Probing feature selectivity of neurons in primary visual cortex with natural stimuli.

T. Sharpee[1-3], H. Sugihara[2,3], A.V. Kurgansky[2,3], S. Rebrik[2,3], M. P. Stryker[1-3] and K. D. Miller[1-3].
Sloan-Swartz Center for Theoretical Neurobiology[1], Department of Physiology[2], Keck Center for Integrative Neuroscience[3], University of California, San Francisco, CA 94143-0444

ABSTRACT

One way to characterize neural feature selectivity is to model the response probability as a nonlinear function of the output of a set of linear filters applied to incoming signals. Traditionally these linear filters are measured by probing neurons with correlated Gaussian noise ensembles and calculating correlation functions between incoming signals and neural responses. It is also important to derive these filters in response to natural stimuli, which have been shown to have strongly non-Gaussian spatiotemporal correlations. An information-theoretic method has been proposed recently for reconstructing neural filters using natural stimuli in which one looks for filters whose convolution with the stimulus ensemble accounts for the maximal possible part of the overall information carried the sequence of neural responses. Here we give a first-time demonstration of this method on real neural data, and compare responses of neurons in cat primary visual cortex driven with natural stimuli, noise ensembles, and moving gratings. We show that the information-theoretic method achieves the same quality of filter reconstruction for natural stimuli as that of well-established white-noise methods. Major parameters of neural filters derived from noise ensembles and natural stimuli, as well as from moving gratings are consistent with one another. We find that application of the reverse correlation method to natural stimuli ensembles leads to significant distortions in filters for a majority of studied cells with non-zero reverse-correlation filter.

Keywords: Primary visual cortex, information theory, white noise, receptive fields, reverse correlation.

1. INTRODUCTION

Neural feature selectivity has traditionally been studied using stimuli that can be described by a small number of parameters, so that the input/output function can be measured with sufficient detail. The classic example is the use of bars and gratings to study neurons in primary visual cortex. Such an approach is contingent on the fact that neurons are responsive to the class of stimuli used [1,2]. Finding an appropriate class of stimuli may require a serendipitous discovery, such as the one made by Hubel and Wiesel for neurons in primary visual cortex [1]. However it is not clear even at the next stage of visual signal processing, what is the most appropriate stimuli class to be used (see, however, Ref. [3,4]). An alternative approach is to use a broad range of stimuli, and, through some sort of statistical averaging, to find out what filter is being used by a particular neuron [5-14]. From the statistical analysis point of view, it is most advantageous to use stimuli with known statistical properties, such as (possibly correlated) Gaussian noise. While it may suffice for the first few stages of visual signal processing, neurons at higher level, whose structure we are most interested in, are left unresponsive by noise ensembles. Neurons at all stages of signal processing are responsive to "natural" stimuli, i.e. those stimuli that are derived from our surroundings. It turns out that such stimuli have specific statistical properties, which vary surprisingly little with a change of environment from which they are derived [15-20]. The most robust features are strong spatiotemporal correlations with characteristic $1/f^2$ spectrum, where f is a spatial frequency, scaling of the probability distribution as a whole, and strong deviation of the probability distribution from a Gaussian one. The non-Gaussian character of correlations invalidates many methods developed to study neural feature selectivity [21]. However, it may be possible to use ideas form information theory to bypass the problem of non-Gaussian effects in natural stimuli, thus allowing their use in large-scale systematic neurophysiological recordings [21].

We decided to test the newly proposed information-theoretic method [21] on neurons in primary visual cortex, whose properties are well studied. Each of the neurons has been probed by white noise and natural stimuli ensembles, so that we can compare the filters obtained with the information-theoretic method to the filters derived from the neuron's responses to white noise. In addition, neural responses where probed with moving (sinusoidal or square) gratings to determine their optimal orientation and spatial frequency. To analyze neural responses to either white noise or natural stimulus ensembles we use the framework of the so-called linear-nonlinear model, according to which the probability of

a response – a spike in our case – given a stimulus $\vec{s}$ is some nonlinear function (whose shape is not *a priori* specified) of projection of the stimulus $\vec{s}$ onto a set of filters $\hat{e}_i$, i=1…k, :

$$P(\text{spike} \mid \vec{s}) = P(\text{spike})\text{f}(\vec{s} \cdot \hat{e}_1, ..., \vec{s} \cdot \hat{e}_k) \tag{1}$$

The most important simplification made in Eq. (1) is that the number of filters k used by the neuron is much less than the dimensionality of the stimulus itself. Even though neurons may have many filters, in this paper we will present the results of simulations for neurons with up to two filters. In analyzing data from neurons in primary visual cortex, we carry out the analysis as if they had just one filter[1]. We concentrate on cells for which the white-noise filter is non-zero. This biases our subset of cells towards simple cells, which are thought to be well described by one filter. However our primary goal in this work is to demonstrate the use of the information-theoretic method on neural data and to verify that the neural filters it provides are consistent with other independent measurements. Under these circumstances the use of mostly simple cells is desirable.

2. METHODS

Experimental recordings. All experimental procedures where conducted under a protocol approved by the University of California, San Francisco Committee on Animal Research and have been described in detail elsewhere [22,23]. Briefly, spike trains were recorded using tetrode electrodes from primary visual cortex of anesthetized adult cats. The visual stimuli $\vec{s}$ were presented as two-dimensional spatiotemporal patterns of light intensities on a video monitor with a refresh rate of 120 Hz. The frame update rate was 60 Hz in the case white noise stimulus ensemble and 30 Hz in the case of a natural stimulus ensemble. In addition cells were probed with full-field moving periodic patterns (gratings). The optimal orientation was determined from cell's responses to a set of 10° spaced orientations and spatial frequency of 0.5 cycles/degree and temporal frequency of 2 Hz. The optimal spatial frequency for a given cell was derived from its responses to a set of moving gratings of optimal orientation and variable spatial frequencies (linearly spaced between 0.1 and 4 cycles/degree). A cell was classified as simple or complex depending on the ratio of the first Fourier component F_1 of its response to a moving grating of optimal orientation and spatial frequency to its mean F_0: a cell was considered simple if $F_1/F_0>1$ and an complex, otherwise. To estimate the information carried by the arrival of one spike, I_{spike}, an 11sec long segment of the natural movie was repeated from 50 -150 times.

Filter reconstruction from neural responses to a white-noise ensemble – reverse correlation method. Generally, the problem of reconstructing neural responses consists of two parts, one of which is to find the underlying linear filters and the other one is to find the nonlinear function $\text{f}(\vec{s} \cdot \hat{e}_1, ..., \vec{s} \cdot \hat{e}_k)$ in Eq. (1) that determines the probability of a spike given stimulus. If stimuli are drawn from a Gaussian distribution, then it is possible to use Wick's theorem to show that the problem of finding the linear filters separates from that of finding the nonlinearity associated with them. In particular, in the case of a white noise ensemble, one may obtain one of the linear filters by computing a correlation function between spikes and stimuli:

$$\hat{e}_{\text{STA}} = \left\langle \vec{s} P(\text{spike} \mid \vec{s}) \right\rangle - P(\text{spike}) \left\langle \vec{s} \right\rangle, \tag{2}$$

which as may be seen from Eq. (2) is computed by taking the average stimulus weighted by the number of spikes it elicits and subtracting the average stimulus multiplied by the overall number of spikes, the procedure also known as the spike-triggered average (STA). The magnitude of the filter is irrelevant, since rescaling of the vector can be accounted for by the appropriate rescaling of the input/output function (1) which converts projections on the relevant filter into spike probability. Therefore, we will normalize all of the derived filters to unit length.

Once the filter has been determined, the nonlinear input/output function can be mapped directly from experimental data, using the Bayes' rule:

$$f(\vec{s} \cdot \hat{e}_{\text{STA}}) = \frac{P(\text{spike} \mid \vec{s})}{P(\text{spike})} = \frac{P(\text{spike} \mid \vec{s} \cdot \hat{e}_{\text{STA}})}{P(\text{spike})} = \frac{P(\vec{s} \cdot \hat{e}_{\text{STA}} \mid \text{spike})}{P(\vec{s} \cdot \hat{e}_{\text{STA}})} \tag{3}$$

---

[1] As might be expected such an approximation is better for some classes of cells than others. For neurons in primary visual cortex, responses of simple cells are thought to be described by one filter, while responses of complex cells require at least two filters for their description. Accordingly we find that the amount of information that can be accounted for by projections onto just one linear filter is significantly higher for simple cells as opposed to complex. The detailed comparison of simple vs. complex cells responses will be given in subsequent work.

We will be using this expression to map the input/output function for any filter found and any stimulus ensemble used.

Filter reconstruction from neural responses to a correlated Gaussian ensemble. In the presence of correlations within the stimulus ensemble, the reverse correlation function (2) is biased towards those Fourier components that occur most often. Therefore the neural filter is found by normalizing expression (2) by the covariance matrix of the stimulus ensemble used:

$$\hat{e}_{\text{Gaussian estimate}} = C^{-1}\left\langle \vec{s}P(\text{spike}\mid\vec{s})\right\rangle - P(\text{spike})\left\langle\vec{s}\right\rangle, \qquad C_{ij} = \left\langle s_i s_j\right\rangle - \left\langle s_i\right\rangle\left\langle s_j\right\rangle, \qquad (4)$$

where notation $\hat{e}_{\text{Gaussian -estimate}}$ notes the fact the filter has been obtained assuming Gaussian statistics for the input signals.

Filter reconstruction from neural responses to a natural stimuli ensemble – most informative dimensions. In cases where the probability distribution of stimuli deviates significantly from a Gaussian one, assuming an incorrect form of the nonlinearity leads to incorrect estimates for the linear filters themselves. One possibility is to re-estimate the nonlinearity using the Bayes' rule of Eq. (3) for each of the trial directions in the stimulus space and find the linear filter used by the neuron through an optimization procedure. This is in effect what is done when different directions $\vec{v}$ in the stimulus space are evaluated according to the Kullback-Liebler distance between the probability distributions of projections $P_{\vec{v}}(x) = \left\langle\delta(x - \vec{s}\cdot\vec{v})\right\rangle_{\vec{s}}$ and $P_{\vec{v}}(x\mid\text{spike}) = \left\langle\delta(x - \vec{s}\cdot\vec{v})\mid\text{spike}\right\rangle_{\vec{s}}$ in the *a priori* and conditional ensembles, respectively [21,24]:

$$I(\vec{v}) = \int dx P_{\vec{v}}(x\mid\text{spike})\log_2\left[\frac{P_{\vec{v}}(x\mid\text{spike})}{P_{\vec{v}}(x)}\right], \qquad (5)$$

where the subscript $\vec{v}$ in probability distributions notes the fact the they depend on vector $\vec{v}$ along which they were evaluated. The information $I(\vec{v})$ along a particular direction $\vec{v}$ quantifies the relevance of this direction for generating spikes. The ratio of the probabilities distributions $P_{\vec{v}}(x\mid\text{spike})/P_{\vec{v}}(x)$ in (5) gives an estimate of the input/output function $f(\vec{s}\cdot\vec{v})$ of Eq. (3). If no change occurs in the probability distribution $P_{\vec{v}}(x)$ upon observation of a spike, then $I(\vec{v}) = 0$ and the corresponding direction $\vec{v}$ is irrelevant for spike generation. The linear filter used by a particular neuron is determined through an optimization procedure where information $I(\vec{v})$ is maximized as a function of the vector $\vec{v}$. The presence of local maxima necessitate the use of Monte Carlo methods, and we implement a simulated annealing scheme, where information is maximized in directions of successive gradients of information (5). Vectors with lower information values are accepted with Boltzman probabilities. We first present the results of the optimization algorithm in reconstructing filters for model neurons, which have that advantage that their filters are known. It is essential that the deviations between the reconstructed and model filters decrease with the number of spikes, the parameter that controls signal-to-noise ratio. Therefore the observed deviations in the reconstructed filter are on average due to noise and not to a failure of the optimization procedure, see below.

Upon completion of the optimization scheme, we will compare the information $I(\vec{v})$ along a particular direction $\vec{v}$ to the overall information carried by the arrival of one spike I_{spike}. The ratio of the two quantities describes the quality of reconstruction. Fortunately, the overall information can be estimated in a model-independent way [25] from repeated presentations of a subset of stimuli in order to sample the probability distribution $P(\text{spike}\mid\vec{s})$:

$$I_{spike} = \left\langle\frac{P(\text{spike}\mid\vec{s})}{P(\text{spike})}\log_2\frac{P(\text{spike}\mid\vec{s})}{P(\text{spike})}\right\rangle_{\vec{s}} \qquad (6)$$

Below we illustrate the results of maximizing information on two examples taken to mimic properties of cells in primary visual cortex.

A model neuron with two relevant directions describing a spatiotemporal receptive field. In preparation to analyze responses of neurons in primary visual cortex, we have analyzed a model where the probability of a spike depends on stimulus history. The model filter $\hat{e}_1$ is a concatenation of 3 successive spatial filters, each one of them is a spatial Gabor function, cf. Fig. 1, upper row. Even though this particular example has a space-time separable receptive field, we make no such assumptions during reconstruction. The relevant direction we are about to find has

approximately 1000 independent parameters. In a model, a given sequence of 3 frames leads to a spike if projection $s_1 = \hat{e}_1 \cdot \vec{s}$ reaches a certain threshold value s_t in the presence of noise, i.e. a spike occurs if $s_1 > s_t + n$, where n is a Gaussian random variable of variance σ^2 and zero mean. Thus, the model is specified by the filter (receptive field) $\hat{e}_1$, the threshold value of projection s_t, and noise variance σ^2. Here, the threshold value was set at $s_t = 2.0$ and a noise standard deviation $\sigma = 0.5$ (both parameters are measured in units of standard deviation of the projection value s_1 onto the model filter). These values set the probability of a spike per frame at ≈ 0.06, which for frames shown at 30Hz corresponds to a firing rate of ≈ 2 Hz. The total number of spikes across all repetitions of the movie was $\approx 10,000$ (the movie duration is approximately ≈ 9min, and it was repeated 10 times). The STA vector or reverse correlation function (2) is shown in the middle row of frames, Figure 1. It represents a broadened version of the underlying relevant direction, a consequence of spatiotemporal correlations present in the stimuli. If stimuli were drawn from a Gaussian probability distribution, the effects of correlations could be removed by multiplying the STA vector by the inverse the a priori covariance matrix, cf. Eq. (4). As mentioned above this procedure fails because natural scenes have higher than second order correlations. In contrast, it is possible to obtain a good estimate of the relevant direction by maximizing information directly. Frames of the recovered relevant direction $\vec{v}_{max}$ found by maximizing information are shown in the bottom row in Figure 1. The quality of reconstruction, as judged by the final value of the projection was on average: $\vec{v}_{max} \cdot \hat{e}_1 = 0.855 \pm 0.007$, with projections of the stimulus ensemble on it accounting for

Figure 1. Analysis of a model neuron with a spatiotemporal filter (upper row) and the corresponding nonlinear function describing the probability of a spike given projection on the model filter. The projection values are given in units of their standard deviation. The STA vector is shown in frames of the middle row, with spike occurring on the right-most frame. It is broadened in both space and time due to spatiotemporal correlations present in the stimulus movie. The corresponding recovered nonlinearity is shown to the right. The reconstructed vector as the most informative direction $\vec{v}_{max}$ is shown in the bottom row. The model filter is normalized to unit length, while the other two vectors are shown relative to their noise standard deviations.
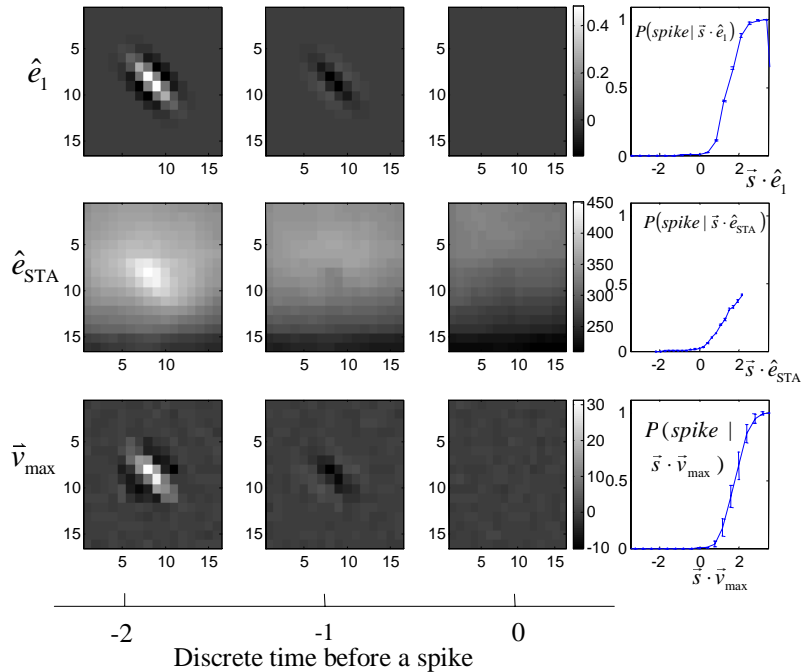
$I(\vec{v}_{max})/I_{spike} = 0.999 \pm 0.002$. The quality of reconstruction in terms of projection between the reconstructed and model filters improves as the number of spike increases, as is discussed in detail below. We have set the model filter to be equal to zero on the frame during which a spike occurs to describe time delay. This and other acausal frames (not shown) could be used to verify algorithm accuracy, as well as typical error values in the reconstructed vector. For example, the STA vector has remaining structure ay the time lag 0, while the reconstructed vector has very little structure at this time lag, as quantified with respect to the noise level.

A model neuron with two relevant directions describing a spatiotemporal receptive field. A sequence of spikes from a model cell with two relevant directions was simulated by projecting each of the stimuli on two vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ which are taken to mimic properties of complex cells found in primary visual cortex. These vectors are shown in two-spatial dimensions and at 3 discrete time lags before a spike in Figure 2(a,b). Their spatial dependence is described by the same Gaussian envelope and with $\pi/2$ difference in their spatial phase (sin and cos). A particular frame leads to a

spike according to a logical OR, that is if either $|s_1|$ or $|s_2|$ exceeds a threshold value s_t in the presence of noise, where $s_1=\mathbf{s}\cdot\mathbf{e}_1$ and $s_2=\mathbf{s}\cdot\mathbf{e}_2$. Similarly to the previous example of a model simple cell, a spike is elicited if either $|s_1|>s_t+n_1$ or $|s_2|>s_t+n_2$ is true. Here, n_1 and n_2 are independent Gaussian variables with zero mean and variance σ^2 modeling the effects of neural noise. The marginal probabilities $P(spike|s_1)$ and $P(spike|s_2)$ are also shown next to the corresponding vectors. The threshold for eliciting a spike was set at 2.0 standard deviations of the projection values s_1 and s_2, and the noise level was 0.5 in the same units. The resulting spike train with respect to the natural movie had an average firing rate of 4 Hz, with the overall number of spikes ~8,000 in the data set. We note that even though a simple logical operation has been chosen to specify the probability of a spike given the two projection values, we do not assume a particular shape of the input/output function during the reconstruction process. Due to the contrast-inversion symmetry, the STA vector calculated according to Eq. (2) would be zero in the case of a white noise stimulus ensemble. In the case of natural stimuli ensemble, the vector of STA is broadened in both space and time due the strong spatiotemporal correlations present in the ensemble, and not surprisingly, bears little resemblance to the underlying relevant dimensions, cf. Figure 2(c). Here we show that maximizing information as a function of direction in the stimulus space allows us to recover one of the relevant dimensions even under more natural stimulus conditions. The result of information maximization is shown in panels (d). The single most informative direction accounts on average for $I(\vec{v}_{max})/I_{spike}=0.55\pm0.02$ of the overall information carried by arrival of single spike, and has projection of 0.83 ± 0.03 onto the plane of the two relevant vectors $(\hat{e}_1,\hat{e}_2)$. Similarly to the case of a model simple cell with one relevant dimension, projection on the relevant plane increases with increasing number of spikes, see below. Even though the single most informative direction accounts for about half of the information in the carried by time of arrival of one spike, it gives values of preferred orientation and spatial frequency which are similar to those of the underlying two model vectors.
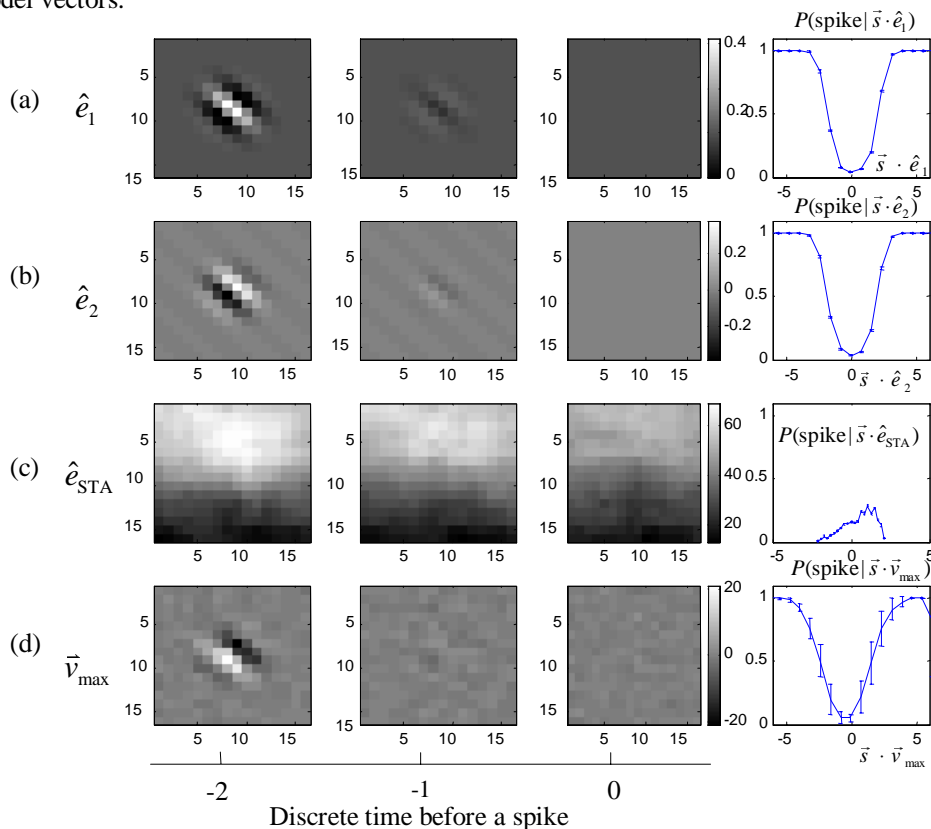

Figure 2. Analysis of a model neuron with two relevant dimensions describing a spatiotemporal receptive field. To describe the time dependence of the receptive field the time axis has been discretized; 3 separate time lags are considered, with a spike occurring on the right most frame at time lag 0. The first and second relevant dimensions of the model neuron are shown in the two upper rows, with marginal spike probabilities given projections of the two vectors. Row (c) describes the spatiotemporal STA with the corresponding spike probability given projection values of stimuli on it. The most informative vector is shown in row (d) and has projection 0.83 ± 0.03 for a data set with ~8,000 spikes.

Signal to noise ratios of reconstruction filters. Our estimate of information $I(\vec{v})$ along a particular direction $\vec{v}$ in the stimulus space is largely limited by the knowledge of the conditional probability distribution $P(\vec{s}\cdot\vec{v}\,|\,\text{spike})$. The number of recorded spikes N_{spike} is the number of available samples from this probability distribution. Therefore one might expect that with increasing number of spikes, the most informative vector obtained for a given spike sequence will approach underlying relevant dimensions used by the neuron. This is indeed the case, and in the case of a neuron with one relevant dimension $\hat{e}_1$ the projection of the (normalized) most informative vector $\vec{v}_{\text{max}}$ on $\hat{e}_1$ approaches one in the limit of large N_{spike} [21,26]: $\hat{e}_1 \cdot \vec{v}_{\text{max}} \approx 1 - \text{Tr}'[A^{-1}]/2N_{\text{spike}} \ln 2$, where A is the Hessian of information evaluated at $\vec{v}_{\text{max}}$ and trace Tr' is taken in the subspace orthogonal to the vector along which Hessian is evaluated. The main conclusion is that the expected errors when maximizing information are well behaved and decrease as $1/N_{\text{spike}}$ with increasing number of spikes. We emphasize that the expected deviation between the most informative vector and the underlying relevant dimension are of the same order of magnitude as errors of the reverse correlation method, where the neural filter is obtained according to Eq. (4) [21,26]. Since the reverse correlation and information maximization methods give similar errors when analyzing neural responses to Gaussian noise ensembles, the errors of the reverse correlation method will be larger when it is applied to a non-Gaussian stimulus ensemble. In Figure 3 we show how the projection value between the reconstructed and model vector increases (error decreases) when the number of repetitions, and therefore number of spikes, increases. Most importantly the reconstruction quality of one filter per cell is similar in the case of a simple and complex cell. Real neurons are likely to have more than one relevant dimension. However, we might be able to describe some of their neural characteristics by finding just one of the relevant filters for each neuron, as demonstrated in Fig. 2 above. Simulations of Figure 3 show that scaling with N_{spike}
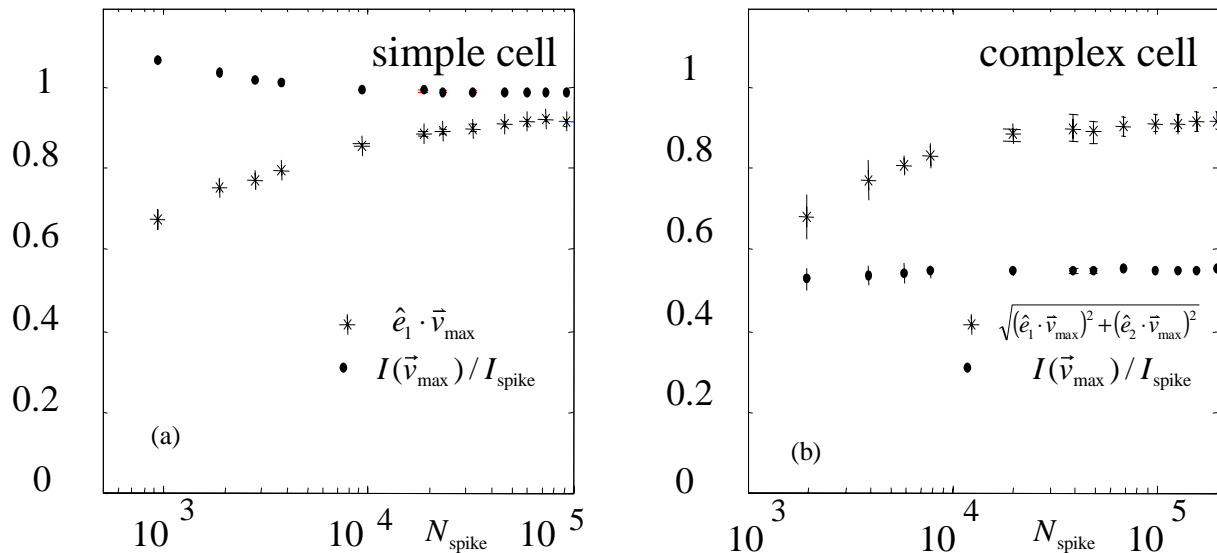

Figure 3. (a) Analysis of a model simple cell with one relevant direction. The reconstruction quality as quantified by $\hat{e}_1 \cdot \vec{v}_{\text{max}}$ increases to one with increasing number of spikes N_{spike}. Regardless of the number of spikes, the most informative dimension accounts for nearly all of the information carried by the arrival of one spike (the overestimate of the fraction of information explained $I(\vec{v}_{\text{max}})/I_{\text{spike}}$ observed for small N_{spike} is due to overfitting.

(b) Analysis of a model complex cell with two relevant dimensions $\hat{e}_1$, $\hat{e}_2$. Similarly to the case of a model neuron with one relevant dimension, the reconstruction quality as quantified by $\sqrt{(\hat{e}_1 \cdot \vec{v}_{\text{max}})^2 + (\hat{e}_2 \cdot \vec{v}_{\text{max}})^2}$ approaches one with increasing number of spikes N_{spike}. Regardless of the number of spikes, the most informative dimension accounts for approximately half I_{spike}, as should be expected given parameters of the model neuron that weighs contributions from the two relevant dimensions equally.

of projection values of the most informative direction onto the relevant plane $(\hat{e}_1, \hat{e}_2)$ of a complex cell is similar to that of projection values onto the single relevant dimension $\hat{e}_1$ of a simple cell. We also note that the dependence of the information values accounted for one most informative dimension on N_{spike} is not as strong as the increase in the projection value $\hat{e}_1 \cdot \vec{v}_{\text{max}}$.

3. RESULTS

We have probed neurons in cat primary visual cortex with three different stimuli ensembles: white noise stimuli, natural stimuli, and moving gratings. For each of the neurons in our dataset, we calculate two estimates of the filter used by the neuron for producing spikes: vector $\hat{e}_{\text{white noise}}$ is obtained as a spike-triggered average for a white noise stimulus ensemble using Eq. (2); vector $\hat{e}_{\text{natural stimuli}}$ is obtained as the most informative dimension for a natural stimulus ensemble. From the two estimates for the neural filter we calculate such major parameters as preferred orientation and spatial frequency. In addition, an independent estimate for preferred orientation and spatial frequency for each neuron is obtained from its responses to moving gratings. Only neurons with non-zero $\hat{e}_{\text{white noise}}$ vector have been included in the current analysis. Our first result is that all three estimates of the preferred orientation and spatial frequency are consistent with one another, cf. Figure 4. The agreement between preferred orientation value derived from natural stimuli filter $\hat{e}_{\text{natural stimuli}}$ and that obtained from grating responses is as good as or better, than agreement between preferred orientation value obtained the white noise filter $\hat{e}_{\text{white noise}}$ thorough a well established method of reverse correlation. Similarly, the values of the preferred spatial frequencies obtained from theses three independent measurements are consistent with one another, cf. Figure 4(b). Thus providing support for the validity of looking for most informative dimensions as a method to recover neural filters for correlated non-Gaussian stimuli ensembles, such those derived from the environment.

We next compare the two estimates of neural filters $\hat{e}_{\text{white noise}}$ and $\hat{e}_{\text{natural stimuli}}$ directly with each other. We also apply the reverse correlation method with corrections for power spectrum (4) to natural stimuli ensemble. The obtained vector $\hat{e}_{\text{Gaussian estimate}}$ would have been a valid estimate for the neural filter if natural stimuli obeyed Gaussian statistics or if a particular neuron under study is strictly linear [21]. We find that the most informative vector $\hat{e}_{\text{natural stimuli}}$ is always closer to $\hat{e}_{\text{white noise}}$ (in a direct Euclidean sense in the space of light intensities) than the corrected for power spectrum reverse correlation vector $\hat{e}_{\text{Gaussian estimate}}$, cf. Figure 5. One might ask whether improvement shown in Figure 5(a) from filter estimate $\hat{e}_{\text{Gaussian estimate}}$ to the information-theoretic one $\hat{e}_{\text{natural stimuli}}$ is significant. In Figure 5(b), we show that in about ~70% of cells this difference is significant. To determine whether the two methods give significantly different results, we calculate distributions of projection values of estimates derived by either one of the two methods from different data segments, and find the mean and standard deviation of projection values $|\hat{e}_{\text{Gaussian estimate}} \cdot \hat{e}_{\text{Gaussian estimate}}|$ and $|\hat{e}_{\text{natural stimuli}} \cdot \hat{e}_{\text{natural stimuli}}|$. The difference between vectors $\hat{e}_{\text{Gaussian estimate}}$ obtained through reverse correlation procedure (4) and $\hat{e}_{\text{natural stimuli}}$ obtained by maximizing information is considered significant if the average projection value $\langle |\hat{e}_{\text{natural stimuli}} \cdot \hat{e}_{\text{Gaussian estimate}}| \rangle$ lies below the average $\langle |\hat{e}_{\text{Gaussian estimate}} \cdot \hat{e}_{\text{Gaussian estimate}}| \rangle$ by more than one standard deviation. The percentage does not change when we distinguish between simple and complex cells, however only cell with non-zero white noise filter where included in the subset.

4. DISCUSSION

Neural feature selectivity can be studied with respect to various stimulus ensembles, be it collection of bars, sequence of gratings, or white noise stimuli. Ultimately, we are only interested in neural responses to artificial stimulus ensembles to the degree to which they provide an approximation of neural responses in their working regime, which for sensory neurons mean processing of signals derived from the environment. Even though one might argue in each
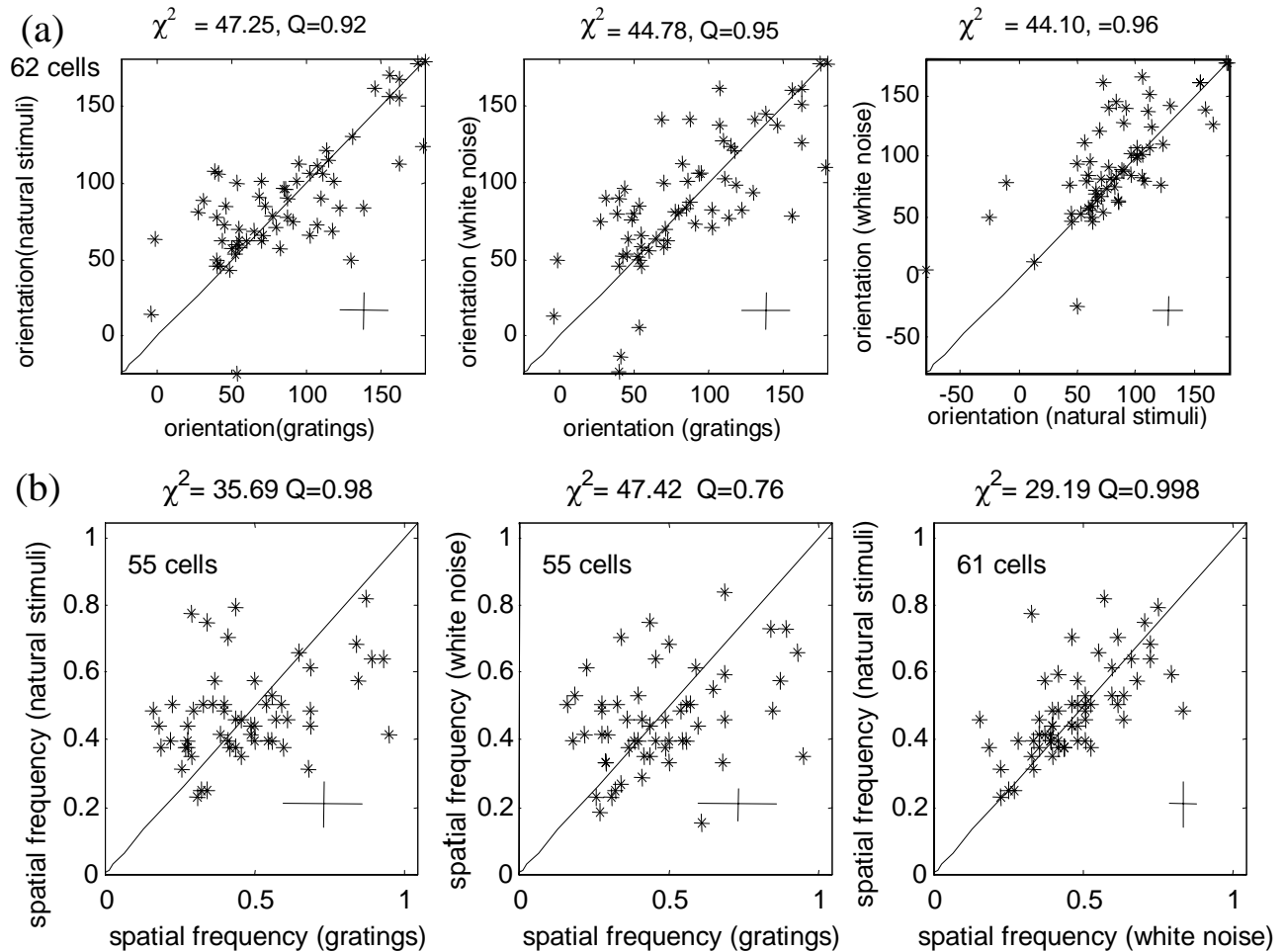
Figure 4. (a) Pairwise comparison of preferred orientation values derived from the natural stimuli filter, the white noise filter, and moving gratings across an ensemble of 62 cells that had non-zero white noise filter. Standard deviations averaged across the ensemble are shown as error bars in a corner of each subplot. Solid lines have slope one. Consistency with a hypothesis that preferred orientation values are the same is evaluated by χ^2-test, and Q gives the probability that the observed discrepancies are chance fluctuations [27]. (b) Pairwise comparison of preferred spatial frequency values derived from the natural stimuli filter, the white noise filter, and moving gratings. Again, only cells with non-zero white noise filter have been included.

particular case whether a chosen stimulus ensemble is "natural enough", our primary purpose in this work was to demonstrate that neural filter properties can be derived with respect to any, sufficiently diverse ensemble, of stimuli that experimenter may think is relevant. Most importantly, in doing so we do not assume that stimuli are Gaussian, since ensembles of both visual and auditory stimuli taken from our environment have been shown to deviate strongly from a Gaussian distribution. Our particular ensemble of natural stimuli has been obtained by walking with a video camera through a park. What we have shown is that despite the complicated spatiotemporal correlations present in our ensemble, the neural filters could be in essence "decorrelated" through the information-theoretic procedure of maximizing information. To some degree this "decorrelation" can be achieved by such standard methods as reverse correlation with power spectrum normalization Eq. (4). However, for majority (70%) of even simple cells, which are thought to be well described by one filter, these leads to significant filter distortions and could make us overestimate the difference between neural filters obtained from neural responses to white noise and more natural stimuli ensembles.

When comparing filters derived from neural responses to white noise and natural stimulus ensembles, we find that their major parameters, such as size, preferred orientation and spatial frequency are consistent with one another and
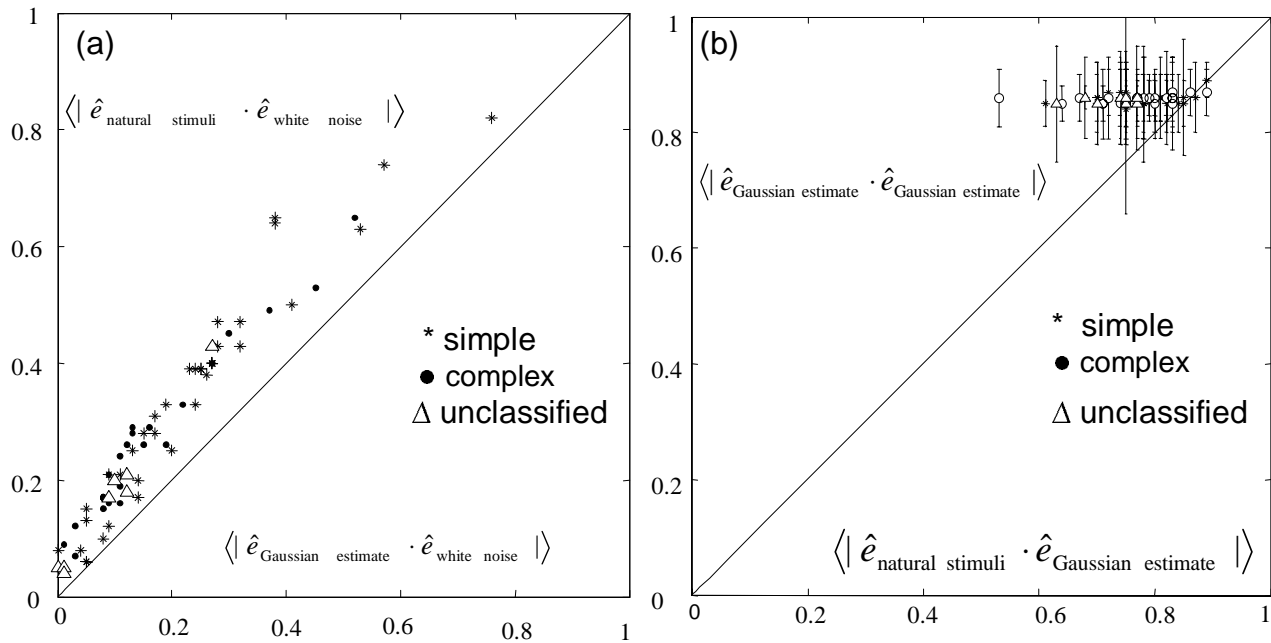
Figure 5. (a) The filter estimate $\hat{e}_{\text{natural stimuli}}$ obtained for natural stimuli ensemble by maximizing information is always closer to the white-noise filter $\hat{e}_{\text{white noise}}$ than the estimate $\hat{e}_{\text{Gaussian estimate}}$ obtained from natural stimuli ensemble by reverse correlation procedure with power spectrum corrections (4). (b) Differences between the two filter estimates $\hat{e}_{\text{Gaussian estimate}}$ and $\hat{e}_{\text{natural stimuli}}$ are significant for 69% of simple cells (36 cells total), 69% of complex cells (23 cells total) and for all 8 of unclassified cells, see text for significance criterion. Only cells with non-zero white noise filter $\hat{e}_{\text{white noise}}$ where included in the analysis.

with direct measurements of these parameters obtained with moving gratings. This can be viewed as additional verification of the algorithm since it is unlikely that major parameters of neural filters (receptive fields) would drastically change when stimulus ensemble is substituted with another. In terms of signal to noise ratios of the two filter estimates, we find that the agreement between parameters obtained from natural stimuli filter and moving gratings is as good as that between parameters obtained form white noise filter and moving gratings. This suggests that the information-theoretic method achieves comparable signal to ratios in filter reconstruction as those of reverse correlation method for noise ensemble. Analytically, the convergence properties for the information-theoretic method have been shown to be comparable those of reverse correlation method [21] for the same number of recorded spikes in a dataset.

Aside from the fact that major parameters of filters derived from white noise and natural stimuli ensemble are the same, one might notice that the direct dot product between the two filters, cf. Figure 5a, is often much smaller than could be explained by the observed noise level, as judged by how different, on average, two white noise filters could be when calculated from different stimulus presentations. This opens a series of questions as to what parameters of neural filters do change with the stimulus ensemble, some of which we hope to address in the subsequent work. The information-theoretic method provides us with means for a rigorous reconstruction of neural responses to natural stimuli.

In conclusion, we demonstrate for the first time on real neural data that it is possible to reconstruct neural filters from natural stimuli without assuming the linearity of neurons under study. In particular, the filter estimate obtained by maximizing information is always closer to the white noise filter than that obtained assuming that the cell is completely linear (reverse correlation function corrected for non-uniform power spectrum). The method does not require extraordinary amount of physiological data and performs as good as standard methods (reverse correlation, spike-triggered average) given similar number of spikes. Across the population of cells with non-zero white noise

filters, major parameters of neural filters derived from noise ensembles and natural stimuli were consistent with one another as well as with direct measurements of these parameters obtained with moving gratings.

ACKNOWLEDGEMENTS

REFERENCES

1. D.H. Hubel and T.N. Wiesel, "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex", *J. Physiol.*, **160,** 106-154 1962.
2. D. L. Ringach, G. Sapiro, and R. Shapley, "A subspace reverse-correlation technique for the study of visual neurons", *Vision Res.*, **37,** 2455-2464, 1997.
3. J. L. Gallant, J. Braun, D. C. van Essen, "Selectivity for polar, hyperbolic and cartesian gratings in macaque visual cortex", *Science*, **259**, 100-103, 1993.
4. J.L. Gallant, C.E. Connor, S. Rakshit, J.W. Lewis, and D.C. van Essen, "Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey", *J. Neurophysiol.*, **76,** 2718-2739, 1996.
5. E. de Boer and P. Kuyper, "Triggered correlation", *IEEE Trans. Biomed. Eng.,* **15,** 169-179 (1968).
6. R. R. de Ruyter van Steveninck and W. Bialek, "Real-time performance of a movement-sensitive neuron in the blowfly visual system: coding and information transfer in short spike sequences", *Proc. R. Soc. Lond. B,* **234**, 379-414, 1988.
7. F. Rieke, D. Warland, R. R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the neural code.* MIT press, Cambridge, 1997.
8. N. Brenner, W. Bialek, and R. R. de Ruyter van Steveninck, "Adaptive rescaling optimizes information transmission", *Neuron,* **26**, 695-702, 2000.
9. F. E. Theunissen, K. Sen, and A. J. Doupe, "Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds", *J. Neurosci.,* **20**, 2315-2331, 2000.
10. K. Sen, F. E. Theunissen, and A. J. Doupe, "Feature analysis of natural sounds in the songbird auditory forebrain", *J. Neurophysiol.,* **86**, 1445-1458, 2001.
11. F. E. Theunissen, S.V. David, N.C. Singh, A. Hsu, W. E. Vinje, and J. L. Gallant, "Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli", *Network: Comput. Neural Syst.,* **12**, 289-316, 2001.
12. E. J. Chichilnisky, "A simple white noise analysis of neuronal light responses", *Network: Comput. Neural Syst.* **12** 199-213 (2001).
13. O. Schwartz, E.J. Chilchilnisky, E.P. Simoncelli, "Characterizing neural gain control using spike-triggered covariance", in *"Advances in Neural Information Processing Systems 14",* edited by T. G. Dietterich, S. Becker, and Z. Ghahramani, MIT Press, Cambridge, MA 2002.
14. J. Touryan, B. Liu, and Y. Dan, "Isolation of relevant visual features from random stimuli for cortical complex cells", *J. Neurosci.,* **22**, 10811-10818, 2002.
15. D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells", *J. Opt. Soc. Am.* **4**, 2379-2394 (1987).
16. D. L. Ruderman and W. Bialek, "Statistics of natural images: scaling in the woods", *Phys. Rev. Lett.*, **73**, 814-817, 1994.
17. D. L. Ruderman, "The statistics of natural images", *Network: Comput. Neural Syst.,* **5**, 517-548, 1994.
18. D. J. Field, "What is the goal of sensory coding?", *Neural Comp.,* **6**, 559-601, 1994.
19. D. W. Dong and J. J. Attick, "Statistics of natural time-varying images", *Network: Comput. Neural Syst.,* **6**, 345-358, 1995.
20. E. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation", *Annu. Rev. Neurosci.,* **24**, 1193-1216, 2001.
21. T. Sharpee, N.C. Rust, W. Bialek, "Analyzing Neural Responses to Natural Signals: Maximally informative dimensions", *Neural Comp.,* **16**, 223-250 (2004).

22. R.C. Liu, S. Tzonev, S. Rebrik and K.D. Miller (2001), "Variability and information in a neural code of the cat latareal geniculate nucleus", *J. Neurophys.,* **86**, 2789-2806.

23. A.A. Emondi, S.P. Rebrik, A.V. Kurgansky and K.D. Miller, "Tracking neurons recorded from tetrodes across time", *J. Neurosci. Methods,* in press (2004).

24. T. L. Adelman, W. Bialek, and R. M. Olberg, "The information content of receptive fields", *Neuron*, **40**, 823-833, 2003.

25. N. Brenner, W. Bialek and R. R. de Ruyter van Steveninck, "Synergy in a neural code", *Neural Computation,* **12**, 1531-1552 (2000).

26. L. Paninski, "Convergence properties of three spike-triggered analysis techniques", *Network: Comput. Neural Syst.* **14**, 437-464 (2003).

27. W. H. Press, S. A. Teukolsky, W. T. Vettenberg, B. P. Flannery, *Numerical Recipes in C: the art of scientific computing,* (Cambridge University Press, 1992).