ENCODING AND PROCESSING OF SENSORY INFORMATION IN NEURONAL SPIKE TRAINS

F. GABBIANI[1] AND W. METZNER[2]

[1]*Division of Biology, 139-74 Caltech, Pasadena, CA 91125, USA* and [2]*Department of Biology, University of California at Riverside, Riverside, CA 92521-0427, USA*

*e-mail: gabbiani@klab.caltech.edu

Summary

Recently, a statistical signal-processing technique has allowed the information carried by single spike trains of sensory neurons on time-varying stimuli to be characterized quantitatively in a variety of preparations. In weakly electric fish, its application to first-order sensory neurons encoding electric field amplitude (P-receptor afferents) showed that they convey accurate information on temporal modulations in a behaviorally relevant frequency range (<80 Hz). At the next stage of the electrosensory pathway (the electrosensory lateral line lobe, ELL), the information sampled by first-order neurons is used to extract upstrokes and downstrokes in the amplitude modulation waveform. By using signal-detection techniques, we determined that these temporal features are explicitly represented by short spike bursts of second-order neurons (ELL pyramidal cells). Our results suggest that the biophysical mechanism underlying this computation is of dendritic origin. We also investigated the accuracy with which upstrokes and downstrokes are encoded across two of the three somatotopic body maps of the ELL (centromedial and lateral). Pyramidal cells of the centromedial map, in particular I-cells, encode up- and downstrokes more reliably than those of the lateral map. This result correlates well with the significance of these temporal features for a particular behavior (the jamming avoidance response) as assessed by lesion experiments of the centromedial map.

Key words: electric fish, *Eigenmannia*, stimulus estimation, signal detection, neural coding, spike train.

Introduction

Establishing a link between the activity of single nerve cells and their performance in conveying information about sensory stimuli has long been a goal of sensory neurophysiology. When attempting to relate the activity of a single nerve cell to performance in the encoding of sensory information, neurophysiologists face several conceptual difficulties. First, it is not always clear which aspect of a stimulus is encoded by a given nerve cell. This problem is particularly acute for neurons belonging to sensory systems analyzing complex natural modalities such as the visual system. Here, a rich variety of stimulus variables such as texture, contrast, motion, edges or color might be encoded by a neuron and used by the animal for visual scene analysis (Victor and Purpura, 1996). Second, the relevant neuronal output variable at subsequent stages of a sensory pathway is not always easy to determine and often a matter of debate (Koch, 1997; Shadlen and Newsome, 1998). In the past 50 years, efforts to characterize the performance of single neurons in encoding sensory information have mainly focused on characterizing discrimination thresholds between two stimuli on the basis of their mean firing rate measured in a fixed time window (for a review, see Parker and Newsome, 1998). The mathematical methods used in such studies are drawn from a particular topic in statistical signal-processing called signal detection theory (Green and Sweets, 1966; Poor, 1994).

In the present work, we review results on the encoding and processing of sensory information by neurons belonging to the first two stages of the amplitude-analyzing pathway in the weakly electric fish *Eigenmannia*. The methods and results presented here differ significantly from the earlier studies by focusing on the performance of single neurons in the encoding of time-varying stimuli. This type of analysis draws on another topic of statistical signal-processing called time-varying parameter or stimulus estimation (Poor, 1994). In a second step, we also apply the more traditional signal detection technique to study the information carried by single spikes and short spike bursts of neurons, rather than their mean firing rate.

One decisive advantage of the electrosensory system in gymnotiform weakly electric fish for this computational analysis is the extreme simplicity of electric stimuli. Only two parameters of the electric carrier signal generated by the electric organ of wave-type fish are thought to be used for electrolocation and electrocommunication: phase shifts and amplitude modulations. A second decisive advantage is that the anatomy, electrophysiology and behavior of wave-type electric fish has been characterized in considerable detail.

The electric organ in *Eigenmannia* discharges at fixed, regular intervals (between 150 and 600 Hz depending on the individual) following a command signal triggered by the medullary pacemaker nucleus. The resulting electric field is quasi-sinusoidal in its time course and spatially dipole-like. Nearby objects differing in impedance from water or conspecifics cause distortions in the electric field, which are detected by tuberous electroreceptor organs (Zakon, 1986).

Tuberous electroreceptor organs are distributed in the skin of *Eigenmannia* with highest density in the head region. They are contacted by two types of sensory afferent fibers which relay electroreceptive information to the brain: T-type and P-type tuberous afferents. T-type afferents fire an action potential on each cycle of the electric organ discharge (EOD), tightly phase-locked to the zero-crossing of the EOD, and are therefore thought to convey phase information to the central nervous system. P-type afferents, in contrast, fire action potentials only loosely phase-locked to the EOD and with a probability smaller than one per EOD cycle (Scheich et al., 1973; Bastian and Heiligenberg, 1980). Both fiber types are preferentially tuned to frequencies close to the individual's own electric organ discharge rate (Hopkins, 1976).

The firing probability per EOD cycle of P-type afferents increases linearly with the logarithm of the peak-to-peak electric field amplitude, suggesting that they are able to encode time-varying amplitude modulations by changes in their instantaneous firing rate (Hopkins, 1976). This point was recently confirmed in a modeling study of P-receptor afferent responses to low-frequency amplitude modulations (<5 Hz; see Fig. 12 of Nelson et al., 1997). To assess how accurately single spike trains of P-receptor afferents are able to convey information about time-varying amplitude modulations, we applied a stimulus estimation method which in essence represents a reverse of the more usual feedforward modeling of neuronal responses (Fig. 1) and which is described next.

Statistical signal detection and stimulus estimation theory

In this section, we will introduce first the signal-detection method used to characterize the performance of the mean firing rate in discriminating two stimuli and then the stimulus estimation method used to characterize the performance of single spike trains of P-receptor afferents in conveying detailed stimulus time-course information. This will allow us to compare and contrast these two statistical methods before applying them, in the following sections, to the analysis of actual neuronal data.

Let us consider a concrete example from a study in the extrastriate middle temporal (MT) area of the monkey cortex (Newsome et al., 1989). In MT area, most neurons are selective for direction of motion and will respond with a different mean firing rate to stimuli moving in their preferred *versus* anti-preferred direction. To characterize the accuracy with which the mean firing rate of a single neuron conveys information about the direction of stimulus motion, one starts by recording the mean response of the cell to many presentations of the two
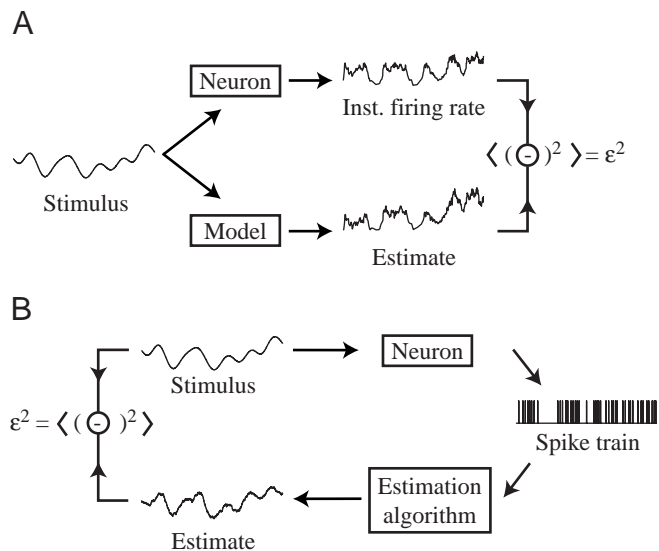

Fig. 1. Comparison between feedforward modeling of neuronal responses and stimulus estimation methods. (A) A model (which might be obtained by computation of Wiener kernels, for instance) attempts to predict the instantaneous (Inst.) firing rate of a nerve cell in response to time-varying stimuli. Model accuracy is quantified by the mean square error, ε^2, between the instantaneous firing rate of the neuron and its estimate. ε^2 is obtained by first subtracting the estimate from the experimental instantaneous firing rate at each time point and then squaring the result. The average squared error is then computed by adding single error values obtained at each time point and dividing them by the number of time points (in this and the following figures, this time-averaging operation is symbolized by < >). (B) For stimulus estimation, an algorithm is designed to produce an estimate of the stimulus from the spike train recorded from the neuron (see Fig. 4). As in the case of forward modeling, the estimation algorithm is designed to minimize the mean square error between the stimulus and the estimate.

moving stimuli. This allows an estimate of the probability density of response to stimuli s_1 and s_2 (preferred *versus* anti-preferred direction of motion in our example; Fig. 2A) to be built up. Typically, these two distributions will overlap, indicating that the mean firing rate is not a perfect indicator of stimulus motion direction. If s_1 and s_2 are now presented randomly with probability 0.5 and the mean spike count is recorded in response to such a stimulus, one can use the following method to 'guess' the direction of stimulus motion on the basis of the neuron's response: fix a threshold spike count k_{thresh}; compare the observed mean spike count with the threshold; and announce s_1 (preferred direction) if the spike count is above threshold and s_2 otherwise (Fig. 2B, upper panel). Two types of errors may occur during this 'guessing' procedure: (1) the stimulus might be announced as moving in the neuron's preferred direction when it was not (a false alarm, with probability P_{FA}) and, conversely, (2) the stimulus might be announced as moving in the anti-preferred direction when it was not (a false miss, with probability P_{FM}). Of course, the frequency of both error types P_{FA} and P_{FM} will depend on the choice of the threshold k_{thresh}; this is usually summarized by
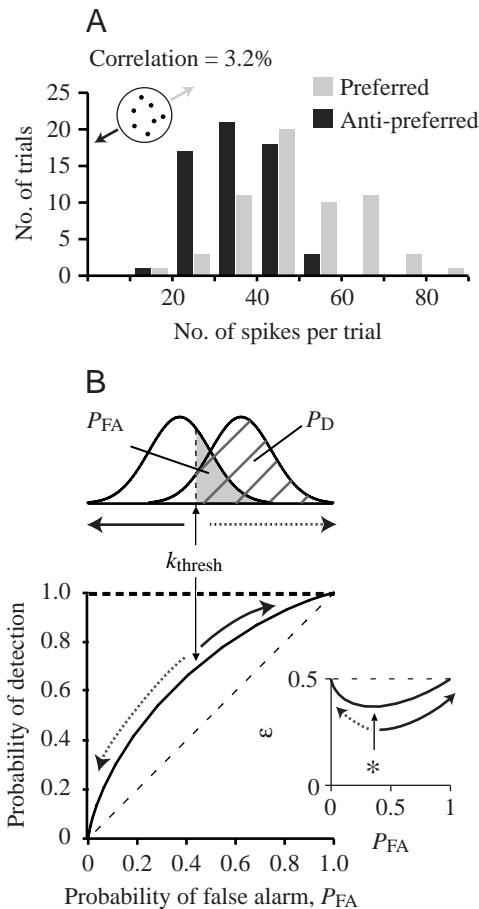
A

Table 1. *Analogies between stimulus estimation and signal detection methods*

	Signal detection	Stimulus estimation
Basic configuration	Two stimuli presented at random	Time-varying random stimuli drawn from a statistical ensemble
Measure of accuracy	Minimal error rate in predicting stimulus 1 and 2	Minimal root mean square error in estimating the stimulus
Minimization parameter	Threshold value k_{thresh}	Estimation filter h(t)

Both stimulus and signal detection require a random input signal (a time-varying stimulus for estimation and two different discrete stimuli for detection). In both cases, a parameter is optimized to minimize an error criterion and the value of the error for the optimal parameter is taken as a measure of performance.

parametric plot; Fig. 2B, lower panel). Such a plot is called a receiver operating characteristic (ROC) curve. For each value of k_{thresh}, the total error rate in guessing the stimulus direction of motion is given by averaging the two possible types of error, $\varepsilon = 0.5 P_{FA} + 0.5 (1 - P_D)$. If we choose the optimal threshold k_{thresh} that minimizes ε (Fig. 2B, right inset), then the corresponding minimal error rate provides an objective measure of the discrimination ability of the neuron's mean firing rate: under fairly general assumptions, it can be shown that no other algorithm will provide a lower error rate. Note that the determination of the minimal error rate relies on knowledge of the statistics of the stimulus (each stimulus is presented with probability 0.5) and of the neuron's response, since the optimal threshold is obtained by taking the cell's firing probability in response to both stimuli into account.

The procedure used to assess the performance of a neuron in conveying detailed time-course information about a time-varying stimulus is analogous to the procedure used for signal detection (Table 1). First, one needs a randomly time-varying stimulus s(t) (this replaces the random stimulus presentations in the signal detection case) and the neuron's response to such stimuli. To 'guess' the value of the stimulus at each time point from the spike train, a waveform h(t) is centered around each spike and these consecutive waveforms are summed (Fig. 3A). This yields a continuous signal which can be compared with the original stimulus. Of course, the accuracy with which the estimated stimulus will reproduce the original one depends on the choice of the waveform h(t). Just as for the threshold k_{thresh} in the signal detection case, the filter h(t) is optimized to minimize the mean squared error ε^2 between stimulus and estimated stimulus. The procedure yielding the optimal filter (due to Kolmogorov and Wiener; see Wiener, 1949) takes both into account, the *a priori* statistics of the stimulus and the statistics of the neuron's response (Fig. 3B). The root mean square error ε between the stimulus and the estimated stimulus for the optimal filter h(t) can then be used as an objective

Fig. 2. Signal detection theory allows the information about two stimuli conveyed by the mean firing rate to be quantified. (A) Distribution of the number of spikes elicited in a middle temporal area (MT) neuron of the monkey extrastriate visual cortex in reponse to dot motion on a screen. A small fraction of the dots (3.2 % in the present case) moved in either the preferred or anti-preferred direction of the cell (left inset), while the majority of dots moved randomly. The two distributions of spikes elicited during a 2 s trial do not overlap entirely, indicating that the mean spike count in this time window carries information about stimulus direction (adapted from Newsome et al., 1989). (B) Discrimination between two stimuli s$_1$ and s$_2$ from neuronal responses is performed by choosing a threshold k_{thresh} and is quantified by computing the probability of false alarm (P_{FA}, grey area in the upper panel of B) and the probability of correct detection (P_D, hatched area). The plot of P_D *versus* P_{FA} (lower panel) for each value of k_{thresh} is called a receiver operating characteristic (ROC) curve. As the threshold value increases (dotted arrow, upper panel), P_D and P_{FA} decrease (dotted arrow, lower panel) and *vice versa* for a decrease in k_{thresh} (filled arrow in both panels). Thin dashed line, chance performance ($P_D = P_{FA}$); thick dashed line, perfect discrimination. Right inset: discrimination error $\varepsilon = 0.5 P_{FA} + 0.5 (1 - P_D)$ plotted as a function of P_{FA}. The minimum value (corresponding to the optimal threshold) is indicated by an asterisk. Dotted and filled arrows have the same meaning as in the two main panels; dashed line, chance performance.

plotting the false-alarm rate P_{FA} *versus* the correct detection rate ($P_D = 1 - P_{FM}$) for all possible values of k_{thresh} (non-
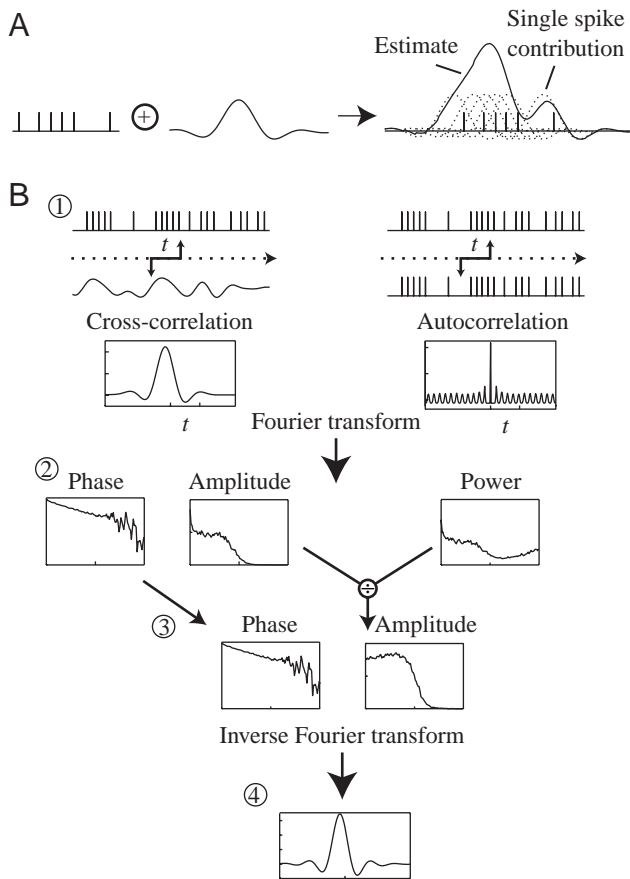
A

B

Fig. 3. Stimulus estimation from neuronal spike trains and computation of the optimal Wiener–Kolmogorov filter. (A) A stimulus estimate is essentially obtained by placing the optimal filter centered around each spike and summing the contributions arising from different spikes (solid line; for important technical details, see Wessel et al., 1996; Gabbiani and Koch, 1998). (B) Computation of the optimal filter corresponds to the following procedure. (1) The cross-correlation between stimulus and spike train as well as the autocorrelation of the spike train are first computed (note the strong periodicity of P-receptor autocorrelation due to phase-locking of the response). t is time. (2) Both functions are Fourier-transformed (yielding frequency-dependent phases and amplitudes in the case of the cross-correlation, and frequency-dependent amplitude only in the case of the autocorrelation; the phase is equal to zero and is independent of frequency). The Fourier transform of the autocorrelation is called the power spectral density. (3) The phase of the Wiener–Kolmogorov filter is identical to the phase of the Fourier-transformed cross-correlation, and its amplitude is obtained by dividing the corresponding frequency-dependent amplitude by the spike train power spectral density. (4) Inverse Fourier transformation yields the optimal filter in the time domain. In practice, steps 1 and 2 are combined for numerical calculations.

measure for the accuracy with which the stimulus is encoded in the spike train. It is easy to see that ε is at most equal to the standard deviation σ of the random stimulus and therefore the root mean square error can be normalized by dividing by σ: $\varepsilon_n = \varepsilon/\sigma$. A normalized error equal to zero means that the estimated stimulus matches perfectly the presented one, while $\varepsilon_n = 1$ signifies that the spike train does not provide any

information about the stimulus time course. Alternatively, $1 - \varepsilon_n$ (called the coding fraction) provides a measure ranging from 0 to 1 of the proportion of the stimulus standard deviation encoded by the spike train.

In addition to the error rate, ε_n, one usually computes the noise in the estimation procedure and compares its frequency-dependence with that of the original stimulus to obtain a frequency-dependent signal-to-noise ratio, SNR(f) (Fig. 4). A value SNR(f)=1 at a given frequency f means that the noise level matches the original signal level in the stimulus estimate (this corresponds to chance performance in the estimation task at that particular frequency), while perfect encoding at a given frequency f requires an infinitely large signal-to-noise ratio.

Fig. 4. Quantification of the accuracy with which spike trains carry information about time-varying stimuli. (1) The stimulus (with zero mean value) and the estimate are subtracted from one another to yield the residual noise in the estimation process. (2) The noise at each time point is squared and time-averaged (step 3) to yield the mean square error, ε^2, in the estimate. In parallel, the same procedure is applied to the stimulus to obtain its variance, σ^2, which is the mean squared error when estimation is at chance level (note that, if the spike train carries no

information about the stimulus, the best possible estimate is the mean value, equal to zero, and the residual noise is therefore equal to the stimulus itself). The square root of ε^2/σ^2 is the error normalized by its chance value: ε_n. In the frequency domain, the stimulus and noise power spectral densities are computed first (steps 4 and 5; Fourier transforms of the stimulus and noise autocorrelations, respectively). The ratio of stimulus power to noise power is then computed at each frequency (step 6), yielding the frequency-dependent signal-to-noise ratio, SNR. In the example depicted here, peak SNR=5, meaning that the noise power level is five times lower than stimulus power level in the estimate (dotted line, chance level, SNR=1).
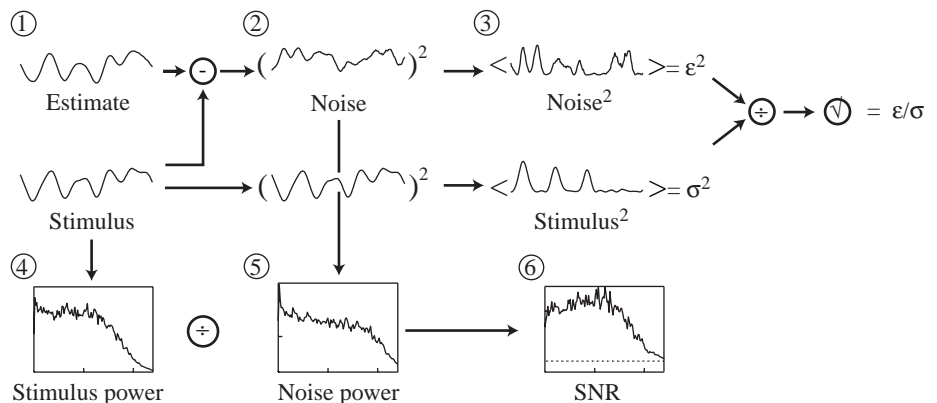
Encoding of random time-varying amplitude modulations in single spike trains of P-receptor afferents

The stimulus estimation method described in the previous section was applied to spike trains of P-type afferents recorded extracellularly in curarized fish (Wessel et al., 1996; Metzner et al., 1998). Curarization strongly attenuates the fish's own EOD, which was replaced by an EOD mimic generated using two electrodes placed in the mouth and near the tail of the fish. The EOD mimic consisted of a sinusoidal carrier signal having the same frequency as the fish's own EOD prior to curarization. The instantaneous amplitude of this carrier signal was pseudo-randomly modulated (random amplitude modulations, RAMs) around its mean value by multiplying it by a random waveform with a flat frequency spectrum up to a variable cut-off frequency (adjusted between 5 Hz and 700 Hz depending on the experiments) and a variable contrast (between 7 and 25 % of the mean amplitude; Fig. 5).

Single P-receptor afferent spike trains were able faithfully to encode RAMs of the carrier signal: an example of a particularly good encoding is illustrated in Fig. 6. In this case, a low cut-off frequency RAM was used to stimulate the recorded afferent fiber. As much as 83 % of the stimulus was encoded by the spike train, with signal-to-noise ratios reaching more than 70:1 in the frequency domain. However, in most afferent fibers, typical peak signal-to-noise ratios and the fraction of the stimulus encoded were usually lower, in the range 5–10 (for SNRs) and 0.4–0.6 (for the fraction of the stimulus encoded).

The fraction of the stimulus encoded depended on the firing rate of the afferent fiber, the contrast of the stimulus and the cut-off frequency of the RAM: the fraction encoded increased as the mean firing rate of the unit increased, as the contrast of the stimulus increased and as its cut-off frequency decreased (Fig. 7A,B,C, respectively; Wessel et al., 1996). Simulations of integrate-and-fire neurons with various noise levels inserted at the spiking threshold mechanism yielded qualitatively similar results (Gabbiani and Koch, 1998). In these models, the input current modulations can be shown to be exactly encoded in the instantaneous firing rate of the neuron, suggesting that an analogous encoding mechanism might be operating in P-receptor afferents for electric field amplitude modulations (see also Fig. 15 and Discussion).

The good performance of P-receptor afferents in encoding RAMs is due in part to their high sustained firing rates. Typically, the fraction of a time-varying stimulus which can be encoded by modulation of interspike interval duration in single spike trains increases with the number of spikes produced per stimulus correlation time (of the order of 10 spikes per correlation time usually yields good estimation results). Therefore, at low cut-off frequencies (<10 Hz), a single P-receptor afferent spike train can convey a fairly accurate representation of the stimulus, while at higher but still behaviorally relevant cut-off frequencies (<100 Hz) one expects that averaging over a small number of afferents (of the order of 10) would be sufficient to yield an accurate representation as well. Multi-unit recordings taking into account possible correlations between P-receptor afferent spike trains will be needed to confirm this point.
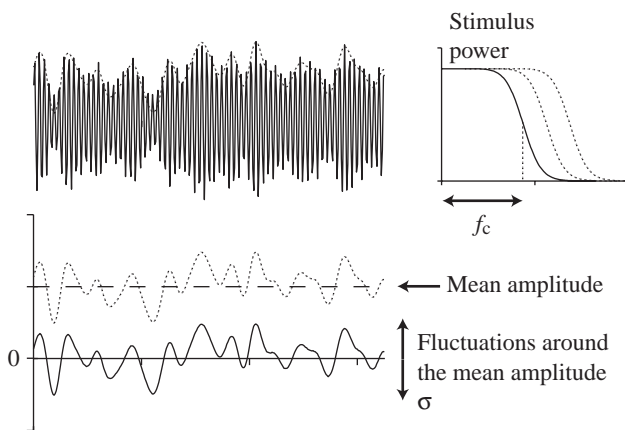

Fig. 5. Experimental stimuli. Top panel: signal played to the fish (sinusoidal carrier with random amplitude modulations, RAMs). The dotted line in the top and bottom panels represents the corresponding random amplitude modulation. Solid line in bottom panel: fluctuation around the mean amplitude value. This is the stimulus estimated from the recorded spike trains of P-receptor afferents. The parameters varied in these experiments were the mean amplitude of the signal (lower panel, horizontal broken line), the contrast σ of the random fluctuations (lower panel, double arrow) and the cut-off frequency of the RAM (right panel, double arrow labelled f_c).
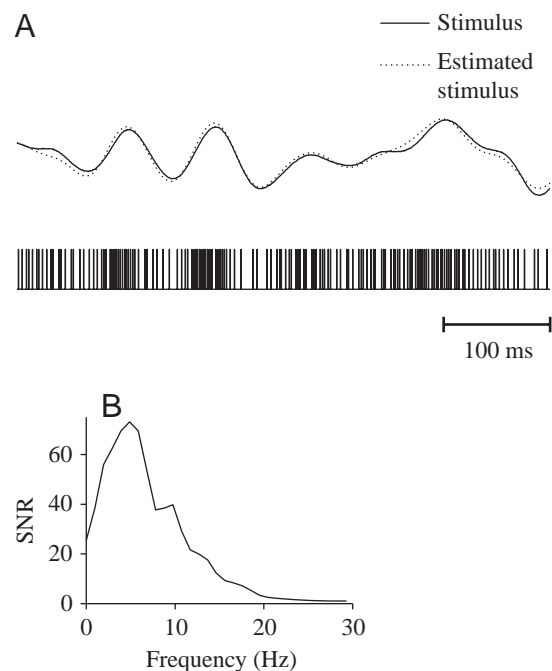

Fig. 6. Example of stimulus estimation from P-receptor spike train. (A) Spike train (bottom panel; mean firing rate 314 Hz), stimulus (solid line, top panel; cut-off frequency 10 Hz, contrast 0.24, mean amplitude 1 mV cm^{-1}) and estimated stimulus (dotted line, top panel; coding fraction 0.83). Note the clear modulation in instantaneous firing rate as the stimulus increases and decreases. (B) Signal-to-noise ratio (SNR) as a function frequency (peak SNR=73).
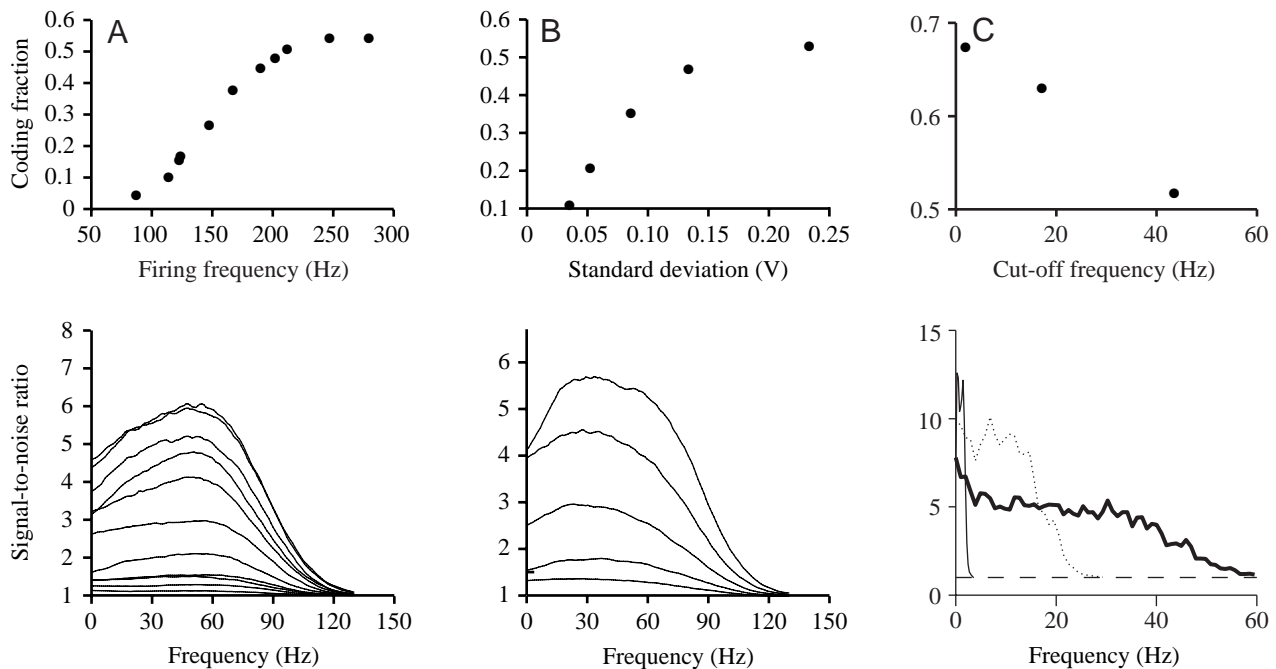
Fig. 7. Typical dependence of encoding accuracy on stimulus parameters (for definitions, see Fig. 5). (A) Fraction of the stimulus encoded as a function of mean firing rate (top panel) and corresponding signal-to-noise ratios (SNRs) as a function of frequency (bottom panel). In the bottom panel, each line represents the SNR at a fixed mean firing rate. The smallest SNR corresponds to the lowest firing rate in the top panel, and increasingly high SNRs correspond to increasingly high firing rates in the top panel. The increase in mean firing rate was obtained by an increase in mean stimulus amplitude. (B) Same as in A, but the stimulus parameter varied was the contrast (or standard deviation, σ) of the random amplitude modulation (RAM). (C) Same as in A, but the stimulus parameter varied was RAM cut-off frequency, f_c. Bottom panel: the thin solid line is the SNR for the lowest cut-off frequency in the top panel, the dotted line for the intermediate value of f_c and the thick solid line for the highest cut-off frequency f_c.

Since P-receptor afferents discharge at most one spike per EOD cycle, the EOD provides a natural clock for this sensory system. It is therefore possible to compare the performance of P-receptor afferents in encoding time-varying amplitude modulations with the best performance that may be achieved by transmitting information at the same rate as the EOD frequency (Fig. 8; Wessel et al., 1996). This reveals that P-receptor afferents are approximately half as accurate in encoding RAMs as an optimal system. While P-receptor afferents might not therefore be optimized to encode amplitude modulations as accurately as possible, their performance is not

too far from the optimum. It is important to note that this comparison assumes that the phase of P-receptor afferent firing within each EOD cycle does not carry information about the stimulus.

Processing of amplitude modulations in the ELL

The ELL is the first nucleus which processes electrosensory information relayed from the periphery by the lateral line nerve. Each tuberous afferent axon enters the ELL, trifurcates and sends collaterals to three independent somatotopic maps

Fig. 8. P-receptor afferent accuracy in encoding random amplitude modulations (RAMs) can be compared with the performance theoretically achievable by optimal encoding. Because P-receptor afferents fire at most one spike per electric organ discharge (EOD) cycle (see schematic spike train on top, the dashed lines represent different EOD cycles), the maximal information that could be transmitted by P-receptor afferents is one bit (corresponding to a spike or no spike) per EOD cycle. The equivalent bit rate (in bits s^{-1}) is therefore equal to the EOD discharge rate (i.e. for an EOD of 400 Hz this correspond to 400 bits s^{-1}). If one is limited to this information transmission rate, the best achievable root mean square error, $\varepsilon_{optimal}$, in estimating the stimulus can be computed exactly (assuming that the stimulus is Gaussian and white, see Kolmogorov, 1956). The ratio $\varepsilon/\varepsilon_{optimal}$ was typically equal to 2 in the experiments reported by Wessel et al. (1996).
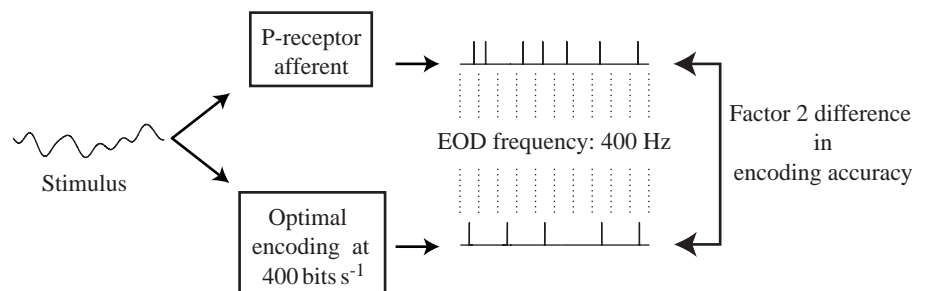
of the body surface called the centromedial (CMS), centrolateral (CLS) and lateral (LS) segments (Carr and Maler, 1986; see also Metzner et al., 1998, Fig. 1A). P-type tuberous afferents synapse directly onto basilar (E-type) pyramidal cells *via* an excitatory synapse and indirectly onto non-basilar (I-type) pyramidal cells *via* an inhibitory interneuron. As their names indicate, E-type ('excitatory') pyramidal cells increase their firing rate with an increase in stimulus amplitude, while I-type ('inhibitory') pyramidal cells are inhibited during an amplitude increase (Bastian and Heiligenberg, 1980; Saunders and Bastian, 1984). These two cell types are the only neurons of the amplitude-analyzing pathway that project outside the ELL to higher-order structures in the fish's brain. Pyramidal cells have been reported to possess different properties along the three body segments both *in vivo* (Shumway, 1989) and *in vitro* (Turner et al., 1996). Typically, the spatial resolution of their receptive fields decreases from the centromedial to the lateral segment, while their temporal resolution increases from the centromedial to the lateral segment. T-type afferents, in contrast, project onto a different set of cells which constitutes the timing pathway in the ELL (for a summary of ELL anatomy, see Berman and Maler, 1999).

To investigate how amplitude modulations are processed by the ELL circuitry, we used intracellular recordings of E- and I-type pyramidal cells from the two 'extreme' maps (CMS and LS) in response to the same RAMs used to drive P-receptor
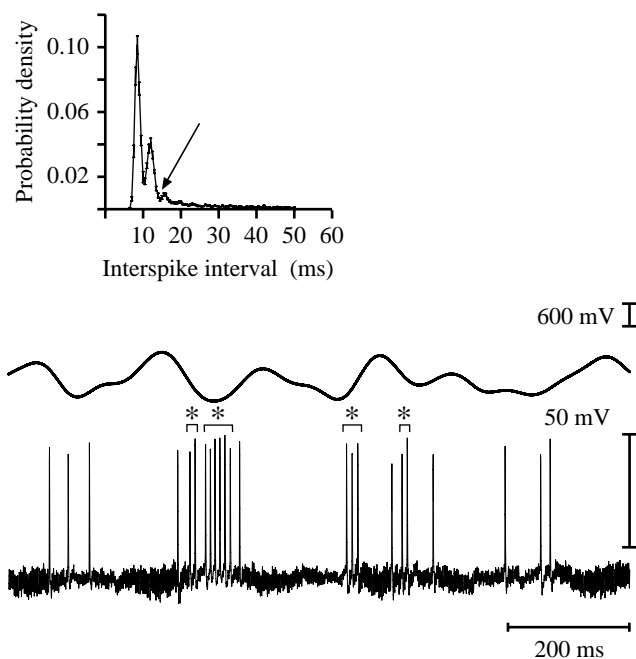

Fig. 9. Pyramidal cells fire in bursts. Bottom panel: intracellular recording from an I-type pyramidal cell in response to the random amplitude modulation stimulus shown on top. Upper inset: corresponding interspike interval (ISI) distribution. Note the two large peaks in the ISI distribution which correspond to spikes marked by an asterisk in the lower panel. Bursts were defined as two or more spikes separated by a time interval smaller than the ISI shown by the arrow in the upper inset.

afferent fibers (Gabbiani et al., 1996; Metzner et al., 1998). The response properties of pyramidal cells to RAMs of the electric field differed significantly from those of P-receptor afferents. Typically, the mean firing rate of pyramidal cells (20–80 Hz) was lower than that of P-receptor afferents. Their responses were also more stable against changes in stimulus parameters such as the mean amplitude and contrast. Estimation of the stimulus or of some simple non-linear function of the stimulus (such as half-wave rectified stimuli) always led to considerably higher normalized errors and lower signal-to-noise ratios than those obtained from P-receptor afferent spike trains, suggesting that, under our experimental conditions, pyramidal cell spike trains conveyed little information about the detailed time course of the stimulus or of some transform of it (Metzner et al., 1998). This was in part due to their lower firing rate, but also to their tendency to fire short bursts of spikes with relatively stereotyped interspike intervals (ISIs), thus precluding the encoding of time-varying information by changes in ISI (Fig. 9).

In spite of this, pyramidal cells were able to respond reliably to RAMs, as illustrated for instance in the example of one I-type pyramidal cell in Fig. 9. In response to the stimulus, this cell was able to signal the occurrence of downstrokes in the RAM waveform with either an isolated or a short burst of spikes. Therefore, under our stimulus conditions, pyramidal cells might attempt to convey reliable information about specific temporal features in the RAM waveform such as upstrokes or downstrokes. To test this idea quantitatively, the stimuli and recorded spike trains were binned in short (5–10 ms) intervals. The stimulus waveforms preceding each bin (ranging from 500 ms to 1 s prior to the bin) were separated into two distributions conditioned on the absence (P_0) or presence (P_1) of a spike in the bin (Fig. 10A). Assuming that pyramidal cell spikes were indeed attempting to convey information about the presence of a temporal feature in the random stimulus, this temporal feature should also optimally discriminate the two distributions of stimuli P_0 and P_1.

We therefore identified the waveform **f** optimally discriminating the two distributions P_0 and P_1 (following the method of Fisher, 1936; Fig. 10B) and, for each stimulus in P_0 and P_1, studied its degree of overlap with **f** (i.e. its projection onto **f**, see Fig. 10B). The optimal feature vector consisted of a large downstroke in the amplitude modulation preceded by a smaller upstroke for I-type pyramidal cells (Fig. 11A) and vice-versa for E-type cells. As illustrated in Fig. 11B, there was a substantial separation between P_0 and P_1 along the feature **f**. More interestingly, when only spikes belonging to bursts were taken into account to build a distribution of stimulus waveforms preceding a spike belonging to a burst, $P_{1\text{burst}}$, the separation increased considerably, thus indicating that bursts conveyed more reliable information about the optimal feature than isolated spikes. This finding was consistent across our pyramidal cell sample and was quantified using the same ROC analysis method described above: we computed the error rate in classifying a stimulus waveform as belonging to P_0 or P_1 (resp. $P_{1\text{burst}}$) based on its projection onto
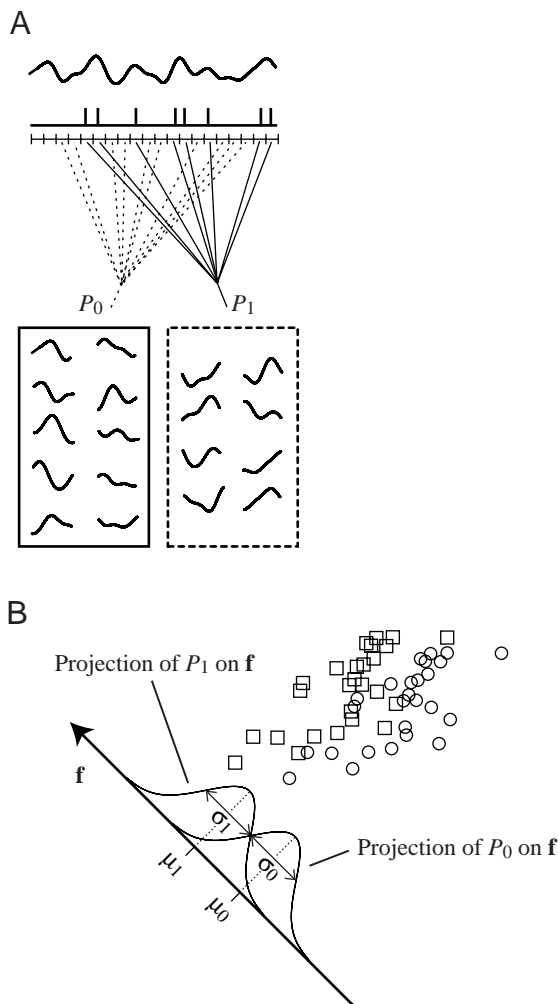
A

B

Fig. 10. Definition of P_0 and P_1 and computation of the optimal feature vector discriminating these distributions. (A) Stimuli and spike trains were first binned (5–10 ms). Depending on whether a bin contained a spike or not, the stimulus waveform preceding the bin (100 time points, corresponding to 500–1000 ms prior to the bin) was placed in two different 'packets' P_0 and P_1 (distributions of stimuli preceding a bin containing no spike and a spike, respectively). The separation of these two ensembles in stimulus space was studied by attempting to classify the stimuli as belonging to P_0 or P_1 using the classification procedure outlined in B. (B) Classification was performed by considering the projection of the stimuli onto a feature vector in stimulus space. The optimal feature vector (or direction in stimulus space) for classification was obtained by maximizing the distance between the mean projected distributions ($\mu_1-\mu_0$) normalized by their averaged standard deviations ($0.5\sigma_1+0.5\sigma_0$; Fisher linear discriminant). This is illustrated here in a hypothetical two-dimensional example (note that, since our stimuli consist of 100 time points or 100 dimensions, they cannot be easily visualized). The optimal direction $\mathbf{f}$ separating the two distributions of squares and circles (P_0 and P_1, respectively) is shown as a thick arrow, and the distributions of stimuli P_0 and P_1 projected onto $\mathbf{f}$ are also shown. Note that the direction perpendicular to $\mathbf{f}$ would not allow discrimination above chance level between P_0 and P_1, since the two distributions projected onto that direction overlap completely.

$\mathbf{f}$ (Fig. 11C,D). When the same analysis was performed on P-receptor afferent spikes, the two distributions P_0 and P_1 were found to be on average less well separated in stimulus space than those of pyramidal cells, corresponding to higher error rates in the classification task. The difference was even more pronounced when compared with the separation between P_0 and P_{1burst} for pyramidal cells (Fig. 12). Since for P-receptor afferents, the optimal feature waveform was similar to that for E-type cells, P-receptor afferent spikes were on average less accurate at conveying information about the occurrence of these features (upstrokes preceded by a small downstroke) than pyramidal cell spikes were for their own features (upstrokes preceded by a small downstroke for E-type cells; downstrokes preceded by a small upstroke for I-type cells). In contrast, the set of stimuli occurring prior to and during a 300 ms silent period in P-receptor afferent spike trains, $P_{silence}$, was well separated from the remaining stimuli, $P_{non-silence}$ (Metzner et al., 1998). Therefore, the encoding of downstrokes by I-type pyramidal cells might rely in part on an accurate encoding of similar waveforms by periods of silence in P-receptor afferent spike trains.

Since the method that we used to discriminate between P_0 and P_1 is based on a simple weighted linear sum of the input signal (the weights being determined by the time components of the optimal feature vector $\mathbf{f}$) followed by a threshold operation, this computation could be implemented in pyramidal cells by linear filtering of the amplitude modulation (or of an estimate computed from P-receptor afferent spike trains, see Discussion) followed by a thresholding of the resulting membrane potential at the spike-initiation zone. This requires detailed stimulus time-course information about the amplitude modulation still to be present in the measured subthreshold membrane voltage of pyramidal cells. We therefore estimated the stimulus from the subthreshold membrane voltage and compared the normalized error with that obtained from pyramidal cell spikes. No statistically significant improvement was observed (Fig. 13), suggesting that the simple model postulated above does not describe the implementation of the temporal feature extraction operation performed by pyramidal cells. However, it has recently been shown that bursts are generated in pyramidal cells by a backpropagation of somatic action potentials into the dendrites and activation of dendritic Na^+ channels (Turner et al., 1994; see Turner and Maler, 1999). In combination, these results suggest that the temporal feature extraction performed by pyramidal cells results from a dendritic computation.

The performance of pyramidal cells in encoding up- and downstrokes of RAMs was also dependent on the cell type and on the map to which the recorded cell belonged. I-type pyramidal cells encoded on average the presence of downstrokes better than E-type pyramidal cells encoded upstrokes. Furthermore, cells from the CMS were on average more accurate than cells from the LS in encoding temporal features of the amplitude modulation (Fig. 14). This last result correlates well with recent lesion studies showing that the CMS is necessary and sufficient for the generation of one particular
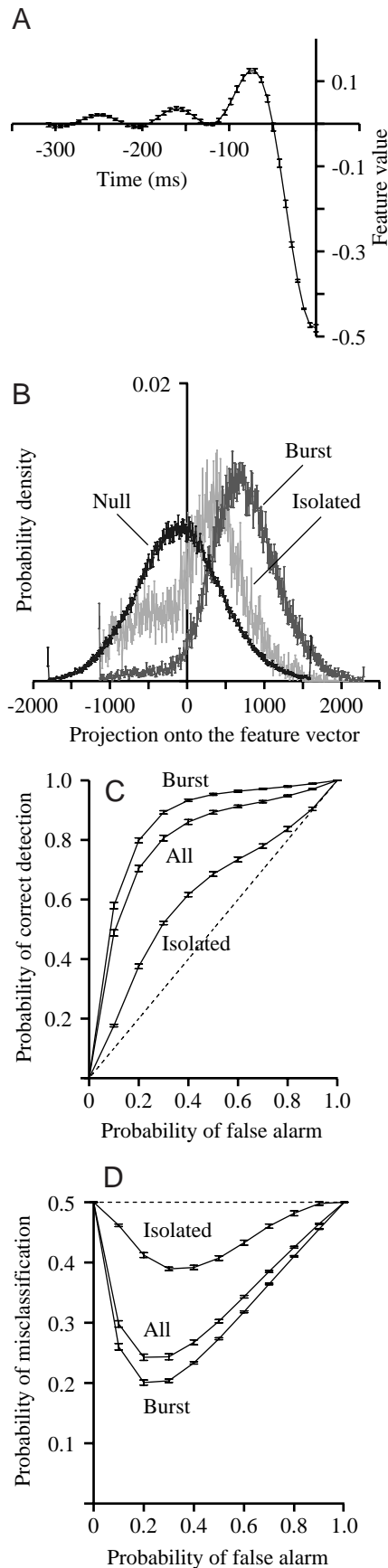
Fig. 11. Quantification of pyramidal cell spike performance in the feature extraction task. (A) Optimal feature computed following the method outlined in Fig. 10 for an I-type pyramidal cell. (B) Distribution of the stimuli projected onto the optimal feature vector $\mathbf{f}$ for a bin containing no spike (P_0, labeled null), a bin containing an isolated spike (isolated) and a bin containing a spike belonging to a burst (burst). When projected onto $\mathbf{f}$, the distribution of stimuli preceding a spike belonging to a burst is better separated from the null distribution than the distribution of stimuli preceding an isolated spike, indicating that spikes belonging to bursts carry more reliable information about the optimal feature than isolated spikes. (C) Receiver operating characteristic (ROC) curve characterizing the discrimination performance between the null distribution and the distribution of isolated spikes projected onto $\mathbf{f}$, all spikes taken together and spikes belonging to bursts, respectively. The ROC curve was computed following the method described in Fig. 2B (broken line: chance level). (D) Corresponding error rates for the three cases considered in C. The minimal error rate is lowest for spikes belonging to bursts (broken line, chance level).

behavior of gymnotiform fish, the jamming avoidance response (JAR; Metzner and Juranek, 1997). The computational rules underlying the generation of the JAR and their neuronal implementation have been extensively investigated in *Eigenmannia* (Heiligenberg, 1991). It is known that fish need to detect amplitude upstrokes and downstrokes combined with phase information to identify successfully whether their EOD frequency should be raised or lowered to avoid jamming by a conspecific. Hence, these lesion studies and the results reviewed here reveal that those pyramidal cells that are most sensitive to upstrokes and downstrokes in RAMs are also those participating in the generation of a specialized behavior known to put these temporal features to use.
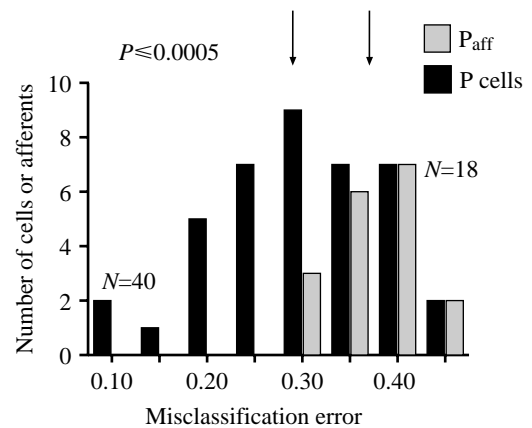

Fig. 12. Minimal classification error (computed as in Fig. 11) for pyramidal cell spikes belonging to bursts (P cells) and P-receptor afferent spikes (P_{aff}). Higher values of the misclassification error mean lower performance in the feature extraction task. P-receptor afferent spikes on average performed less well than pyramidal cell spikes in the feature extraction task (vertical arrows, median value of the misclassification error for both groups; significance level of the Wilcoxon rank-sum test 0.0005).
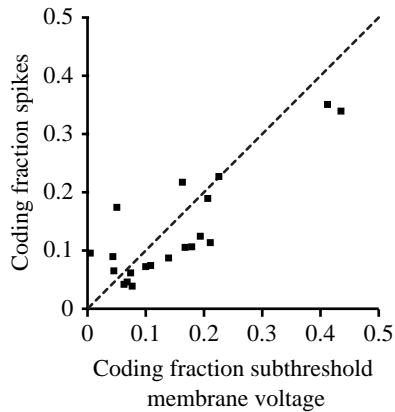
Fig. 13. Fraction of the detailed stimulus time course encoded in the subthreshold membrane voltage of pyramidal cells plotted against the fraction encoded in spikes. No significant improvement is observed, as would be predicted from the simple feature extraction model discussed in the main text. Dashed line, identical performance.

Discussion

The results of the preceding two sections suggest a considerable transformation in the encoding of time-varying amplitude modulations between the first two stages of the amplitude-analyzing pathway in *Eigenmannia*. The accurate time-course information conveyed by P-receptor afferents is processed in the ELL to extract the occurrence of specific temporal features in the amplitude modulation waveform, such as upstrokes and downstrokes of RAMs. These features are then explicitly represented in the output spike trains of E- and I-type pyramidal cells, preferentially by short spike bursts, and relayed to higher stations in the fish brain for further processing. Presumably, these features could be used to elicit different behaviors from the fish electrocommunication and electroreception repertoire. Since we have little evidence as to which features might be involved in most of these behaviors and as to how they are effectively used (except for the JAR), it is of interest to compare our current understanding of the processing of sensory information at the first two stages of the electrosensory pathway with the processing that might be performed by an engineer on similar input signals and with the processing of time-varying stimuli in other sensory systems.

The encoding of time-varying amplitude modulations by P-receptor afferents is most reminiscent of one-bit analog-to-digital (A/D) converters commonly used in the digitization and encoding of audio frequency band signals called oversampled sigma-delta modulators (Norsworthy et al., 1997). A simple diagram of such an A/D converter is illustrated in Fig. 15 (Inose and Yasuda, 1963): the input signal is integrated and compared with a threshold at each cycle of a synchronizing clock; a spike is emitted if the output of the integrator is above threshold and is subtracted from the input *via* a feedback circuit. The digitizer has a very low resolution since it can output only two different levels in the face of a continuous input signal. This very low resolution is
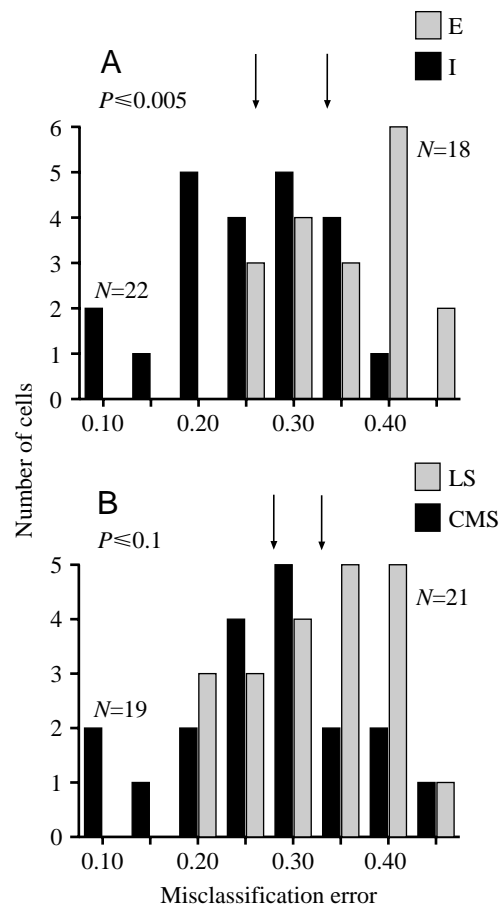

Fig. 14. Comparison of the performance of different pyramidal cell types in the feature extraction task. (A) I-type pyramidal cell spikes belonging to bursts on average performed the feature extraction task better than E-type cell spikes (vertical arrows, median value of the misclassification error for both groups; significance level of the Wilcoxon rank-sum test 0.005). (B) Centromedial segment (CMS) cell spikes belonging to bursts on average performed the feature extraction task better than lateral segment (LS) cell spikes (vertical arrows, median value of the misclassification error for both groups; significance level of the Wilcoxon rank-sum test 0.1). For a more detailed explanation of the origin of the difference in significance level and performance observed in A and B, see Table 1 and Discussion in Metzner et al. (1998).

compensated by a high sampling clock rate (thus explaining the use of the term 'oversampled'); usually the clock is run at a rate more than ten times higher than the cut-off frequency of the input signal. Important advantages of oversampled sigma-delta modulators are the simplicity of the underlying algorithm and its robustness against possible material defects (leading to leaky integration or threshold variations for example). The technique works best for low-bandwidth inputs and, in such cases, yields a very high final digitization resolution (Aziz et al., 1996). Furthermore, the input signal may easily be recovered from the digitizer output stream by simple low-pass filtering. From an engineering point of view, accurate sampling of the amplitude-modulated input signal is

Fig. 15. Schematic circuit representation of an oversampled sigma-delta modulator. The stimulus is passed through an integrator, and pulses are generated if the output of the integrator is above threshold at the end of each clock cycle (clock cycles are represented by the dotted lines below the output pulses). The output pulse is subtracted from the input *via* a feedback circuit so that effectively the difference between input and output is integrated. The input signal can be recovered by simple low-pass filtering.
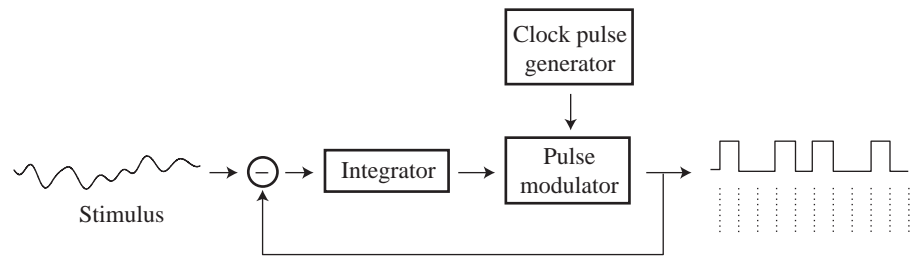
a natural strategy if subsequent stages of the system analyzing amplitude modulations are to extract various specific features from their input: it is known that under fairly general conditions optimal feature extraction algorithms must rely on such estimates of the input signal (Kailath, 1969).

As mentioned above, the presumed high firing rate of P-receptor afferents *in vivo* suggests that they might be specialized for accurately sampling the detailed time course of amplitude-modulated signals. The strategy adopted by the electrosensory system to convey information from the periphery to more central processing stations might be very different from that adopted in higher vertebrates to convey auditory signals in the auditory nerve or visual signals in the optic nerve, for example. In the case of audition, the presence of very high frequency components (much higher than the firing rates achievable by single neurons) precludes a detailed encoding at the level of single spike trains. For vision, the richness of the sensory input compared with the information-carrying capacity of the optic nerve has long been interpreted to imply a substantial preprocessing of photoreceptor input signals at the level of the retina, leading to a compression of visual information by redundancy reduction.

Whether the information conveyed by P-receptor afferents is used postsynaptically by pyramidal cells or by other ELL interneurons to estimate amplitude modulations effectively (by low-pass filtering for instance) remains an open question. Our intracellular recordings close to the spike-initiation zone of pyramidal cells suggest that such information is not present there, but the situation might be different further away along the basilar dendrites of pyramidal cells for instance. In the absence of *in vivo* intradendritic recordings, which appear extremely difficult, useful clues to answer this question would come from estimates of the number of afferent fibers converging onto a basilar pyramidal cell, the number of synaptic boutons per afferent axon and compartmental modeling of these neurons (Berman and Maler, 1998). The dynamics of synaptic responses, their reliability and the time course of excitatory postsynaptic potentials will also significantly affect the signal available postsynaptically.

It is interesting to note that, in the fly visual system, wide-field tangential neurons have been shown to encode significant information about time-varying random velocity fluctuations in their subthreshold membrane potential as well as in their spiking responses (Haag and Borst, 1998). The performance of these tangential neurons is poorer than that of P-receptor afferents, presumably because motion information has to be computed from nearby photoreceptor inputs. In the case of spiking neurons such as the H1 cell, their firing rate is also typically lower than that of P-receptor afferents. Of course, these studies do not directly address the question of whether the estimated velocity (or some simply related function) is used by the animal to generate optomotor responses (Gabbiani and Koch, 1996), just as our own study does not directly imply that the accurate amplitude modulation information conveyed by P-receptor afferents is explicitly used at the level of the ELL. In this respect, stimulus estimation techniques are no different from the signal detection techniques used to estimate the information carried by the mean spike count: both methods quantify the performance of an ideal observer of the recorded spike train in a task defined by the experimenter. Only a correlation with behavior can lead to insight as to whether this information is used by the animal. One advantage of stimulus estimation methods is their ability, now clearly demonstrated, to reveal detailed encoding of time-varying information from single spike trains and to quantify it. This represents progress not available through earlier methods.

The work reviewed here leaves open many additional interesting questions concerning the processing of amplitude modulations at the level of the ELL. One important open question concerns the encoding of RAMs by pyramidal cells when the electrical stimulus is spatially localized. Such electric field distortions are more natural than whole-field amplitude modulations in the context of electrolocation and are expected to stimulate differently the center-surround receptive fields of pyramidal cells. Another important question that has not been addressed in the present study is the differential sensitivity of pyramidal cells to various frequency bands of amplitude modulations from the CMS to the LS. Since our stimuli had a flat frequency spectrum, they were not very appropriate to quantify the performance of cells in the different segments to different frequency bands of amplitude modulations. Using different band-pass random stimuli would yield more appropriate information on this point. To formulate a concrete hypothesis: is it possible for pyramidal cells to convey detailed stimulus information about a specific frequency band of amplitude modulations when stimulated locally? Up to now, it has been shown using

extracellular recordings that pyramidal cells in the CMS behave more as low-pass filters, while LS cells behave more as high-pass filters (Shumway, 1989). A similar gradient in sensitivity among different cell types in the torus semicircularis has also been demonstrated recently (Fortune and Rose, 1997). Because these studies used deterministic sinusoidal amplitude modulations, they did not address the question of whether the spike train ISIs carry information about the detailed time course of the stimulus as could in principle be achieved using random-time varying stimuli and stimulus estimation methods.

We expect that further applications of statistical signal-processing techniques to the analysis of electrosensory processing, combined with anatomical, electrophysiological and behavioral approaches, will lead to further advances in the understanding of the neuronal basis of electrocommunication and electrolocation behaviors. Because of its relative simplicity, the electrosensory system in weakly electric fish represents an ideal model system for investigating the general issue of how information is encoded and processed in neuronal spike trains.

References

Aziz, P. M., Sorensen, H. V. and Van der spiegel, J. (1996). An overview of sigma-delta converters. *IEEE Sign. Proc. Mag.* **13**, 61–84.

Bastian, J. and Heiligenberg, W. (1980). Neuronal correlates of the jamming avoidance response of *Eigenmannia*. *J. Comp. Physiol.* **136**, 135–152.

Berman, N. J. and Maler, L. (1998). Distal versus proximal inhibitory shaping of feedback excitation in the electrosensory lateral line lobe: Implications for sensory filtering. *J. Neurophysiol.* (in press).

Berman, N. J. and Maler, L. (1999). Neural architecture of the electrosensory lateral line lobe: adaptations for coincidence detection, a sensory searchlight and frequency-dependent adaptive filtering. *J. Exp. Biol.* **202**, 1243–1253.

Carr, C. E. and Maler, L. (1986). Electroreception in gymnotiform fish, central anatomy and physiology. In *Electroreception* (ed. T. H. Bullock and W. Heiligenberg), pp. 319–373. New York: John Wiley.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Ann. Eugenics* **7**, 179–188.

Fortune, E. S. and Rose, G. J. (1997). Passive and active membrane properties contribute to the temporal filtering properties of midbrain neurons *in vivo*. *J. Neurosci.* **17**, 3815–3825.

Gabbiani, F. and Koch, C. (1996). Coding of time-varying signals in spike trains of integrate-and-fire neurons with random threshold. *Neural Comput.* **8**, 44–66.

Gabbiani, F. and Koch, C. (1998). Principles of spike train analysis. In *Methods in Neuronal Modeling*, 2nd edn (ed. C. Koch and I. Segev), pp. 313–360. Cambridge, MA: MIT Press.

Gabbiani, F., Metzner, W., Wessel, R. and Koch, C. (1996). From stimulus encoding to feature extraction in weakly electric fish. *Nature* **384**, 564–567.

Green, D. M. and Sweets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: John Wiley.

Haag, J. and Borst, A. (1998). Active membrane properties and signal encoding in graded potential neurons. *J. Neurosci.* **18**, 7972–7986.

Heiligenberg, W. (1991). *Neural Nets in Electric Fish*. Cambridge, MA: MIT Press.

Hopkins, C. D. (1976). Stimulus filtering and electroreception: Tuberous electroreceptors in three species of gymnotoid fish. *J. Comp. Physiol.* **111**, 171–207.

Inose, H. and Yasuda, Y. (1963). A unit bit coding method by negative feedback. *Proc. IEEE* **51**, 1524–1535.

Kailath, T. (1969). A general likelihood-ratio formula for random signals in Gaussian noise. *IEEE Trans. Inf. Theory* **15**, 350–361.

Koch, C. (1997). Computation and the single neuron. *Nature* **385**, 207–210.

Kolmogorov, A. N. (1956). On the Shannon theory of information transmission in the case of continuous signals. *IRE Trans. Inform. Theory* **2**, 102–108.

Metzner, W. and Juranek, J. (1997). A sensory brain map for each behavior? *Proc. Natl. Acad. Sci. USA* **94**, 14798–14803.

Metzner, W., Koch, C., Wessel, R. and Gabbiani, F. (1998). Feature extraction by burst-like spike patterns in multiple sensory maps. *J. Neurosci.* **18**, 2283–2300.

Nelson, M. E., Xu, Z. and Payne, J. R. (1997). Characterization and modeling of P-type electrosensory afferent responses to amplitude modulations in a wave-type electric fish. *J. Comp. Physiol.* A **181**, 532–544.

Newsome, W. T., Britten, K. H. and Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature* **341**, 52–54.

Norsworthy, S. R., Schreier, R. and Temes, G. C. (1997). *Delta-Sigma Data Converters. Theory, Design and Simulation*. Piscataway, NJ: IEEE Press.

Parker, A. J. and Newsome, W. T. (1998). Sense and the single neuron: probing the physiology of perception. *Annu. Rev. Neurosci.* **21**, 227–277.

Poor, H. (1994). *An Introduction to Signal Detection and Estimation*, 2nd edn. New York: Springer.

Saunders, J. and Bastian, J. (1984). The physiology and morphology of two types of electrosensory neurons in the weakly electric fish *Apteronotus leptorhynchus*. *J. Comp. Physiol.* A **154**, 199–209.

Scheich, H., Bullock, T. H. and Hamstra, R. H. (1973). Coding properties of two classes of afferent nerve fibers: high frequency electroreceptors in the electric fish, *Eigenmannia*. *J. Neurophysiol.* **36**, 39–60.

Shadlen, M. N. and Newsome, W. T. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation and information coding. *J. Neurosci.* **18**, 3870–3896.

Shumway, C. (1989). Multiple electrosensory maps in the medulla of weakly electric gymnotiform fish. I. Physiological differences. *J. Neurosci.* **36**, 39–60.

Turner, R. W. and Maler, L. (1999). Oscillatory and burst discharge in the apteronotid electrosensory lateral line lobe. *J. Exp. Biol.* **202**, 1255–1265.

Turner, R. W., Maler, L., Deerinck, T., Levinson, S. R. and Ellisman, M. H. (1994). TTX-sensitive dendritic sodium channels underlie oscillatory discharge in a vertebrate sensory neuron. *J. Neurosci.* **14**, 6453–6471.

Turner, R. W., Plant, J. R. and Maler, L. (1996). Oscillatory and burst discharges across electrosensory topographic maps. *J. Neurophysiol.* **76**, 2364–2382.

Victor, J. D. and Purpura, K. P. (1996). Nature and precision of

temporal coding in visual cortex – a metric space analysis. *J. Neurophysiol.* **76**, 1310–1326.

Wessel, R., Koch, C. and Gabbiani, F. (1996). Coding of time-varying electric field amplitude modulations in a wave-type electric fish. *J. Neurophysiol.* **75**, 2280–2293.

Wiener, N. (1949). *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. New York: John Wiley.

Zakon, H. H. (1986). The electroreceptive periphery. In *Electroreception* (ed. T. H. Bullock and W. Heiligenberg), pp. 103–156. New York: John Wiley.