide a hybrid approach (Torrence and Compo, 1998; Soucek et al., 2004). Wavelet transforms create time–frequency representations, with optimal localization in each domain. A method that uses wavelets can predict first-order kernels under a wide range of conditions. First, starting with known kernels, numerical computations demonstrate that they can be reconstructed rapidly from presentation of several kinds of noise stimuli, with excellent tolerance when responses have additive noise. Second, results are presented from testing visual neurons for which the

kernels are unknown *a priori*, comparing the wavelet analysis with conventional methods.

Only temporal analyses are treated here. Spatial aspects take the simplest form, as temporal response profiles are derived independently at each position. The underlying reason for this choice is that neuronal inputs are localized in space, in that they have compact spatial receptive fields. However, the methods described below could be applied to experiments where spatial receptive fields need to be derived from spatially extended stimulation (Enroth-Cugell and Robson, 1966; Aertsen and Johannesma, 1981; deCharms et al., 1998; Depireux et al., 2001; David et al., 2004; Prenger et al., 2004).

Neurons are driven by multiple inputs. Responses from neurons, or from the whole retina in the case of multifocal electroretinography (Sutter, 2001), for example, represent some additive combination of signals from spatially separated receptive fields. One can simultaneously measure the temporal properties at each spatial position by presenting noise stimuli that are statistically independent across space. From the point of view of any given position, the total response looks like the response evoked by the stimulus at that position, plus noise from the activity generated from all of the other positions. Analysis methods that extract temporal kernels in the presence of noise are therefore needed.

The goal of this report is to illustrate the ease with which neurophysiologists can implement a method that permits extraction of system kernels under a wide variety of situations. The use of

* Tel.: +1 706 721 0695; fax: +1 706 721 7913.
E-mail address: asaul@mcg.edu.

wavelets neither optimizes systems identification nor provides clean analytic tools for treating theoretical issues. The demonstrations here are strictly numerical, and the focus is pragmatic.

2. Methods

2.1. Wavelet correlations: non-technical summary

The goal of systems analysis is to estimate how a system converts arbitrary inputs into outputs, often in the interest of discerning mechanisms. In the present case, I obtained responses of model and neuronal systems to controlled stimuli, and derived estimates of the systems' structures by correlating the stimuli and responses. These correlations were performed after transforming both stimuli and responses into time–temporal frequency representations, using a wavelet transform. At each time bin, a stimulus and its corresponding response contain amplitudes and phases at each of a range of temporal frequencies. The correlation between the stimulus and response consists of dividing the response by the stimulus. The results of these divisions are then accumulated over time.

Inputs were taken as 5 s trials, to match the physiological protocol described below. Inputs were samples of pseudorandom noise. The noise was either white, with equal amplitude across linear frequency bins, or natural, with amplitude skewed toward low frequencies, which tends to have equal amplitude over logarithmically spaced frequency bins.

The input was passed through a kernel (convolving with the impulse response function or multiplying by the frequency-domain filter; the former of these equivalent methods was used below). This provided the linear response corresponding to the input. In some cases this linear response was then modified by a static nonlinear function to produce the final response. Noise was also added in some cases. The response was a function of time, over 5 s, just like the stimulus.

Stimulus and response were transformed by a complex Morlet wavelet. This is equivalent to a set of convolutions with Gabor functions of varying frequencies. In standard terminology, this yields complex functions of time and scale. Wavelet scales correspond to reciprocals of temporal frequencies, and I will describe the analyses in terms of frequency rather than scale. Fig. 1 (D, F, G, and I) shows an example of the amplitude and phase components of stimulus and response after wavelet transformation. The vertical axis is temporal frequency, and the horizontal axis is time. The response to this stimulus was computed by convolving the stimulus with the kernel shown in black in Fig. 1B. Because of the bandpass amplitude tuning of this kernel, responses to low and especially high frequencies are attenuated. The kernel shifts the phase, as can be noticed on inspection of the colors in the phase plots in (G and I). Phase is given in cycles (H) throughout.

The original kernel can be reconstructed, therefore, by comparing the response and stimulus. The amplitude tuning of the kernel is derived from the ratio of the response and stimulus amplitudes, and the phase of the kernel from the difference of the response and stimulus phases. These separate computations are actually combined by taking the ratio of the complex numbers formed from the amplitude and phase pairs. For a linear

system, these comparisons are made at common frequencies in stimulus and response, since linear systems do not evoke any interactions across frequencies. In fact, the wavelet transform, which is incompletely localized in both time and temporal frequency, introduces overlap between nearby frequencies, causing small errors. The time dimension amounts to a set of samples, and a statistic must be used to compile the response/stimulus comparisons across time. I computed the mean across time of the amplitude/phase values at each frequency, in the complex plane (Fig. 1J–M).

Nonlinear systems can produce interactions across frequencies. For instance, a squaring or other multiplicative nonlinearity leads to responses at sums and differences of the stimulus frequencies. Because the wavelet representation encompasses a wide range of frequencies, interactions can be computed at each combination of these frequencies (Li et al., 2007). The present study will be restricted to the linear computations.

2.2. Details of computation

For most of the results below, the frequency range used was about 0.25–64 Hz. Step size was a quarter octave over these 8 octaves. These choices induce redundancy in the wavelet representation, which has advantages and disadvantages, and is standard practice (Soucek et al., 2004).

The full calculation consisted of iteratively repeating the following steps:

- (1) (Fig. 1A and C) starting with stimulus and response arrays that comprised 800 points in 6.25 ms steps;
- (2) (Fig. 1D, F, G, and I) the stimulus and response are transformed to time–frequency representations, using calibrations to map wavelet scales to temporal frequencies and to determine wavelet amplitudes at each frequency;
- (3) (Fig. 1J–M) the complex arrays of stimulus and response values are divided;
- (4) at each frequency, the amplitudes across time are median-filtered to avoid spurious correlations, as discussed below;
- (5) (Fig. 1J–M, green discs) a single complex value for the correlation at each frequency is obtained by averaging across time in the complex plane;
- (6) (Fig. 1E and H, red curves) the resultant array of complex values across frequencies are added to an estimate that accumulates over trials, providing estimates of the kernel temporal frequency tuning and timing;
- (7) (Fig. 1B, red curve) a complex array that has the dimensions of the Fourier transform of the original stimulus and response arrays is filled with values interpolated from the computed array from step 6, and this array is inverse Fourier transformed to yield an impulse response function with 6.25 ms steps.

Stimuli can be arbitrary, but here consist of pseudorandom noise of several types. The experimenter must know the stimulus, either by generating it or recording it. The response to the stimulus is assumed here to come about through a quasilinear process. The system is described by a kernel that accounts for

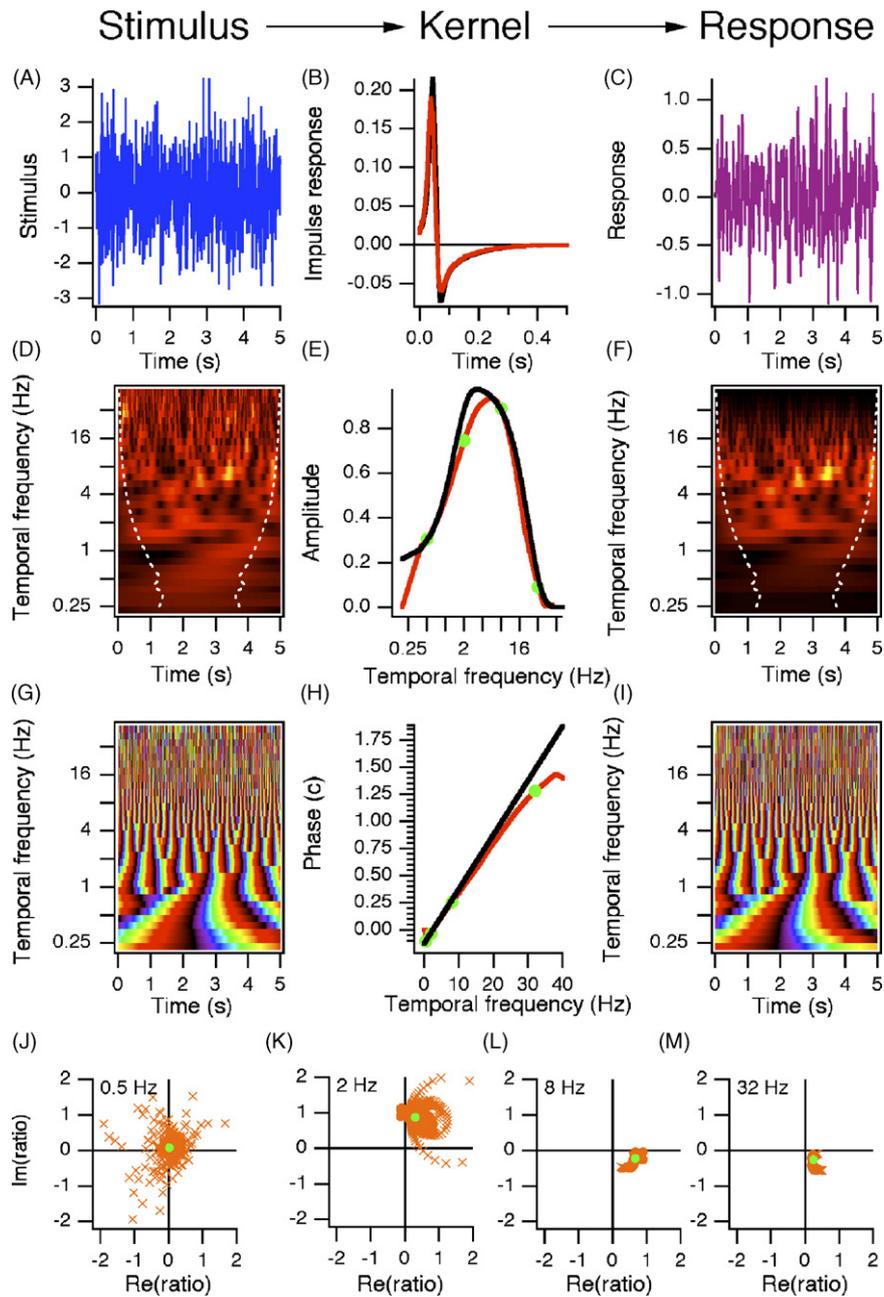


Fig. 1. Wavelet correlation procedure. The stimulus (A) was Gaussian white noise, and the response (C) was the linear transform of this stimulus by the filter shown in (B) (black trace). In the amplitude plots in (D and F), the color scale ranges from black to yellow. In the phase plots in (G and I), the color cycles from black through violet, blue, yellow, and red and back to black. The wavelet representation in (D, F, G, and I) provides amplitude and phase as a function of time and frequency. The white dashed traces in (D and F) indicate the cone of influence. The kernel with the time domain shape in (B) has the frequency-domain representation shown in (E and H), in black. The unit of phase here is cycles (c). The division in time–frequency space of response by stimulus yields a complex number at each frequency and time. Examples at 4 frequencies are shown in (J–M), with the complex numbers displayed in the plane by an orange cross for each time point. The average of the points in these graphs is shown as a green circle. The amplitude and phase values corresponding to these averages at each frequency are shown in (E and H) as the red curves, with the 4 example frequencies marked with green circles. These frequency-domain results lead to the kernel estimate in (B) (red trace).

the linear portion of the stimulus-response relation. For many neural systems, temporal kernels consist of smooth functions of time (impulse responses) with durations less than several hundreds of milliseconds. In the frequency domain, these kernels are tuned to frequencies between 0.25 and 32 Hz, often with bandpass tuning but sometimes with lowpass tuning, and have an approximately linear relation between phase and frequency. The slope of the phase vs. frequency line is a measure of *latency*,

which is typically less than hundreds of milliseconds. The phase at low frequencies (*absolute phase*, the extrapolation of phase to 0 Hz) determines the shape of the impulse response, and, for example, whether the kernel leads to transient or sustained responses to step stimuli: absolute phase leads nearing a quarter cycle give transient responses and biphasic impulse responses, whereas absolute phases near 0 cycles give sustained responses and monophasic impulse responses.

Model kernels in this study were constructed by selecting values of 5 parameters that describe the kernel in the frequency domain (I refer to this as a kernel for convenience; usually, the frequency-domain version is called the modulation transfer function): phase was given by $\varphi_0 + L\omega$, where ω is temporal frequency, φ_0 is absolute phase and L is latency, and amplitude was given by $e^{-1/2(\omega/E_w)^2} - I_a e^{-1/2(\omega/I_w)^2}$, where E_w is excitatory width, I_w is inhibitory width, and I_a is inhibitory amplitude. E_w varied between 4 and 20 Hz, I_w varied between 0.1 Hz and $E_w - 2$ Hz, I_a varied between 0.05 and 0.99, absolute phase varied over 1 cycle, and latency ranged from 30 to 200 ms. The frequency domain and time domain versions of the kernel were obtained from each other via Fast Fourier Transforms (FFTs) or Inverse Fast Fourier Transforms (IFFTs).

Responses were obtained by convolving the stimulus s with the system kernel k , $k \otimes s$. In some cases, noise was then added to the response, to investigate how well the kernel could be reconstructed in the presence of noise. The response can be written as $r(t) = k \otimes s(t) + v(t)$, where v is the additive noise. The standard way to obtain the kernel k is to convolve both sides with s : $r \otimes s = (k \otimes s) \otimes s + v \otimes s$. The convolutions $(k \otimes s) \otimes s$ become the kernel convolved with the stimulus autocorrelation, $k \otimes (s \otimes s)$. If the stimulus is white, its autocorrelation is a delta function, and the convolution with the kernel is exactly the kernel. This calculation is standard in neurophysiology, where it is termed the spike-triggered average. If the stimulus is not white, however, one needs to deconvolve its autocorrelation from the kernel.

Because the added noise v is assumed to be uncorrelated with the stimulus s , that term vanishes (uncorrelated means that the convolution is zero). This leaves $k \otimes (s \otimes s) = r \otimes s$, which can be solved in the frequency domain, where $r \otimes s$ is transformed to RS^* , the product of the Fourier transforms of stimulus and response (S^* is the complex conjugate of S). For the conventional “spike-based” reverse correlations below, stimulus and response were transformed via FFTs, and the quotient $K = RS^*/SS^*$ was computed in the frequency domain, then an IFFT transformed the result back to the time domain impulse response function. Regularization was added by smoothing the power spectrum SS^* prior to the division, or by adding a regularizing parameter to the power spectrum before dividing.

The wavelet transform (step 2 above) was implemented using the CWT operation in Igor Pro 6 (WaveMetrics, Lake Oswego OR). Similar algorithms are available in most recent numerical analysis software, or could be coded based on descriptions such as [Vrhel et al. \(1997\)](#). The continuous wavelet transform of a function $f(t)$ by a “mother wavelet” ψ is

$$F(\omega, t) = \sqrt{\omega} \int f(\tau) \psi^*(\omega(\tau - t)) d\tau.$$

This wavelet representation is given in terms of temporal frequency ω and time t , as used below. However, most implementations transform the one-dimensional input function to a function of “scale” and time. Wavelet scales are reciprocals of temporal frequencies ([Torrence and Compo, 1998](#)). The correspondence between frequencies and scales was calculated empirically for each set of parameters, via a calibration routine that also measured scaling of amplitudes, and the cone of

influence outside of which the wavelet-transformed values are less reliable. The integral in the transform formula is taken over the real line. However, the key is that the mother wavelet ψ has particular properties, including effectively vanishing outside a finite interval. I use the complex Morlet wavelet here, which is a Gabor function:

$$\psi(u) = \left(\frac{1}{\pi^{1/4}} \right) e^{i\omega_0 u} e^{-u^2/2},$$

in other words, a complex sinusoid windowed by a Gaussian. The wavelet transform is a set of local Fourier transforms. Note that the transformed signal is complex, with amplitude and phase. In the application here, the phase values are key. The parameter ω_0 is chosen to obtain a reasonable number of oscillations within the Gaussian window, and a value of 5 is used below, except when its variation is considered. The mother wavelet requires some normalization in addition to that given here, which is obtained numerically by the calibration routine. For the purposes of computing correlations, however, the amplitude normalization is not necessary, because of the division described next.

Having transformed the stimulus and the response, the relationship between them can be discovered by comparing their relative amplitudes and phases. At each time and frequency, the response is divided by the stimulus (step 3). Because the stimulus amplitude (i.e. contrast) can be small at some points, this division can lead to large quotient amplitudes if the response depends on anything more than the stimulus itself (which occurs for the cases of added noise and multi-input systems). In order to avoid having these large artifacts skew the results, I eliminate them by, at each frequency, comparing the ratios across times. Because the 800 time points should, in principle, yield similar amplitudes and phases at any given frequency, the outliers due to low contrasts stand out. I used a median filtering algorithm that looked for amplitudes exceeding the median by at least one standard deviation over the 1250 ms interval surrounding each point, and reset outliers to the median value (step 4).

The goal is to obtain a description of the system kernel. In the frequency domain, that means having values of amplitude and phase at each frequency. At this point, we still have 800 values of amplitude and phase at each frequency, and need to perform some type of averaging across the 800 time points. If amplitudes and phases were averaged separately, even a random set of 800 complex numbers would be biased toward large amplitudes, because amplitudes are non-negative. It is crucial to measure how well the sample of complex numbers coheres, that is, points in a definite direction in the plane. I simply calculated the mean of the 800 complex numbers in the plane (step 5), but one could instead compute other statistics such as the mode. Because the wavelet transform is subject to edge effects, I discarded time points near the beginning and end of each trial, based on the cone of influence that is a function of frequency (white dashed traces in [Fig. 1D](#) and [F](#)). My criterion for the cone of influence was to include those points whose amplitudes for a sinusoidal input were within one standard deviation of the maximum amplitude at that frequency (the ratio of this criterion amplitude to the maximum amplitude was 0.93 ± 0.01). This is similar to the criterion given by [Torrence and Compo \(1998\)](#).

At this point, one has the amplitude and phase at each frequency, which was the goal. However, these apply to the current trial, and should be combined with all of the other trials. I simply averaged the complex function of frequency across trials. Alternatives exist, such as averaging all of the time points from all of the trials, or forming a matrix whose rows describe the frequency-domain responses for each trial or at each time point, then performing principal components analysis and projection onto a subspace (Schwartz et al., 2006). Finally, the kernel is obtained for the entire run. The frequency-domain version can be inverse Fourier transformed to give the impulse response function that describes the system. In order to do this, a new array must be created in the frequency domain, since the wavelet frequencies are logarithmically spaced, and the IFFT expects a linearly spaced array. One must guard against extrapolation errors in the process of filling this array due to the typically bandpass tuning, but also when tuning is non-monotonic at the extremes. I set amplitudes at low and high frequencies to zero, which creates inaccuracies, including edge effects that might be reduced by using a smooth window.

For neuronal data, the response arrays were histograms of spikes during the trial, with 6.25 ms bins. Because the histograms have only non-negative values, they contain a strong component at low frequencies that does not exist in the model responses. These histograms were therefore subjected to a high-pass filter, multiplying the amplitude in the frequency domain by $1 - e^{-\omega/0.4}$ in order to eliminate the DC and reduce the very low frequency components. For some comparisons below, I computed conventional “spike-based” reverse correlations as well, using the method described above. Cells were tested at multiple positions across space. These positions were always tested independently and no dependencies were introduced by the analysis methods. Simulations of multi-input neurons were not performed because they are equivalent to the simulations of single-input neurons with noise.

All work was performed in Igor Pro 6 (example code to regenerate the figures is available at <http://www.igorexchange.com>). Neuronal data were obtained from recordings in awake behaving rhesus monkeys (*Macaca mulatta*). Details of the animal care, surgery, and recording setup are identical to those in Tang et al. (2007). All procedures adhered to NIH guidelines and were approved by the Animal Care and Use Committee of the Medical College of Georgia. Monkeys performed a fixation task during 5 s trials, and stimulus position was adjusted based on signals from an eye position monitor to compensate for fixational eye movements (Tang et al., 2007). Stimuli were presented on a CRT monitor at a frame rate of 160 Hz, accounting for the 6.25 ms binwidth used throughout.

3. Results

I first address basic issues for these analyses, by examining first-order responses derived from linear transformations of Gaussian white noise. Then, kernel estimates are derived in the presence of response noise, for nonwhite inputs, and for rectified responses. Finally, results of the analyses on real neurons are shown.

3.1. Basic findings: convergence, accuracy, flexibility, noise tolerance

Does the wavelet technique provide accurate estimates of system kernels? How much stimulation time does it need to yield adequate predictions? Are the estimates biased? Which parameters affect accuracy? Can kernels be extracted in the presence of noise? I address these questions by first considering simple situations where known linear systems are tested with white noise stimuli. Repeated presentations of 5 s trials with Gaussian white noise produced estimates that could be compared to the known kernel.

Given enough time, most methods are capable of generating accurate estimates of system structures. Obtaining these estimates after limited testing is one of the most important practical considerations. Wavelet correlations excelled on this score. Predictions converged rapidly toward model systems. Typically, tens of seconds of stimulation were sufficient to approach asymptotic points for first-order kernels. Examples are shown in Fig. 2A. Predicted temporal frequency timing (phase) and tuning (amplitude) and the corresponding impulse responses are shown for the first 10 iterations. The earlier predictions (greenish hues) are noisy but otherwise deviate little from the later predictions (bluish hues), so that they overlap almost completely except in the amplitude plot. The estimate of the target phase (black line) has small systematic errors. The deviation at high frequencies, above 20 Hz, is associated with the low amplitudes at those frequencies, however. Means and their standard errors after 10 iterations are shown in (B). The estimated phase values are accurate, though with a smaller slope. The amplitude estimates are less reliable. The impulse response is captured well.

Arbitrary kernels were matched by this process. I generated a variety of first-order kernels by randomly choosing parameters that describe amplitude and phase as functions of temporal frequency (described explicitly in Section 2 and below). Phase was linear, with parameters of slope (*latency*) and intercept (*absolute phase*). Amplitude was given by a difference of Gaussians with 3 parameters, excitatory half-width, inhibitory half-width, and the ratio of inhibitory to excitatory amplitudes. Fig. 3 shows how kernels that differed only in absolute phase (the phase at 0 Hz) were estimated. Phase values shifted as expected, with all estimates being close to the actual kernel phases at frequencies between 2 and 10 Hz, where amplitudes were strongest. Transient kernels (± 0.25 c) produced underestimates of amplitude across moderate frequencies, and sustained kernels (± 0.05 c) produced overestimated amplitudes at low frequencies and underestimated amplitudes at moderate frequencies. Latency is shorter than in the model, as the large phase shifts at high frequencies are underestimated.

Most variations in kernels had minimal effects on the accuracy of the kernel estimates. A sensitivity analysis was performed by varying 5 parameters that were used to generate kernels: absolute phase, latency, excitatory width, inhibitory width, and inhibitory strength. Phase was given by $\varphi_0 + L\omega$, where ω is temporal frequency, φ_0 is absolute phase and L is latency, and amplitude was given by

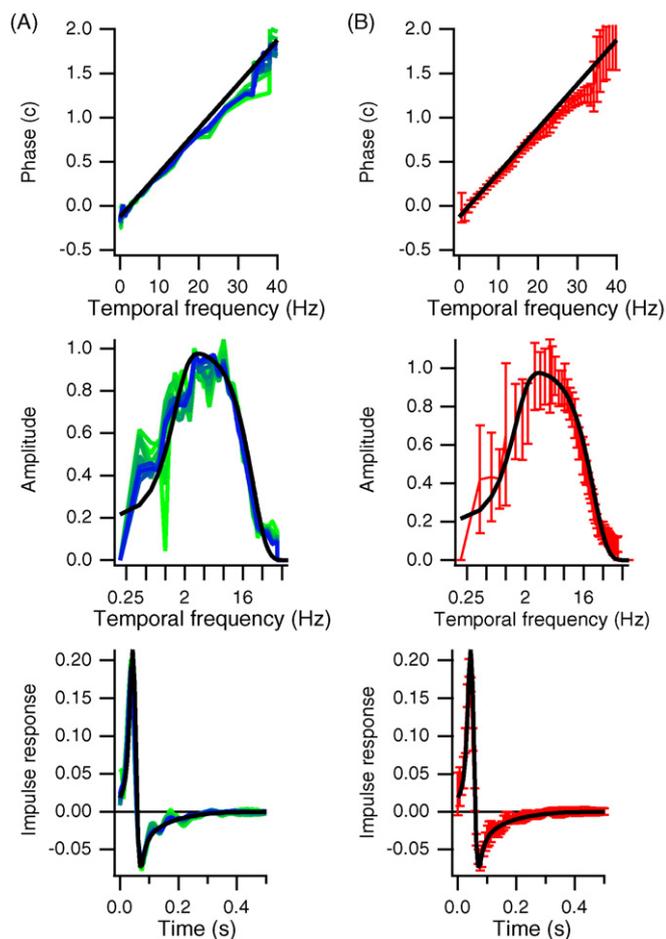


Fig. 2. Convergence and reliability. A model kernel, shown in black, was used to generate responses to Gaussian white noise inputs, with noise added to the responses. Estimates of first-order kernel phase and amplitude were obtained over a range of temporal frequencies from 5 s trials. Impulse responses were computed from the frequency-domain results. (A) Estimates are shown after each of the first 10 trials, as averaging accumulated. The first trial is in bright green, and subsequent trials are shown with increasingly blue traces. (B) The means and standard errors for these 10 trials are shown in red. A subset of the phase and amplitude points is shown here to reduce overcrowding in the plots.

$e^{-1/2(\omega/E_w)^2} - I_a e^{-1/2(\omega/I_w)^2}$, where E_w is excitatory width, I_w is inhibitory width, and I_a is inhibitory amplitude. E_w varied between 4 and 20 Hz, I_w varied between 0.1 Hz and $E_w - 2$ Hz, I_a varied between 0.05 and 0.99, absolute phase varied over 1 cycle, and latency ranged from 30 to 200 ms. Sample size was 2000 random kernels. The accuracy of the estimates was evaluated by computing the correlation between the model kernel and its estimate; specifically, the impulse response functions between 20 and 500 ms were compared to generate an *estimate accuracy* index. Correlation coefficients between estimate accuracy and each of the 5 parameters were 0.05 for absolute phase, -0.67 for latency, -0.47 for excitatory width, -0.28 for inhibitory width, and -0.17 for inhibitory amplitude. As either latency or excitatory width increased, estimate accuracy deteriorated significantly ($P < 0.001$, Pearson correlation test; Fig. 4). Kernels with both large E_w and L were poorly predicted. No significant correlations were seen for the other 3 parameters. Increases in the slope of the phase vs. temporal frequency rela-

tion, that is, the latency, caused estimates to deteriorate because the stimulus/response correlations shifted to attenuated levels of the Gaussian envelope of corresponding wavelets, and to neighboring wavelets in time, losing the connection that was assumed between the phases of stimulus and response wavelets at the same position in time. That is, the localization inherent in time–frequency analysis underlies this defect (which could be corrected with more complicated analyses). The effect of excitatory width is related to this, because the effect of latency is felt at high frequencies, which are emphasized more as excitatory width increases. In real cells, there is a tradeoff between these two parameters, as cells tuned to higher frequencies tend to have shorter latencies. In particular, such cells have latencies closer to 50 ms, where the wavelet reconstructions remained accurate.

The estimate accuracy measure is particularly sensitive to high frequencies. This can be shown by looking at the correlation between estimate accuracy and the errors in the various frequency-domain parameters. Estimates were computed for a set of 100 random kernels for each of 11 values of the mother wavelet parameter ω_0 . The estimate accuracy measure was then compared to the errors in each of 6 parameter estimates: amplitude (original kernels always had an amplitude of 1, but estimates could vary), inhibitory amplitude, excitatory width, inhibitory width, latency, and absolute phase. The only significant correlations were found between estimate accuracy and the error in the estimation of the E_w parameter ($r = 0.81$, averaged across ω_0 values between 5 and 15; the positive correlation is because the errors in E_w are consistently an underestimate of the actual value). That is, most of the errors in the estimation of the impulse response functions could be attributed to poor estimates at high frequencies.

The estimate accuracy measure improves as ω_0 increases, though not monotonically. Fig. 5 illustrates this dependence. One should think of this parameter as reflecting the localization in the analysis, with increasing values meaning less localization. As the number of cycles in the envelope increases, the envelope must enlarge at a given actual frequency. In the limit of large values, the wavelet analysis is no longer localized, and becomes a Fourier analysis. An oscillation occurs across even and odd values. As ω_0 increases, estimates at low frequencies are no longer computed. I chose a value of 5 for the rest of the work here in order to capture low frequency behaviors. This choice sacrifices accuracy as evaluated by the estimate accuracy measure, but localizes the analysis. Choosing estimate accuracy to judge the accuracy of estimates therefore biases the results against the wavelet method.

Real responses contain not only the signal evoked by the stimulus, but additional influences from other inputs (see below) and intrinsic activity. I modeled these influences as additive noise, so that the response r is determined by both the effect of the system kernel k on the stimulus s , and random noise v : $r = k \otimes s + v$. These responses were correlated with the stimuli using the wavelet analysis, and the estimated impulse response was compared to the model impulse response by computing their correlation between 20 and 500 ms. As the noise level grows, estimates should deteriorate. Fig. 6 plots the estimate accuracy for the wavelet analysis after 100 iterations against the

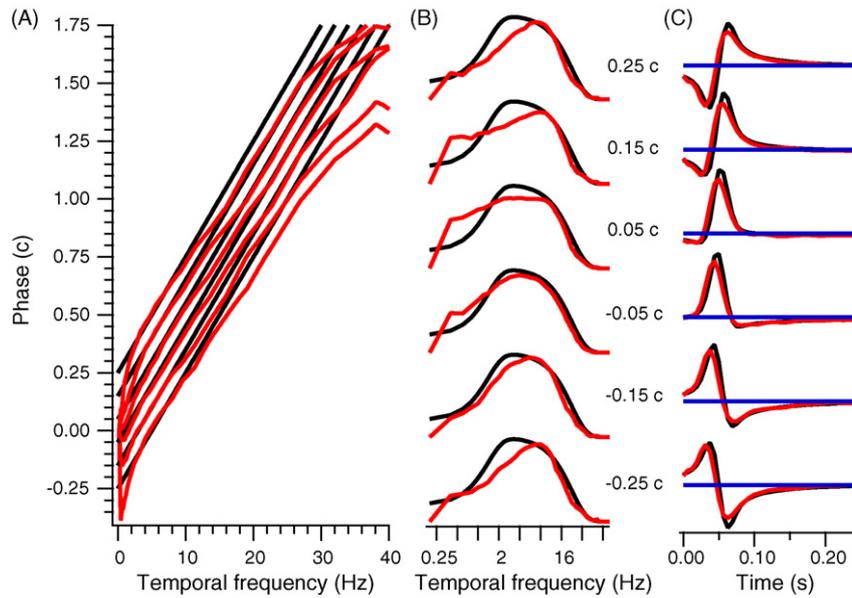


Fig. 3. Estimates of different kernels. Six model kernels are shown in black, differing only in absolute phase, which ranged from -0.25 to $+0.25$ c. (A) Phase vs. temporal frequency. (B) Amplitude tuning. (C) Impulse response functions. Estimates are shown in red, after 10 5 s trials for each estimate. Stimuli were Gaussian white noise, and responses were simple convolutions of the stimuli with the impulse responses.

signal-to-noise ratio. Parameters of model kernels were varied randomly, as described above. The signal-to-noise ratio is the ratio of the root mean square magnitude of $k \otimes s$ over 5 s to the root mean square magnitude of the added noise, which was varied. Estimates remained accurate until the signal-to-noise ratio dropped below 0.1. Data were fit with a Naka–Rushton function, $0.5 * a(1 + \tanh(b * \log(x/c)))$, where a is the saturation level, b reflects the rate of rise, and c is the signal-to-noise ratio at half-saturation. For white stimuli (red curves), the parameter values

for the best fit were $a = 0.93$, $b = 1.8$, $c = 0.008$. As discussed below, cells are normally tested with multiple stimuli presented simultaneously at different locations. Responses evoked by stimuli at one location appear to be noise when considering other locations. Resistance to noise is therefore an important property of an analysis technique.

Compared to the conventional correlation method, the wavelet method converged more rapidly, but did not achieve as much accuracy. The curves in Fig. 7 show the average correlations between the predicted and actual impulse response functions for each method as a function of the number of 5 s trials. Correlations were computed for different amounts of additive noise. When noise was added to the responses, the wavelet method was more accurate over the early iterations, but the standard correlation technique did better eventually, as well as when little noise was present. For low signal-to-noise ratios, as illustrated in Fig. 7 with the 0.125 level (purple traces), the wavelet

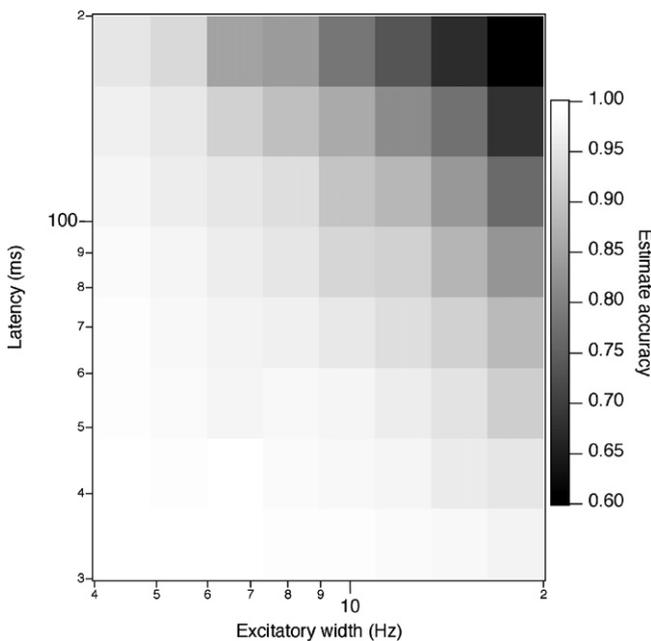


Fig. 4. Joint dependence of the accuracy of kernel estimates on excitatory width and latency. Estimate accuracy is plotted with a gray scale, binned as a function of the excitatory width parameter and the latency parameter.

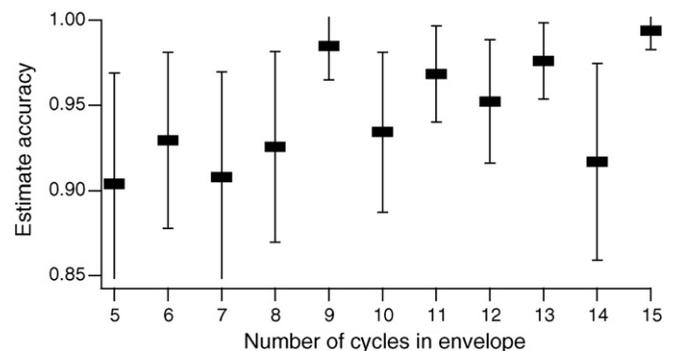


Fig. 5. Accuracy vs. localization. Estimate accuracy, the correlation between model and estimated impulse response functions, was calculated for 100 random kernels for each of 11 values of the ω_0 parameter. Means and their standard errors are shown.

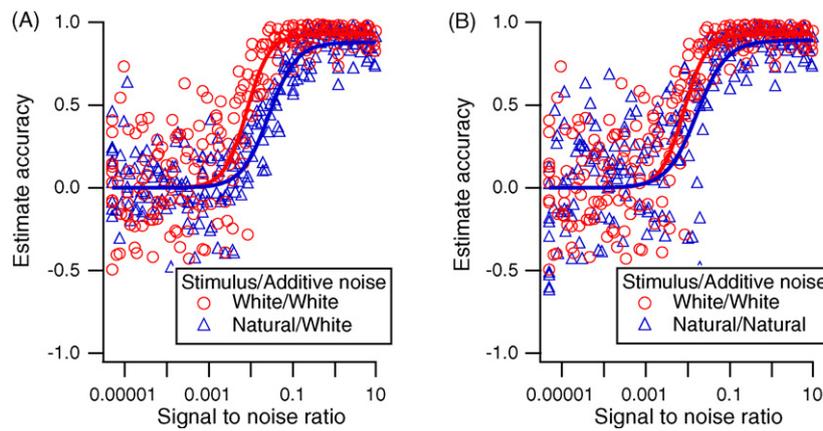


Fig. 6. Resistance to noise. Predictions of model impulse response functions were evaluated by the estimate accuracy measure. Noise was added to the responses after convolving the stimulus with the kernel. As noise levels increased, the signal-to-noise ratio and the estimate accuracy declined. Model kernels ($N=221$) were chosen randomly, and estimates were obtained from 100 iterations of 5 s trials. White stimuli generated red markers, and natural stimuli produced points marked in blue. In (A), the additive noise was white. In (B), for the natural stimuli the additive noise was also natural. Curve fits to each data set are shown as solid traces. Parameter values (saturation level, rate of rise, half-saturation signal-to-noise) for these fits were: (0.93, 1.8, 0.008) for white/white, (0.88, 1.5, 0.027) for natural/white, and (0.89, 1.4, 0.016) for natural/natural.

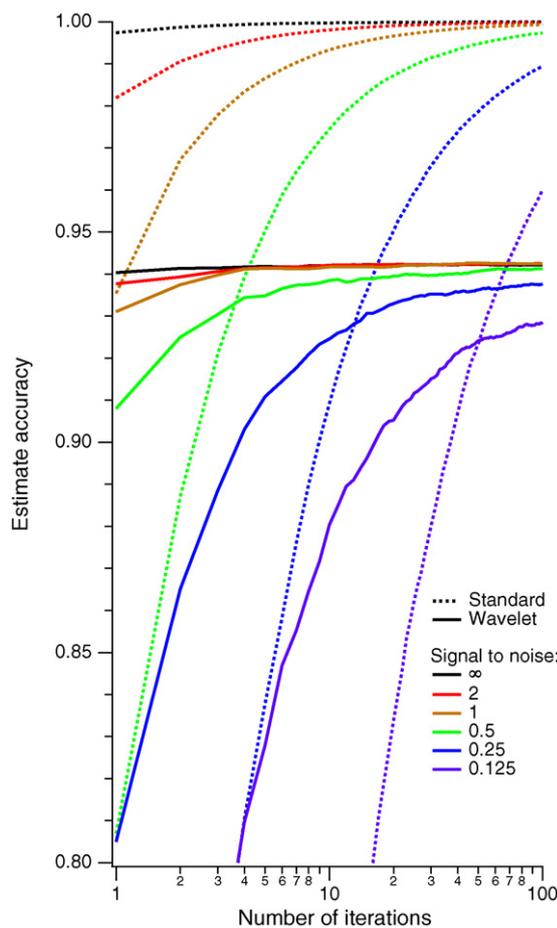


Fig. 7. Convergence and accuracy. Noisy responses to stimuli passed through different kernels were correlated with the stimuli using either conventional correlations (dotted traces) or wavelet-based correlations (solid traces). The accuracy of the kernel estimates was averaged over a set of 100 random kernels for each of 100 successive iterations. Six different signal-to-noise ratios were tested, as indicated by colors. Averages were compiled over the same 100 random kernels for each trace.

method was more accurate even beyond 40 iterations. The standard method attempts to “fit” the noise (Willmore and Smyth, 2003), as is especially obvious in the impulse response at late times (Fig. 12C). The amplitude of the additive noise had no effect on the asymptotic accuracy of the estimates, and had relatively little impact on the initial estimates for the wavelet method (when the signal-to-noise ratio was 2 the results were as good as the no-noise case). In summary, wavelet reverse correlations converge rapidly, are tolerant of noise, but contain intrinsic approximations that reduce their accuracy. Realistic situations with neuronal data contain noise and are often limited in the amount of data available, as noted below for several examples. Although the asymptotic accuracy is slightly lower with the wavelet method, the difference is minor, and is compensated by the more rapid convergence.

3.2. Nonwhite stimuli

The wavelet technique normalizes the spectra of the stimuli effectively. Fig. 8 shows an example where natural noise stimuli were used. That is, the stimulus on each trial was generated from the inverse Fourier transform of an amplitude spectrum given by $100/(1 + \omega)$ and a random phase spectrum. The wavelet transform of this stimulus in Fig. 8C can be compared with the white stimulus in Fig. 1D. Even though these stimuli were dominated by low temporal frequencies, only a few iterations (the illustrated case is the result of 10 iterations, that is, 50 s) were needed to obtain a good estimate of the model kernel. No changes were made to the calculations to enable estimates with arbitrary stimuli. When the stimulus was deprived of some frequency range, estimates deteriorated, but low amplitudes such as seen for high frequencies in Fig. 8 sufficed to maintain good predictions.

I compared the accuracy of estimates made with white and natural stimuli. Over 50 randomly chosen kernels, predictions

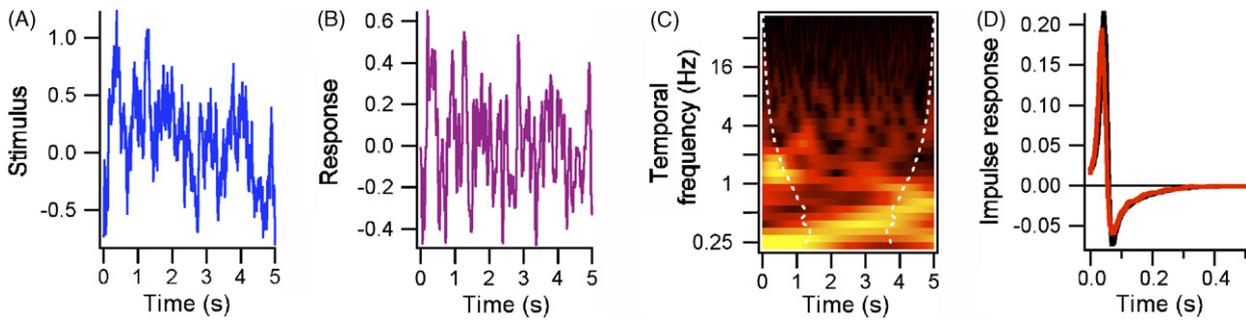


Fig. 8. Nonwhite stimuli. The test stimuli were natural noise samples, as illustrated in (A) and in (C) by its amplitude spectra across the trial. The response to this trial is shown in (B). The predicted impulse response is shown in (D), along with the model kernel, which is the same as that used in Fig. 1.

made with white stimuli had an accuracy of 0.98, as opposed to 0.96 with natural stimuli. This difference between white and nonwhite stimuli was significant ($t = 4.3, P < 0.001$), and was due to the weakness of the predictions at high frequencies for natural stimuli. The main finding, however, is that predictions made with natural stimuli are accurate. Predictions calculated with natural stimuli deteriorated more quickly in the presence of noise than did those from white stimuli (Fig. 6A). The half-saturation signal-to-noise parameter shifted more than an octave higher. This result is expected because the measure of signal-to-noise used here averages across frequencies, but for natural stimuli the noise has especially strong impact at high frequencies because the signal is weak there. In the situation where multiple positions are tested simultaneously (see below), when testing with natural stimuli the added noise will also be natural, as it consists of responses to the natural stimuli from other positions. In this case (Fig. 6B), less difference was observed between the white and natural stimuli in the presence of noise. The half-saturation point shifted by one octave. The additive noise was strongest at low frequencies, where the stimulus was strongest, and therefore did not outcompete the stimulus at higher frequencies as in Fig. 6A.

3.3. Nonlinearities

A full discussion of how wavelet correlations are used to estimate nonlinearities is beyond the scope of this report, but it is worth illustrating how first-order kernels can be estimated when subjected to static nonlinearities, that is, a nonlinearity that does not alter timing. Many static nonlinearities preserve

a monotonic input–output relationship, and therefore a strong linear component. A basic model for neuronal filtering involves a linear filter followed by rectification, since neurons do not respond with negative firing rates. For this model, the wavelet technique yields the linear filter, as shown in Fig. 9A. Half-squaring, meaning squaring the response after rectification, does not interfere considerably with the kernel estimate (Fig. 9B). As expected, full squaring eliminates the linear component that is estimated by the first-order kernel (Fig. 9C).

3.4. Non-stationary systems

The method used here performs the division of response by stimulus prior to averaging, $\langle R/S \rangle$, where S and R are the wavelet-transformed stimulus and response functions of frequency and time, and the average (denoted by $\langle \rangle$) is taken over time. This does not optimize kernel estimates in the sense of minimizing the error between the actual and estimated kernels (Fig. 7). For stationary systems, where the kernel does not vary over time, an alternative method using the wavelet transform is to compute $\langle RS^* \rangle / \langle SS^* \rangle$ as in the conventional spike-based method. This method was superior to the method where the division is performed first (Fig. 10, data fall mostly above the diagonal), with the advantage increasing with increasing noise levels. However, both methods performed well for kernels with short latencies and when signal-to-noise was high.

Some systems vary in their temporal properties because of stimulus or response history, or because of external influences. For instance, adaptation shifts timing in a highly specific manner

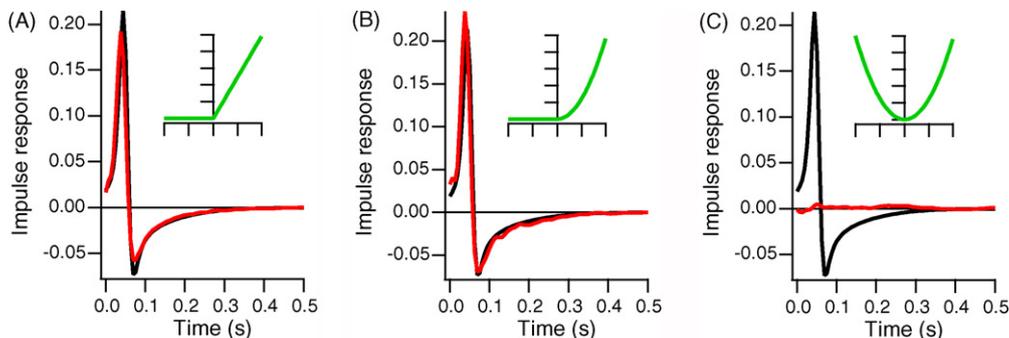


Fig. 9. First-order kernels estimated from linear/nonlinear models. The predictions are shown in red, with the model kernel in black. The nonlinearity that followed the linear kernel is shown in the insets. (A) Responses were given by $\max(0, 2k \otimes s)$ where k is the linear kernel and s is the stimulus, with \otimes denoting convolution. This is simple rectification. (B) Responses were generated by $\max(0, 2k \otimes s)^2$. This is known as half-squaring. (C) Responses were $(2k \otimes s)^2$, full squaring.

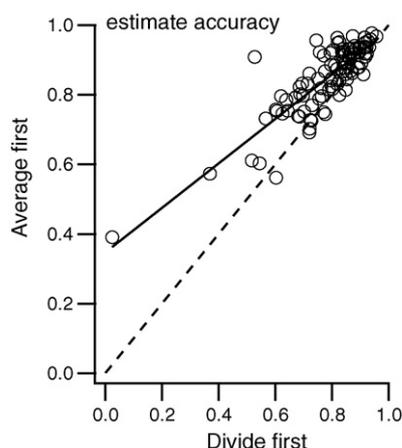


Fig. 10. Averaging methods. Accuracy of kernel estimates was measured for two methods of computing the kernels using wavelet transforms. The horizontal axis shows results using the method used elsewhere in this paper. The vertical axis averaged the stimulus-response product, then divided by the stimulus power. Over 100 random kernels, with a signal-to-noise ratio of 0.055, the regression line (solid line) had an intercept of 0.35 and a slope of 0.64, $r=0.85$.

in visual cortical neurons (Saul, 1995). For these non-stationary systems, it can be inappropriate to average over time. The wavelet method used here is easily modified to handle this case, by computing the quotient R/S at each time point, without averaging. The kernel estimates at each time can be compared across time in various ways (as shown in Fig. 15), including projecting them into subspaces that can reveal different inputs to the system, as in visual cortical complex cells (Ringach et al., 1997; Touryan and Dan, 2001; Schwartz et al., 2002; Touryan et al., 2002; Rust et al., 2005). In short, averaging can be avoided, and stimulus-response correlations can be derived over short time samples.

To illustrate the simplicity of the wavelet correlation analysis on non-stationary systems, Fig. 11 presents an example where two different kernels are applied to the stimulus during each

trial. For the first 2.5 s, kernel1 (solid red trace in (A)) was used to generate responses, and during the final 2.5 s of each trial, kernel2 (solid green trace in (A)) was used. The quotient between response and stimulus after wavelet transformation is shown in (B) (amplitude) and (C) (phase). This was computed without averaging over time during each trial. The time–frequency data were averaged across 100 trials. The change that occurs at 2.5 s is obvious in both of these plots, as the two kernels were chosen to be dissimilar, somewhat orthogonal. Kernel estimates are thus obtained at each of the 800 points in time. I performed some slight averaging over time by computing the mean of the complex values across 5 bins, which is 31.25 s, although this is not necessary. For each of the 160 remaining time points, the frequency-domain data were inverse Fourier transformed to obtain a real-valued impulse response function. For times between 0 and 625 ms, those 160 impulse response functions were then assembled into a matrix, with each row of the matrix one of the impulse response functions. A standard form of principal components analysis, singular value decomposition, provided orthogonal components that are similar to the original kernels (dotted traces in (A)). The estimates deviate from the model kernels because the model kernels are not actually orthogonal. In the general case, additional methods are used to project the 160 impulse responses into a subspace, where they often tend to cluster around the actual kernels in a structured way across time (in this case, the two halves of the trial). For visual cortical complex cells, clusters are often observed in rough antiphase along some direction in the subspace, corresponding to the overlapping ON and OFF responses.

3.5. Neuronal data

Cells in the lateral geniculate nucleus (LGN) and primary visual cortex (V1) were tested with various types of visual noise stimuli. Since we no longer have a model to compare to the results of the wavelet analysis, we compare different

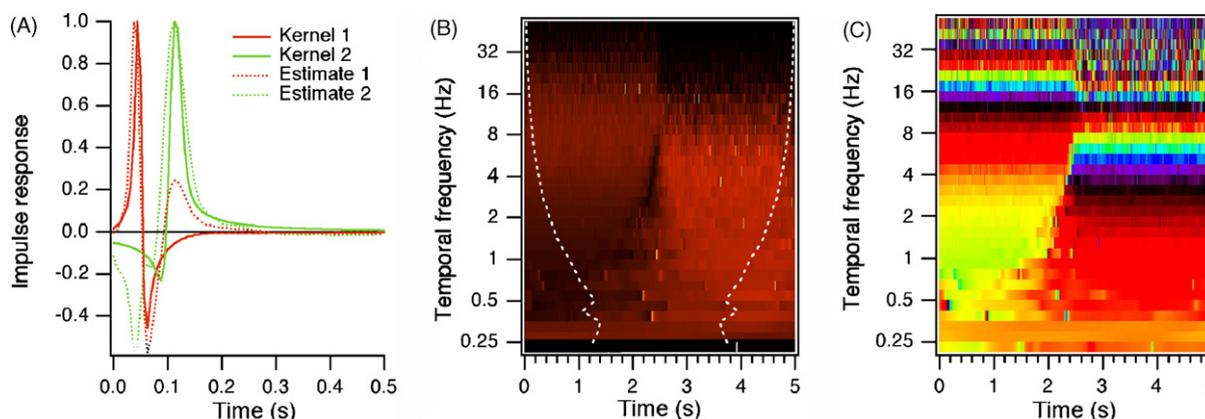


Fig. 11. Non-stationary system. During each 5 s trial, the response was given by the stimulus convolved with kernel1 (solid red trace in (A)) for the first 2.5 s, then with kernel2 (solid green trace) for the last 2.5 s. The wavelet correlation was performed without averaging across time during each trial, producing the amplitude and phase plots in (B and C). These frequency-domain data were averaged over 5 time bins (31.25 ms) and inverse transformed for each bin to generate a set of impulse response functions. Singular value decomposition of the matrix formed from those impulse responses provided the estimates shown (dotted traces in (A)), as the first two eigenvectors. The second kernel (green) had a long latency, which is reflected in the more rapid change in phase with frequency during the second half of the trial in (C). Latencies were 50 ms and 110 ms, absolute phase values were -0.15 and 0.1 c, and kernel2 was tuned to lower frequencies than kernel1. This caused its responses to bleed into earlier times at low frequencies.

techniques. First-order kernels were estimated with both conventional spike-triggered techniques and wavelet correlations. Rather than the single-input system considered above, the neuronal data were obtained with multi-input stimulation. That is, each of many positions across the receptive field was independently tested simultaneously. Because activity can be evoked from stimulation at one position while stimuli are also presented at another position, spurious correlations can occur. If the contrast of the stimulus for one of these spurious correlations is extremely small, the division of the response by the stimulus yields an especially large value. Many techniques can be employed to eliminate these artifacts. I used median filtering, where amplitudes that exceeded one standard deviation above the median over the surrounding 1000 ms were reset to the median value. This extra computation did not noticeably affect the kernel estimates illustrated above, because artifacts rarely occurred. However, the estimate accuracy when using median filtering was slightly lower than if median filtering was not used (0.982 vs. 0.987, $t = 8.8$, $N = 100$, $P < 0.001$).

Single neurons in LGN and V1 of awake behaving monkeys were isolated, and tested with noise stimuli. The stimuli consisted of bars arranged in either a one-dimensional row across the receptive field, with the bars oriented optimally, or a two-dimensional grid that covered the receptive field. On each frame (i.e. at 160 Hz), or on every other frame (80 Hz), each bar's luminance and/or color was set independently of the other bars. The temporal modulation applied at each position could be chosen from several types, such as binary (bright or dark) or ternary (bright, dark, and background) or Gaussian white noise, or natural noise (with correlations between the luminance across successive frames, but still without spatial correlations). On each 5 s trial, a random sequence was chosen based on an initial seed for the random number generator. I used sets of 10–20 seeds, with 3–10 repetitions of each sequence over the run. These repetitions were used in order to enable testing of reproducibility of responses to the random sequences. Spikes were timed with high precision, but added to histograms with 6.25 ms bins.

To validate the wavelet technique, I show examples of comparisons between spike-based (i.e. normalized spike-triggered averages, computed by correlating responses with stimuli in the frequency domain) and wavelet-based correlations for neurons stimulated with these noise patterns. An LGN cell that was excited by blue stimuli in its receptive field center, and inhibited by yellow stimuli, is illustrated in Fig. 12. It was tested with an array of square bars, with the color on every other frame chosen from a uniform distribution across the monitor's RGB space. That is, the color consisted of a trio of values that each ranged between 0 and 255, with (0, 0, 0) being black, (255, 255, 255) being bright white, (255, 0, 0) being bright red, etc. At each position, responses to each of the monitor's guns (R, G, and B) were derived separately, applying both a conventional spike-based reverse correlation method and the wavelet correlation method. The maps in Fig. 12A show the red, green, and blue responses as a function of space at the peak response time of 37 ms. The black areas in the middle of the red and green maps indicate that the cell responded when the red and green guns were given low values. This is equivalent to saying that the

cell was suppressed by yellow stimuli, or that it responded to the offset of yellow stimuli.

Fig. 12B provides more detailed maps, showing the time-course of the responses at each position. The position with the strongest peak response is highlighted with colors. Downward on these traces indicates that responses occurred to darker values. The small responses to bright blue can be glimpsed in the plot on the right. The vertical scale bars vary in height between these 3 plots, with the green map having the highest amplitudes, followed by the red then the blue plots. This is primarily caused by the fact that the green monitor phosphor emits many more photons than the red and blue phosphors for the same RGB value. In particular, the short-wavelength cones that presumably excite this cell are relatively weakly stimulated by the blue gun compared to how the red and green guns stimulate the medium- and long-wavelength cones (Horwitz et al., 2005).

The main point is that the kernel estimates are structured in space and time as expected. Strong responses were evoked only from the receptive field center, where consistent timing was observed over just a few pixels. The rest of the tested positions showed only weak responses. Spurious correlations were avoided by median filtering without sacrificing the ability to see the valid correlations.

The maps in Fig. 12A and B were derived using wavelet-based correlations. Fig. 12C illustrates the spike-triggered average responses at the peak position (the one highlighted by colors in (B)). The corresponding wavelet-based impulse responses are shown in (D). These curves are similar to each other (as were the responses at all positions), showing that the cell had sustained OFF responses to red and green, and a lagged ON response to blue. The spike-based profiles have DC offsets that were explicitly eliminated during the wavelet-based calculations. Eliminating this offset and smoothing the spike-based curves makes them resemble the wavelet-based curves more closely. In particular, the timing is almost identical for the predictions of these analysis methods.

The wavelet method normalizes the stimulus temporal spectrum, as illustrated in Fig. 8. This is especially clear with real data. Testing cells with natural stimuli that have spectra that fall off in amplitude with increasing frequency can cause problems for conventional analysis methods (Baddeley et al., 1997; Gallant et al., 1998; Smyth et al., 2003; Theunissen et al., 2000, 2001; Ringach et al., 2002; Willmore and Smyth, 2003). Fig. 13 compares impulse responses derived either from wavelet correlations or from spike-based correlations. Two successive runs were obtained from the same cell with all parameters identical except for the temporal statistics of the stimulus. In one run, luminance was modulated in a ternary white sequence, and in the other run it was modulated in a natural sequence. The particular sequence chose the value of luminance based on the 2 previous values: $val[n] = 0.6 * val[n - 1] + 0.15 * (val[n - 1] - val[n - 2]) + 0.1 * u$ where u is a uniformly distributed random value between -1 and 1 . Contrast was then enhanced by applying a sigmoidal function to the value.

The parvocellular LGN neuron in Fig. 13 gave sustained OFF responses, with a latency to peak of 42 ms. The two analysis

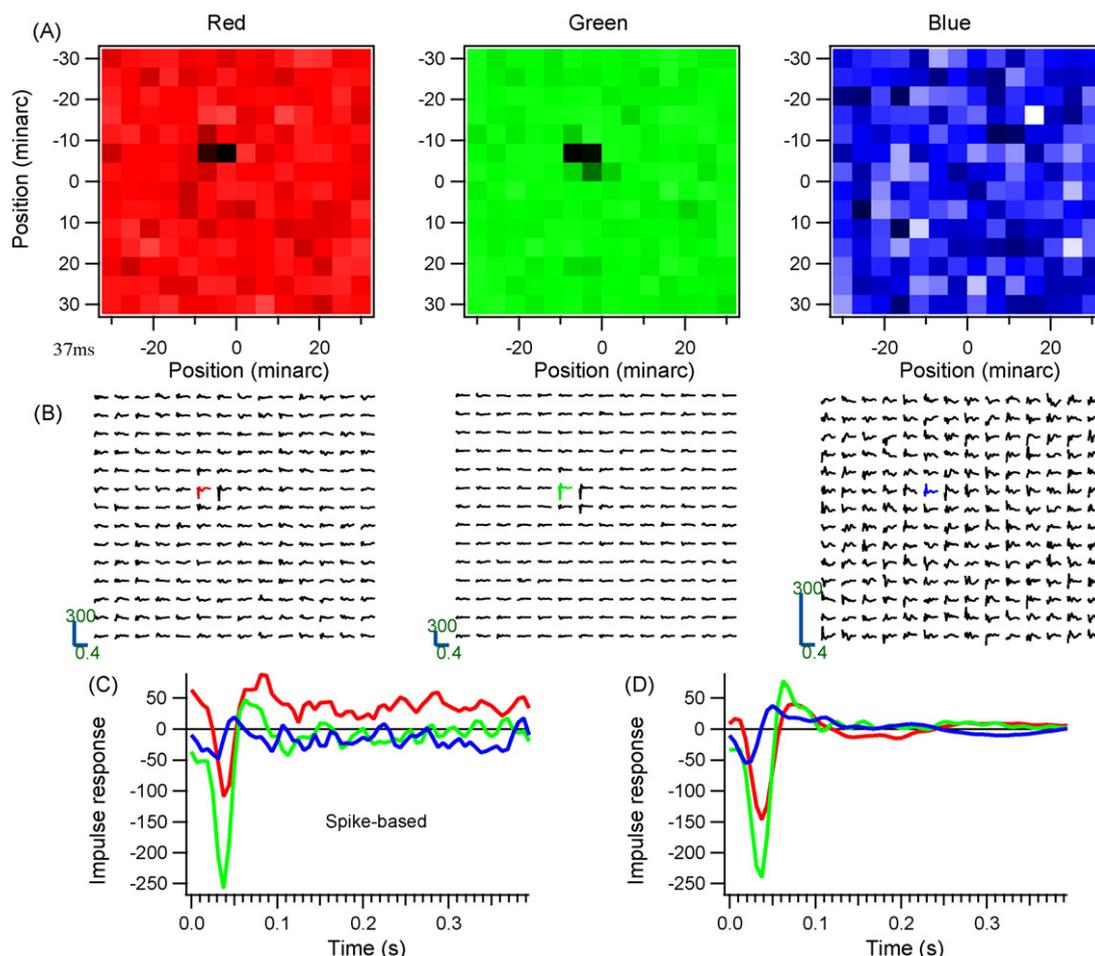


Fig. 12. Comparison of kernels estimated by standard spike-based reverse correlation and wavelet correlation. Responses to red, green, and blue stimuli from a Blue-ON cell in awake monkey LGN are shown, computed from spike-triggered averages (C) and from wavelet correlations (A, B, and D). Slices of the data at 37 ms are shown in (A), with both the vertical and horizontal axes spanning 1° . Impulse response functions derived from the wavelet method are shown for each position and color in (B). Colored traces are from the pixel with the strongest response. Impulse responses from this position are shown magnified in (C and D). The scaling on the vertical axis has units of spikes per second per full range of luminance modulation for each color, with negative values corresponding to decreased luminance. This run was 79 trials long, or 395 s.

methods yielded almost identical results for the ternary white noise (red and blue solid traces in Fig. 13) as well as for the natural noise (green and black dashed traces). The correlations in the stimulus were normalized, and both methods gave estimated kernels that closely resemble those obtained from white noise.

As another example, consider a V1 neuron tested with spatially one-dimensional noise stimuli. Optimally oriented bars positioned at 16 points across the receptive field were modulated either with a binary white sequence or the same sort of natural sequence as described above. This cell was direction selective, and showed some spatiotemporal orientation corresponding to this property. The maps in Fig. 14 show kernels as a function of space (horizontal axis) and time (vertical axis). Panels (A and B) show the results from the binary white noise run. Similar results were obtained with the two analysis methods. For the natural noise run in (C and D), the spike-based method gave noisier results, and did not pick up the portion of the kernel around -10 minarc. The wavelet method, however, produced a clear, well-oriented map with at least as good resolution as seen in the white noise maps.

Many systems are not stationary. Kernels can vary over time because of external events, for example. The wavelet method facilitates the analysis of such event-related behaviors. An example is how visual receptive field kernels vary around saccades. Even while fixating, small saccades are made, and I asked how timing in LGN cells varies with respect to when saccades occur (testing theoretical predictions such as in Dong et al., 2003). This is addressed by averaging the ratio of the wavelet-transformed response and stimulus as a function of time relative to each saccade (Fig. 15A and B). Averages were computed for brief time intervals at different points around saccades. Kernels were then derived at each point in time around saccades (Fig. 15C). The phase plot in Fig. 15A contains an obvious break near 4 Hz around the time of saccades, and phase values before and after saccades differ from each other at some frequencies. This is reflected in the impulse response functions (Fig. 15C) that shift in phase across the saccade. Far from the times of saccades (± 0.9 s), the impulse responses are similar to each other. Just before saccades (-0.1 s), the impulse response degenerates, suggesting that the effect is not purely due to stimulus movement

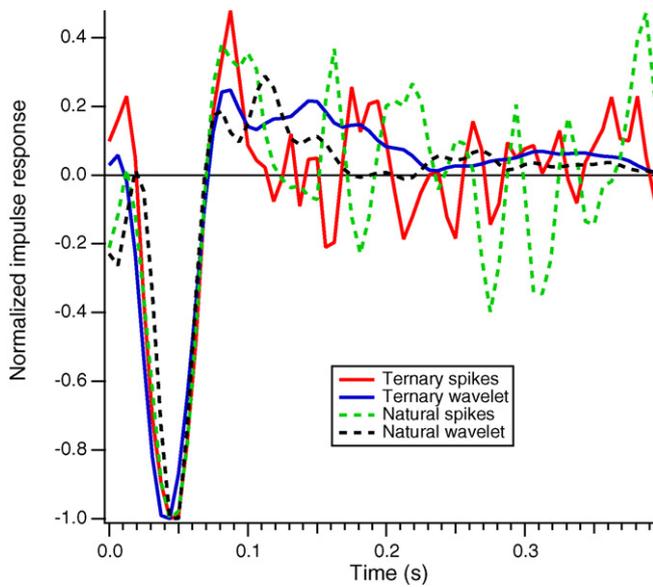


Fig. 13. Impulse response functions derived from spike-based and wavelet-based correlations. A parvocellular LGN neuron was tested with ternary white noise and natural noise in separate runs. The illustrated profiles were measured at the pixel with the strongest responses, i.e. the receptive field center. These impulse response functions were normalized to the same lowest point, which occurred at 42 ms in all cases. The ternary run had 24 trials, and the natural run had 43 trials.

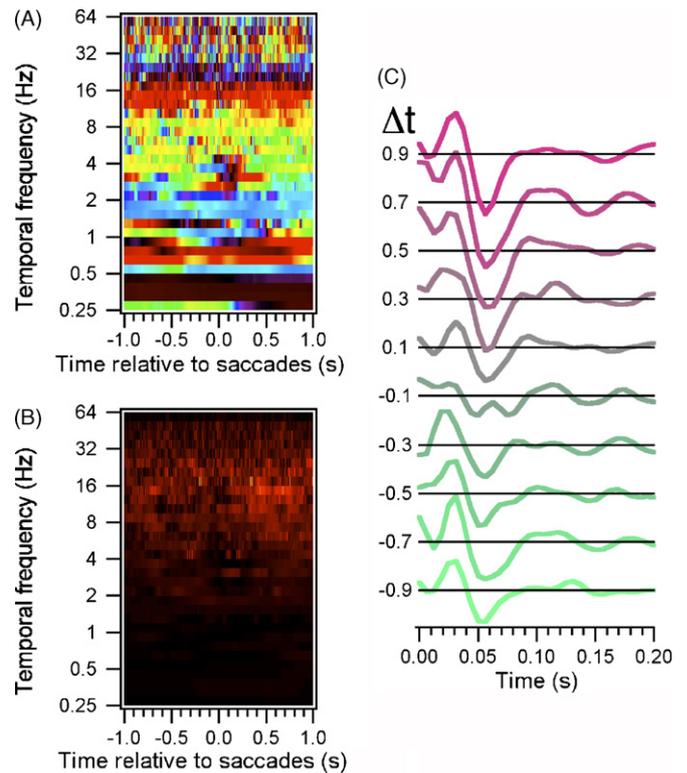


Fig. 15. Saccade-triggered averaging. The ratio of response to stimulus after wavelet transformation was accumulated as a function of distance in time from each saccade. The pixel with the strongest response from an LGN cell's receptive field was used for these temporal measurements. The stimulus was natural noise, as in Fig. 13. Phase (A) and amplitude (B) of the kernels are shown as functions of time from saccades (horizontal axis) and temporal frequency (vertical axis). Color scales are as in Figs. 1, 8 and 11. (C) Impulse response functions averaged over 200 ms intervals at a series of distances in time from saccades. The numbers at the left give the center of each averaging interval, in seconds. The run had 32 trials.

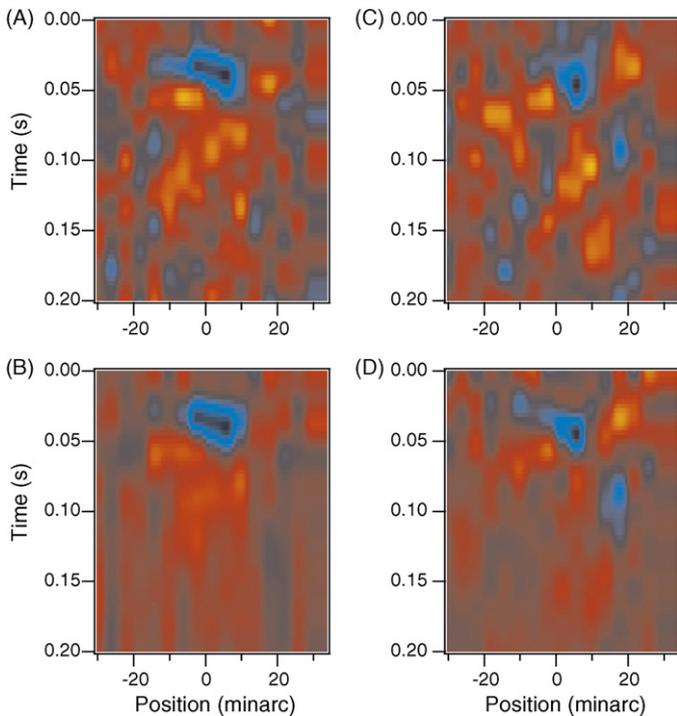


Fig. 14. Comparing spike-based and wavelet-based analyses. A direction selective V1 cell was tested with binary and natural noise in consecutive runs. These were each analyzed with either spike-based reverse correlation or with wavelet correlations. These maps show space on the horizontal axis and time on the vertical axis, with cold colors representing dark-excitation and warm colors bright-excitation. (A) Binary noise analyzed with spike correlations. (B) Binary noise analyzed with wavelet correlations. (C) Natural noise analyzed with spike correlations. (D) Natural noise analyzed with wavelet correlations. The binary noise run had 24 trials and the natural noise run had 36 trials.

on the retina. After the saccade (0.1–0.7 s), the second phase of the impulse response dominates. Again, this does not reflect movement across the retina, since during most of the times involved (e.g., 200–400 ms after saccades for the +0.3 s case) the eyes were nearly stationary, and small retinal movements were compensated by moving the stimulus to match. Although similar calculations could be made by computing spike-triggered averages contingent on distance from saccades, the wavelet method is well-suited to these sorts of analyses that depend on time.

4. Discussion

Another method for estimating receptive field structure is described here. The wavelet transform isolates, in time, the spectral components of the stimulus and response that are to be correlated. Arbitrary stimuli can be applied, and response noise, including that evoked by multi-input stimulation, is handled well. Perhaps most importantly, results are obtained with relatively limited amounts of data. The disadvantages include inaccuracy of the estimates due to approximations made by not

localizing in either the frequency domain or the time domain, as well as long computation times. It typically takes about 3 times as long to compute the wavelet correlations compared to the spike-based correlations.

The wavelet transform provides a representation of its input that is convenient, but that can be problematic in certain respects. For instance, input at a single frequency is spread out over neighboring frequencies. The representation is redundant, which is useful numerically but makes theoretical analysis less tractable. For stimuli that are somewhat sparse in the frequency domain, conventional methods require additional techniques to avoid singularities, but this is built into the redundancy of the wavelet method.

With conventional reverse correlation, normalization is required when using nonwhite stimuli, and regularization is sometimes applied to reduce the noise that is evident in the examples above. Typically, physiologists have not routinely applied either technique when performing reverse correlation. If algorithms are modified to include normalization and regularization, accuracy of kernel estimates can exceed that of the wavelet method, with faster computational speeds. The wavelet technique converges more quickly, and is easy to implement in a general form. It can also be used to derive non-stationary system behavior in a straightforward way. An example where this could be useful is in determining the behaviors of dynamic synapses, investigating the effects of synaptic depression and facilitation under physiological conditions.

The wavelet method can be applied to analog data. The simulations shown in this report used analog stimuli and responses. A real-world example is in measurement of multifocal electroretinograms (mfERGs). Recordings of retinal responses are obtained from a corneal electrode. The mass responses from the whole retina are then decomposed to reveal localized function across the retina. Multifocal ERGs have been obtained primarily with m-sequences (Sutter, 2001). Nonwhite stimuli can be applied as well, and analyzed with wavelet correlations.

In summary, a simple technique permits rapid estimation of first-order kernels from measurements of responses to arbitrary stimuli. The stimulus and the response are both transformed into their instantaneous frequency-domain representations at each point in time, via a wavelet transform. The kernel is then obtained by dividing response by stimulus, filtering the results to eliminate artifacts, and averaging across the time samples. Working in the frequency domain, and in particular computing the phase differences between response and stimulus, provides accurate reconstructions with little effort. Nonstationarity is managed by using the multiple time samples. Physiologists can easily perform system identification in this manner across a broad spectrum of preparations.

References

- Aertsen AMHJ, Johannesma PIM. The spectro-temporal receptive field. A functional characterization of auditory neurons. *Biol Cybern* 1981;42:133–43.
- Baddeley R, Abbott LF, Booth MC, Sengpiel F, Freeman T, Wakeman EA, et al. Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc Biol Sci* 1997;264:1775–83.
- Christakos CN, Rost I, Windhorst U. The use of frequency domain techniques in the study of signal transmission in skeletal muscle. *Eur J Physiol* 2004;400:100–5.
- David SV, Vinje WE, Gallant JL. Natural stimulus statistics alter the receptive field structure of V1 neurons. *J Neurosci* 2004;24:6991–7006.
- deCharms RC, Blake DT, Merzenich MM. Optimizing sound features for cortical neurons. *Science* 1998;280:1439–43.
- Depireux DA, Simon JZ, Klein DJ, Shamma SA. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. *J Neurophysiol* 2001;85:1220–34.
- Dong DW, Simpson GB, Weyand TG. No suppression, only dynamic decorrelation: saccadic effects on the visual responses to natural time-varying images. *J Vision* 2003;3:45a.
- Enroth-Cugell C, Robson JG. The contrast sensitivity of retinal ganglion cells of the cat. *J Physiol (Lond)* 1966;187:517–52.
- French AS. Practical nonlinear system analysis by Wiener kernel estimation in the frequency domain. *Biol Cybern* 1976;24:111–9.
- Gallant JL, Connor CE, Van Essen DC. Neural activity in areas V1, V2 and V4 during free viewing of natural scenes compared to controlled viewing. *Neuroreport* 1998;9:2153–8.
- Horwitz GD, Chichilnisky EJ, Albright TD. Blue-yellow signals are enhanced by spatiotemporal luminance contrast in macaque V1. *J Neurophysiol* 2005;93:2263–78.
- Li X, Xin Y, Fox J, Jefferys JG. Interaction dynamics of neuronal oscillations analysed using wavelet transforms. *J Neurosci Meth* 2007;160:178–85.
- Luczak A, Hackett TA, Kajikawa Y, Laubach M. Multivariate receptive field mapping in marmoset auditory cortex. *J Neurosci Method* 2004;136:77–85.
- Marmarelis P, Marmarelis V. Analysis of physiological systems: the white noise approach. New York: Plenum; 1978.
- Nishimoto S, Ishida T, Ohzawa I. Receptive field properties of neurons in the early visual cortex revealed by local spectral reverse correlation. *J Neurosci* 2006;26:3269–80.
- Prenger RJ, Wu MCK, David SV, Gallant JL. Nonlinear V1 responses to natural scenes revealed by neural network analysis. *Neural Netw* 2004;17:663–79.
- Ringach DL, Sapiro G, Shapley R. A subspace reverse-correlation technique for the study of visual neurons. *Vision Res* 1997;37:2455–64.
- Ringach DL, Hawken MJ, Shapley R. Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *J Vis* 2002;2:12–24.
- Rust NC, Schwartz O, Movshon JA, Simoncelli EP. Spatiotemporal elements of macaque V1 receptive fields. *Neuron* 2005;46:945–56.
- Saul AB. Adaptation aftereffects in single neurons of cat visual cortex: response timing is retarded by adapting. *Visual Neurosci* 1995;12:191–205.
- Schwartz O, Chichilnisky EJ, Simoncelli EP. Characterizing neural gain control using spike triggered covariance. *Adv Neural Inf Process Syst* 2002;14:269–76.
- Schwartz O, Pillow JW, Rust NC, Simoncelli EP. Spike-triggered neural characterization. *J Vision* 2006;6:484–507.
- Smyth D, Willmore B, Baker GE, Thompson ID, Tolhurst DJ. The receptive-field organization of simple cells in primary visual cortex of ferrets under natural scene stimulation. *J Neurosci* 2003;23:4746–59.
- Soucek J, Dudok de Wit T, Dunlop M, Décréau P. Local wavelet correlation: application to timing analysis of multi-satellite CLUSTER data. *Ann Geophysicae* 2004;22:4185–96.
- Sutter EE. Imaging visual function with the multifocal m-sequence technique. *Vision Res* 2001;41:1241–55.
- Tang Y, Saul A, Gur M, Goei S, Wong E, Ersoy B, et al. Eye position compensation improves estimates of response magnitude and receptive field geometry in alert monkeys. *J Neurophysiol* 2007;97:3439–48.
- Theunissen FE, Sen K, Doupe AJ. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci* 2000;20:2315–31.
- Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, Gallant JL. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network* 2001;12:289–316.

- Torrence C, Compo GP. A practical guide to wavelet analysis. *Bull Am Meteorol Soc* 1998;79:61–78.
- Touryan J, Dan Y. Analysis of sensory coding with complex stimuli. *Curr Opin Neurobiol* 2001;11:443–8.
- Touryan J, Lau B, Dan Y. Isolation of relevant features from random stimuli for cortical complex cells. *J Neurosci* 2002;22:10811–8.
- Victor JD. Nonlinear systems analysis: comparison of white noise and sum of sinusoids in a biological system. *Proc Natl Acad Sci USA* 1979;76:996–8.
- Victor JD, Shapley RM. Receptive field mechanisms of cat X and Y retinal ganglion cells. *J Gen Physiol* 1979;74:275–98.
- Vrhel M, Lee C, Unser M. Rapid computation of the continuous wavelet transform by oblique projections. *IEEE Trans Signal Process* 1997;45:891–900.
- Willmore B, Smyth D. Methods for first-order kernel estimation: simple-cell receptive fields from responses to natural scenes. *Network* 2003;14:553–77.