JOURNAL OF NEUROPHYSIOLOGY
Vol. 66, No. 5, November 1991. *Printed in U.S.A.*

Representation of Sensory Information in the Cricket Cercal Sensory System. II. Information Theoretic Calculation of System Accuracy and Optimal Tuning-Curve Widths of Four Primary Interneurons

FRÉDÉRIC E. THEUNISSEN AND JOHN P. MILLER
Department of Molecular and Cell Biology, Division of Neurobiology, University of California, Berkeley, California 94720

SUMMARY AND CONCLUSIONS

1. Principles of information theory were used to calculate the limit of accuracy achievable by a subset of the wind-sensitive primary interneurons in the cricket cercal sensory system. For these calculations, an ensemble of four neurons was treated as an information channel, which encoded the direction of air-current stimuli for a defined range of air-current velocities. The specific information theoretic parameter that was calculated was the "transinformation" or "mutual information" between the air-current directions and the neuronal spike trains, which were characterized in the preceding report. Under the assumptions used for these calculations, the ensemble of four interneurons was demonstrated to be capable of encoding between 4.2 and 3.5 bits of information about wind direction. This corresponds to an average directional accuracy of 4.7 and 7.7°, respectively.

2. The same principles were applied to estimate the extent to which any variation in the width of the tuning curves would affect the transfer of information. As the widths of simulated tuning curves were varied, the mean ensemble accuracy showed a clear global maximum. This maximum corresponds to tuning curves widths of 110° wide (at half maximum), which was remarkably close to the actual mean widths of the tuning curves observed in the cricket of 130°.

3. The effect of varying the parametric "spacing" of the tuning curves within the stimulus range was also examined through a series of simulations. The configuration allowing the maximum information transfer corresponded to equal spacing of the tuning curves around the stimulus range (i.e., 90° separation of peak sensitivity points). This theoretically optimum spacing corresponded exactly to the values observed in the experiments presented in the preceding report.

4. These simulations also showed that the degradation in the accuracy resulting from a shift in the tuning-curve spacing would depend on the plasticity of the higher order decoder of directional information. If there were no plasticity in the interneurons making up the higher order decoder, then the accuracy would be degraded by 50% for a mean tuning-curve shift of only 3.5°. However, if the higher order decoding network were capable of being reoptimized to any arbitrary shift in tuning curves, the degradation in attainable accuracy would be much less severe as shifts of up to 10° would result in virtually no degradation in the accuracy.

5. From these results, two general conclusions can be drawn about the coding of specific stimulus parameters by arrays of sensory cells. First, the effectiveness of the coding of a stimulus parameter by an ensemble of cells with broadly overlapping tuning curves is strictly limited by the intrinsic variance in the cells' responses. Second, the robustness of any sensory system to shifts in tuning-curve characteristics can be greatly enhanced by allowing the higher order "decoder" networks to be capable of reoptimization.

INTRODUCTION

In the preceding paper, six interneurons of the cricket cercal sensory system were characterized in terms of their responses to wind stimuli of different peak velocities and directions. The cells were divided into two classes, on the basis of their wind velocity sensitivity. The "low-velocity" class is composed of four interneurons that had sensitivity to stimuli spanning the entire 360° range of possible directions in the horizontal plane. The stimulus-response curve of each cell had a characteristic "tuning-curve" (or "receptive-field") shape typical of many sensory systems. For each of these four cells, the response was maximal for a particular air-current stimulus direction, and the response decreased in a systematic fashion as the stimulus was rotated from this optimal direction. The tuning curves of these four cells were broad and equally spaced in angular separation. It is assumed that the ensemble response of these four cells encodes information about the direction and the velocity of air-current stimuli, and questions concerning the directional accuracy achievable by this subsystem of four cells were raised.

The goal of the study reported here was to determine the limits of directional accuracy, taking into account not only the characteristic responses of these four cells but also the variance in their responses. Further, we wanted this estimation of the limiting accuracy to be model independent in the sense that it would not depend on any specific model for "decoding" by higher order neurons. A measure that satisfies all these requirements is the statistical quantity defined as "transinformation" in information theory. This measure is uniquely dependent on the conditional probabilities of the response for specific stimuli and yields a quantitative measure of the maximum discriminatory resolution of the system. In the case of this sensory system, this measure can be translated into degrees of directional accuracy.

Principles of information theory were also used to investigate the optimality of certain parameters characterizing the operation of this system. Previous theoretical studies have demonstrated the relatively high degrees of accuracy and noise tolerance achievable within an array of sensory

cells that have broad, overlapping response curves. This form of coding is sometimes referred to as "coarse coding" (Hinton et al. 1986; Heiligenberg 1987). Inspired by such studies, we carried out simulations to investigate the dependence of system accuracy on the widths and the relative spacing of the tuning curves. In all cases the actual values of the parameters observed in the real system were found to be very close to the values that yield the theoretical maximum information transfer for this system.

METHODS

Information theory

The principles of information theory were described by Shannon (1948). His purpose was to find a quantitative measure for the amount of information carried by the "symbols" used in any kind of communication scenario, and to calculate the rate at which information could be transmitted by a temporal sequence of such symbols. This is essentially the same problem we considered in this study, i.e., we wished to determine the amount of information about stimulus direction encoded in the neural spike train responses of four sensory interneurons in this system. In the following sections we briefly summarize the aspects of information theory relevant to our investigation, review recent applications of information theory to similar questions in other neurophysiological studies, and then present in that theoretical and historical context the methods and assumptions on which our own studies are based.

ENTROPY AND TRANSINFORMATION. Consider a source of symbols (i.e., a "transmitter"), which we will call X, and its corresponding set of symbols $\{x_j\}$ used in the communication. An essential idea in information theory is that the amount of information conveyed by a symbol is inversely related to its probability of occurrence. In other words, receiving a symbol that a priori is known to have very low probability of occurrence will yield a large amount of information about the state of the source.[1] The information conveyed by one particular symbol x_j is written in binary units as

$$i(x_j) = \log_2\left(\frac{1}{p(x_j)}\right) = -\log_2 [p(x_j)] \qquad (1)$$

where $p(x_j)$ is the probability of occurrence of symbol x_j.

A quantity that is important for our subsequent analysis is the *average* amount of information conveyed by each of the different symbols transmitted by the source X. This average is simply the sum, over all possible symbols, of the information conveyed by each symbol times that symbol's probability of occurrence

$$H_x = \sum_{j=1}^{n} -p(x_j) \log_2 [p(x_j)] \qquad (2a)$$

where n is the total number of discrete symbols in the set. This quantity H_x was defined by Shannon as the "entropy" of the transmitter X, because it corresponds to the amount of a priori uncertainty about the source (and is also equivalent to the a posteriori average amount of information that can be obtained from its set of symbols $\{x\}$). H_x can also be thought of as representing the fractional number of binary digits (or bits) needed to encode the symbols of X with an optimal encoder.

[1] This makes intuitive sense, considering our own written language. The letter "x" occurs in many fewer words in a standard English dictionary than does the letter "e." If you were presented with a word having all but one letter masked out, you could offer a much better "guess" about what that word was if the exposed letter were an "x" rather than an "e."

In many cases the symbols of interest cannot normally be represented as a finite set of discrete symbols $\{x\}$ but constitute instead an infinite ensemble representing the values of a continuous parameter. In the case of the cercal sensory system, for example, the symbols corresponding to the air-current stimulus direction should, presumably, be allowed to represent any arbitrary angle in the horizontal plane. In such cases involving continuous variables, $p(x)$ becomes a continuous probability function and *Eq. 2a* can be rewritten in its equivalent integral form

$$H_x = \int -p(x) \log_2 [p(x)] dx \qquad (2b)$$

where the integration is carried out over the whole range of the parameter to be represented.

Although it may seem inappropriate to characterize aspects of neurally coded information in terms of binary bits, it is actually quite reasonable from a biological perspective. Consider, for example, our problem of characterizing the intrinsic resolution with which the cercal sensory system could represent the direction of an air current. If the system's resolution was limited to the segregation of air currents directed at the front of the animal from air currents directed at the rear of the animal, then the system would effectively be able to divide the whole stimulus range of 360° into only two "bins." In other words the system could be thought of as resolving wind direction with one bit of accuracy. If the stimulus range could be divided into four bins, encoding the direction information would require two bits. Every doubling in the number of bins would imply one more bit of information. Fractional binary quantities are also possible: a division of the range into five bins corresponds to 2.32 bits.

In general, however, sensory systems should not be thought of as segregating the ranges of relevant stimulus parameters into fixed numbers of bins with discrete boundaries. Rather, any system will display intrinsic limitations in the *reliability* or *probability* with which two different stimuli having slightly different parameters can be distinguished. To illustrate this point, consider a typical series of stimulus-response measurements from the preceding report (Miller et al. 1991), in which 32 identical air-current stimuli were directed at a cricket from a single direction. A *range* of responses were elicited that were distributed around the *mean* response according to a Gaussian probability function. Consequently, the corresponding set of 32 estimates of stimulus direction derived from this probabilistic response set by any subsequent optimal decoder would itself be characterized by a probability distribution. In general, the probability distribution of these decoded direction estimates would be centered at the correct direction and would have a spread that depended on *1*) the spread of the probability distribution of the neuronal responses and *2*) the rate at which the mean ensemble response changed as a function of stimulus direction. In such cases, the accuracy of the system in distinguishing different stimuli is usually defined in terms of the spread of this probability function characterizing the decoder estimates of a stimulus direction: the wider the spread, the lower the corresponding accuracy. Roughly speaking, a stimulus range can be thought of as being fractionated with a resolution equivalent to the width of this probability distribution, and the mean resolution of the system in bits is therefore approximately equal to the logarithm (base 2) of the extent of the stimulus parameter range divided by the mean width of this probability distribution. A rigorous derivation of the relation that is relevant to our case (i.e., relating the width of a Gaussian probability function to the information theoretic quantities in bits) is summarized later in the text and is presented in detail in the APPENDIX.

The different forms of *Eq. 2* have been used in several theoretical studies to calculate the upper limit of the information transmission rates in several neural systems (for example, see McKay

and McCulloch 1952; Rapaport and Horvath 1960). Such calculations were carried out by multiplying the mean information per symbol (i.e., H_x) by the presumed rate at which the symbols were being transmitted. Such calculations, however, required ad hoc assumptions about the nature of the symbols used for the neural coding scheme and were therefore extremely problematic. These calculations yielded upper limits to the information transfer rate that may be extremely and unreasonably high, and therefore essentially meaningless from a physiological standpoint. The basis for this overestimation is that the summation (or integration) limits over which H_x was calculated spanned the entire range of possibly "decodable" symbols in the transmitted signals, without taking into account the "meaning" associated with each symbol.[2]

A more appropriate and model-independent quantity related to neural processing is the amount of information transferred *between* layers of neurons, rather than the total entropy of the signal stream transmitted *by* one layer. To calculate this quantity of information transfer, the proportion of the observed signal variance that does *not* encode independent information about the transmitter must somehow be subtracted from the entropy of the true symbol set itself. Such a calculation can be accomplished in a straightforward manner. Because we wish to characterize a transfer of information from one layer of neurons to another, a second set of symbols at a "receiver" must be defined. Assume that the meaningful information contained within the whole set of transmitted symbols $\{x\}$ is encoded as a set of symbols $\{y\}$ at some subsequent level in the system. The "conditional entropy"

$$H_{y|x} = \int\int -p(x, y) \log_2 [p(y|x)] dxdy \qquad (3)$$

measures the variance or uncertainty that remains in the encoded set of symbols $\{y\}$ when the identity of the symbols of the source X is known. The transinformation, then, can simply be defined as

$$T_x = H_y - H_{y|x} \qquad (4)$$

In other words, H_y is a measure of the total range of symbols used by the encoder (including all redundant and/or "nonsense" symbols), and $H_{y|x}$ is the portion of this range of symbols that is *not* correlated to the set of symbols $\{x\}$. T_x is therefore the relevant biological quantity and measures the information about the source X that is actually transmitted through the channel.

This measure of transinformation, also called "mutual information," can be applied to any stochastic system that has a well-defined input and output. In neurobiological systems, the output is appropriately defined as some parameter of the electrophysiological signal (i.e., membrane potential or spike pattern) recorded from a single neuron or a group of neurons. The input to the system can correspond either to some relevant electrophysiological parameters of presynaptic neurons or, in the case of sensory systems, to the range of relevant sensory stimuli. Note that *Eq. 3* is written in its integral form, which is appropriate for cases in which the x and y symbols correspond to parameters of continuous functions. In the general formulation of information theory, x and/or y could represent discrete symbols, in which case the integration signs would be changed to summation signs.

PREVIOUS CALCULATIONS OF TRANSINFORMATION IN NEURAL SYSTEMS. In their pioneering work on the application of transinformation in neurobiology, Eckhorn and Pöpel (1974, 1975) calculated the net transinformation and the rate of transinformation flow at the level of the retinal ganglion cells and lateral geniculate nucleus (LGN) cells of the cat. The experimental input to the system was a flashing light (with a pseudorandom rate) represented as a binary function of time. The output of the system was taken as the neural impulse trains, also represented by a binary function of time. By taking into account the conditional probability of spike occurrence given previous spike activity, Eckhorn and Pöpel obtained measurements of information transfer without making any arbitrary assumptions about the nature of the neural code. Along the same lines, Richmond and Optican (1990) showed that a significant part of the information conveyed by single cells of V1 about a complete set of multidimensional spatial visual stimuli was present in the temporal patterns of the spike trains. To get around the overwhelming dimensional size of the representation when a binary code is used, and to gain more insight on the actual encoding of the information, they used a principal component decomposition of the temporal spike pattern. They found that three to four principle components were sufficient to encode all the relevant structure of the spike pattern.

In these studies listed above, the approach taken in examining neural information processing was to calculate the information contained in the ensemble of spike-train patterns elicited by specific stimulus sets. De Ruyter van Steveninck and Bialek (1988) used the same transinformation principles but applied those principles in a way that can be thought of as being the inverse of the above approach. In their studies, de Ruyter Van Steveninck and Bialek calculated the information contained in all sensory stimuli that could have evoked specific, observed spike-train responses. By doing the calculations over all possible response patterns of up to three spikes, they provided some insight into the actual neural code used in their preparation and addressed questions related to real-time decoding of neural response patterns.

Fuller and Looft (1984) extended the principles defined in the original work of Eckhorn and Pöpel to the multineuron case. They calculated the transinformation for a group of cutaneous receptor cells of the cat to test the two prevalent assumptions: *1*) that the receptor responses are conditionally independent and *2*) that the information about the stimulus is encoded by the mean rate of the neuron spike train. They found that the first assumption is approximately true, but that a mean rate code will only transmit part of the information.

TRANSINFORMATION IN THE CRICKET CERCAL SYSTEM. The first goal of the work presented here was to determine the limit of accuracy with which four sensory interneurons in the cricket cercal system could encode information about wind stimulus direction. To determine this limit, the transinformation between the sensory stimuli and interneuronal output was calculated, on the basis of experimental measurements presented in the preceding report (Miller et al. 1991). For these studies, the source or transmitter X was considered to be the air-current stimulus generator, and the set of symbols $\{x\}$ corresponded to all possible *directions* of the transient, unidirectional air currents used in these experiments (see Fig. 1). In other words, the input symbol variable x in *Eqs. 1–4* corresponds to the *continuous range of stimulus directions* in the horizontal plane, which we will subsequently refer to

[2] To illustrate this problem with an example using a set of discrete symbols, consider the task of deciphering a handwritten message. If the written message contained randomly intermixed cursive script and roman printed characters, then a reader unaware of the equivalence between the 2 sets of symbols would calculate H_x using *Eq. 2a* summing over a possible set of 52 characters instead of the actual "meaningful" number of 26. In other words, even if 2 different symbols have the same meaning, they would each contribute independently to the net mean entropy calculated by *Eq. 2a*. Furthermore, in previous studies using H_x as a means to quantify neural information, the maximum information transmission rates were estimated under the assumption that all possible neural symbols had the same probability of occurrence. In our illustrative example about the handwritten message, this would be equivalent to assuming that all of the letters of the alphabet have equal probability of occurring in each word. This mistaken assumption would further contribute to an overestimation of the information transmission rate. Because we have no way of knowing with any certainty what the significant symbols are in neural spike trains, then the use of H_x to characterize neural information transmission will lead to an equivalent overestimation problem.

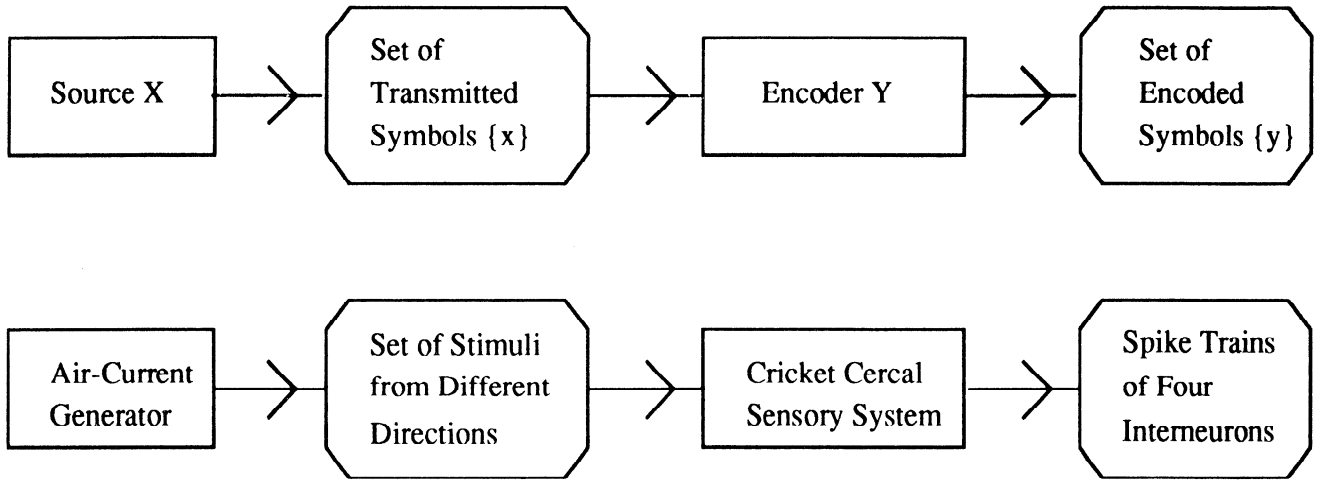

FIG. 1. Schematic flow chart of the information transfer in the cricket cercal system and the corresponding mathematical symbolic notation used in the transinformation calculation.

as θ. The "output" symbol variable y in *Eqs. 3* and *4*, which represents the encoded information about stimulus direction, corresponds to the *ensemble response* of a set of interneurons during an interval after the stimulus, which we will subsequently refer to as r. The interneurons for which we performed the transinformation calculation is the four-member class of 10-2 and 10-3 cells described in the preceding report (Miller et al. 1991).

This transinformation calculation is considered as representing the *limit* to the system's accuracy, because it can be proven that no possible neural (or engineered) decoder could extract more information from the spike trains (under our set of limiting assumptions) than can be calculated to be "encodable" according to *Eq. 4*. (Of course, the neural circuitry at higher levels might not be capable of decoding all of the available information; only behavioral studies could ultimately test this possibility.) It is in this limiting sense that the transinformation calculation yields a model-independent measure of the accuracy. However, as is the case for any measure of sensory performance, the transinformation calculation is very much dependent on the assumptions made when choosing the relevant coding parameters of the spike train, as well as any assumptions that would limit the set of possible stimuli to the system. A list of our assumptions, and a consideration of how they could have influenced our calculations of the accuracy of this system, are as follows.

Assumptions

REPRESENTATION OF THE STIMULUS DIRECTIONS AS THE INPUT SYMBOLS. For our analysis of directional accuracy to be rigorous, the ensemble response of the system should, in theory, be observed for all possible wind-current velocity profiles at all directions, and the contribution of each different stimulus-response measurement to the overall transinformation calculation should be weighted by that stimulus profile's actual probability of occurrence in the cricket's real environment. In practice, such a complete set of experimental measurements was not possible. Therefore we restricted our stimulus shape to the half-cosine waveforms described in the previous report, and the peak velocities of the waveforms were restricted to the optimal sensitivity range within which the responses of the cells varied as the log of velocity (Fig. 3, Miller et al. 1991). By restricting our calculations to this limited set of inputs, we made the following implicit assumptions: *1*) the general aspects of the tuning curves (i.e., their shapes, widths, and variances) were relatively independent of the precise velocity profiles of different air-current stimuli, and *2*) other aspects of the neuronal responses within the system (for example, the relative activity levels between the 9-2 and 9-3 cell class and the 10-2 and 10-3 cell class) encoded independent information about air-current direction and velocity for stimuli with velocities outside of the "log-linear" operating range of the 10-2 and 10-3 cells. The first assumption was observed to be valid with respect to two particular classes of stimuli: the half-cosine waveforms (used for the calculations presented here) and for constant velocity "step" waveforms (like those used for Fig. 2*A* of Miller et al. 1991). The second assumption also seems reasonable and is considered in greater detail in the DISCUSSION.

REPRESENTATION OF THE CELL RESPONSES AS THE OUTPUT SYMBOLS. As indicated above, the output variable r must be some representation of the recorded activity of the nerve cells, which has relevance in the information transfer process. As shown elsewhere (Eckhorn et al. 1976), a "signal code" (binary function of time of the spike occurrences) would be the most complete representation. In our study the analysis was limited to the mean elicited spike counts in an interval after the stimulus. We believe that this is a very reasonable first-order representation for the cells in this particular neural subsystem, because their spike-train responses displayed very regular firing rates (see Miller et al. 1991). The integration time for the estimation of the mean spike count was 100 ms. The choice of the length of the integration time constitutes an additional assumption about the resolution of the system. We are, in effect, assuming that higher decoding neurons could be able to integrate this response for ~100 ms. If the integration time is reduced, the information transfer would also be reduced, but this effect is less significant for regularly firing units such as these cells.

For our calculations, therefore, the output symbol variable y in *Eqs. 3* and *4* can be substituted with the ensemble response vector $r = (r_1, r_2, r_3, r_4)$ where r_1 is the mean spike count (MSC) of 10-3 left (10-3L), r_2 is the MSC of 10-2L, r_3 is the MSC of 10-2 right (10-2R), and r_4 is the MSC of 10-3R. Each of the four cells had a maximum MSC for its optimal stimulus direction, and the responses fell off symmetrically on either side of that optimal direction (Fig. 5 in Miller et al. 1991). The curves of MSC versus wind direction are therefore considered to represent "directional sensitivity curves," or "directional tuning curves."

The calculation of transinformation and systematic accuracy required a knowledge of two features of the cells' response characteristics: *1*) the shape of these directional tuning curves, plotted as the mean spike count values and *2*) the values for the response variance at each stimulus orientation for each cell. These data were presented in the preceding paper and are dealt with in our computations as follows.

DIRECTIONAL TUNING CURVES OF THE FOUR INTERNEURONS. As described in the preceding report, analytic functions corresponding to truncated cosine functions were fitted to the 10-2 and 10-3 cells' directional tuning curves. For reasons of computational efficiency, the calculations presented in this report were carried out on the basis of these analytic functions, rather than with the observed data points themselves. The use of continuous functions allowed us to interpolate reasonable values for the MSCs of the interneurons at intermediate stimulus directions and for a range of velocities for which real experimental data were not obtained. The use of these model curves was considered valid, because *1*) the fitted curves were statistically indistinguishable from the real data curves, and *2*) small deviations from those curves would not in any case be expected to significantly alter the results of our calculations.

The general form of the analytic function used to represent the stimulus-response characteristics of each individual cell in the four cell ensemble was as follows

$$r = [r_{max}/(1 - a_t)] * [\cos(\theta - \theta_{max}) - a_t] \quad \text{for} \quad \cos(\theta - \theta_{max}) > a_t \quad (5a)$$

and

$$r = 0 \quad \text{for} \quad \cos(\theta - \theta_{max}) < a_t \quad (5b)$$

where r is the response amplitude (i.e., mean spike count), equivalent to the y symbol; r_{max} is the maximum response amplitude; θ is the stimulus direction in radians, equivalent to the x symbol; θ_{max} is the optimal stimulus direction evoking the maximum response; and a_t is the threshold stimulus value for evoking spike activity.

In this function, the response r can be thought of as being the resultant sum of excitation from sensory receptor afferents and inhibition from local interneurons. The threshold value is presumably determined by the relative amount of excitation and inhibition, as well as by the spike threshold of the interneurons. The $[r_{max}/(1 - a_t)]$ coefficient can be thought of as a gain factor.

As well as being plausible from a physiological standpoint, this formulation of the equation is convenient for parametric variation simulations. Operationally, changing the value of the threshold parameter a_t modifies the width of the tuning curve but maintains the maximum response at a constant value by adjusting the gain.

Because the experimental directional tuning curves were only measured for one particular peak stimulus velocity (i.e., the velocity at which the MSC response was at ~60% of saturation), the parametric representation allowed us to simulate the response of the system for a range of velocities that were not actually characterized with full tuning curves. To carry out these simulations, r_{max} was increased or decreased proportionally to the log of the stimulus velocity, according to the relationship demonstrated in the previous report. In doing so, we effectively assumed that, within this log-linear region of the cells' velocity sensitivity, the shape of the tuning curves varied only by a scaling factor.

Regardless of the plausibility of *Eq. 5* as a suggestive model for directional selectivity, the form of the equation is essentially arbitrary; a number of other functions could have been used to fit the observed data equally well. A significant concern does arise, however, when considering how the shapes of these different functions change as the relevant width parameter is changed for the parametric variation simulations. To verify that the results were not an artifact of the particular analytic function used to represent the tuning curves, some calculations were repeated with the use of a function that maintained a steeper slope at the tails of the tuning curves as widths were increased. The function we chose was a cosine with varying period and zero threshold. The MSC response in this case was modeled as

$$r = r_{max} * \cos[(2\pi/\Omega) * (\theta - \theta_{max})] \quad (6)$$

where Ω was the period in radians.

RESPONSE VARIANCE OF THE FOUR INTERNEURONS. As discussed above, calculations of the transinformation and system accuracy require a knowledge of the variance in the responses of the individual cells to identical repeated stimuli. This is intuitively understandable. Consider a function constructed from the sum of two Gaussian functions that have equal amplitudes but different mean values. The ability to resolve the two functions depends not only on the *separation* of their means, but also on the relative *widths* of the functions. The case of our neural tuning curves is entirely equivalent: the ability to distinguish the directions of different stimuli on the basis of the mean firing rates of a neuron (or neurons) is dependent on the variability in responses from one stimulus presentation to the next.

In terms of the information theoretic calculations, the variance in a cell's response r to repeated stimuli at a direction θ and a velocity v is used to define the conditional probability distribution $p(r|\theta, v)$. With the use of the simplest assumptions, consistent with the variance measurements reported in the preceding report, we assumed that each cell's responses varied around it's mean response for that particular direction and velocity with a discrete normal distribution. Because the MSC could never be negative, the distributions were truncated at zero. In other words, $p(r|\theta, v) = 0$ if any of the components of the ensemble response r were <0. The spread of the distribution was determined by the variance of the MSCs reported in the preceding paper (Miller et al. 1991). As shown in Fig. 6 of that paper, graphs of the standard deviation of the MSC versus the amplitude of the MSC from each cell could be fitted with a linear function. Thus we could estimate the spread of $p(r|\theta, v)$ by first calculating the MSC at that particular direction and velocity and then obtaining the variance from its linear dependence on the MSC. To estimate the directional accuracy in the real system, we chose the variance relationship from the cell having the highest variance.

The deviation from the normal distribution was assessed by calculating the skewness coefficient. In all cases, the skewness was always less in absolute value than the expected variance of the skewness for a normal distribution calculated from 32 points of data.

Note that the functional representation of the conditional probability distributions as normal distributions was not strictly necessary for the directional sensitivity calculation. The probability distribution could also have been extracted directly from the data used to calculate the variance. Our functional representation simply smoothed out these probabilities and allowed a rigorous and consistent model for the distributions at intermediate values of the MSC.

COVARIANCE IN THE RESPONSES OF DIFFERENT CELLS. Because of the limitation of recording from only one cell at a time, any possible covariance that might have existed between cells could not be measured. Thus an implicit assumption of these calculations was that the cells responded independently for a given directional stimulus. A consideration of the covariance in cell activities would, in effect, have allowed a determination of the proportion of the variance in the MSCs that was correlated between cells. An optimal decoder could, in certain cases, use such information about the covariance to enhance accuracy. (Consider, as an engineering example, a bipolar amplifier set to reject the "common mode noise" on both inputs.)

By assuming no covariance between activities of different cells, by using only the mean firing rates for our representation of the spike train, and by choosing the largest variance values measured in our experiments, we are, in effect, calculating the lowest bound of the maximum directional accuracy of our system. On the other hand, our choice of the integration time over which to estimate the MSC and the restriction of the stimuli to a limited velocity range might have lead to overestimations of this lowest bound.

Calculation methods

The directional transinformation ($T_\theta = H_r - H_{r|\theta}$) can also be written as

$$T_\theta = \int p(\theta) \left[\int \log_2 \left(\frac{p(r|\theta)}{p(r)} \right) p(r|\theta) dr \right] d\theta \quad (7)$$

or

$$T_\theta = \int p(\theta)\, T(\theta)\, d\theta \quad (8)$$

where

$$T(\theta) = \int \log_2 \left(\frac{p(r|\theta)}{p(r)} \right) p(r|\theta)\, dr \quad (9)$$

is the transinformation for a particular input symbol θ (Mansuripur 1987). $T(\theta)$ is also referred to as the "partial transinformation." This is the form that was used in our calculations. In these equations, the conditional probability $p(r|\theta)$ was calculated by taking the average of $p(r|\theta, v)$ over all velocities in the velocity interval of interest[3]:

$$p(r|\theta) = \int p(r|\theta, v)\, p(v) dv \quad (10)$$

To calculate the partial transinformation for a given direction θ, a four-dimensional nonanalytic integral had to be calculated (i.e., the integrand in *Eq. 9* is a function of r, the 4-dimensional response vector representing the ensemble MSC of the 4 interneurons). The fastest numerical method for estimating $T(\theta)$ is a Monte Carlo integration where $p(r|\theta)$ is used as the "importance function" (Davis and Rabinowitz 1984; Gillespie 1975). Finally, to obtain the overall transinformation, a simple one-dimensional averaging integral (*Eq. 8*) had to be performed. We calculated the overall transinformation, which we will subsequently refer to as T_θ, by obtaining $T(\theta)$ for enough values so that their mean value would not change. We found that simulating the responses at 128 directions was sufficient.

Correspondence between T_θ in bits and angular resolution in degrees

As discussed earlier, it is reasonable to consider the value of transinformation in bits as being directly related to the angular resolution of the system. For example, if the average directional transinformation is approximately four bits, the simplest interpretation is that the total stimulus range (i.e., 360°) could be divided by the four cells into 2^4 (i.e., 16) different directional bins, each of which would be 22.5° wide (corresponding to a standard deviation of 6.5°). A more rigorous interpretation of the measure of transinformation is to equate it to the "angular resolution" or "standard deviation" of the error with which an optimal decoder could compute the stimulus direction from the mean ensemble spike count response of these four cells. The formal relation between the *transinformation values* we calculated and the equivalent *accuracy of the direction calculations* by any presumed optimal decoder depends on the shape of the probability function that characterizes the output of an optimal decoder to identical repeated stimuli. Assuming a Gaussian distribution for this probability function (which was, in fact, a good approximation for a simulated decoder based on the minimum mean square error), the number of bits of transinformation in the response with standard deviation σ is approximately given by

$$T = \log_2 (\sqrt{2\pi}/\sigma) - 1/[2 \ln(2)] \quad (11)$$

as long as σ is less in magnitude than 2π radians (see APPENDIX). Four bits of resolution, for example, would correspond to a Gaussian error with a width of 5.4°.

RESULTS

Directional accuracy for stimuli with a "known" velocity

The directional transinformation T_θ was calculated according to *Eq. 9*, using stimulus-response parameter values equivalent to the experimental observations presented in the preceding report (Miller et al. 1991). For this analysis, the calculations of transinformation *at each direction* were based on the cells' responses to *stimuli having known direction and velocity*. (The stimulus velocity used for the experiments was set to elicit a response equal to ~60% of a cell's "saturating" response in its optimal direction.) The curves plotted with solid lines in Fig. 2 are the model functions representing the cells' directional tuning curves at this velocity, calculated from *Eqs. 5a* and *5b*, with parameters set to best fit the physiological data. The transinformation (in bits) is plotted as the dashed curve. Transinformation varied between 3.7 and 4.8 bits, with a mean of 4.22 bits. This fluctuation would presumably result in a slight unevenness in the attainable accuracy at different stimulus directions.

Directional accuracy for stimuli with undetermined velocities

Additional calculations were performed according to *Eqs. 9* and *10* to determine the extent to which *uncertainty in stimulus velocity* would decrease system accuracy in de-

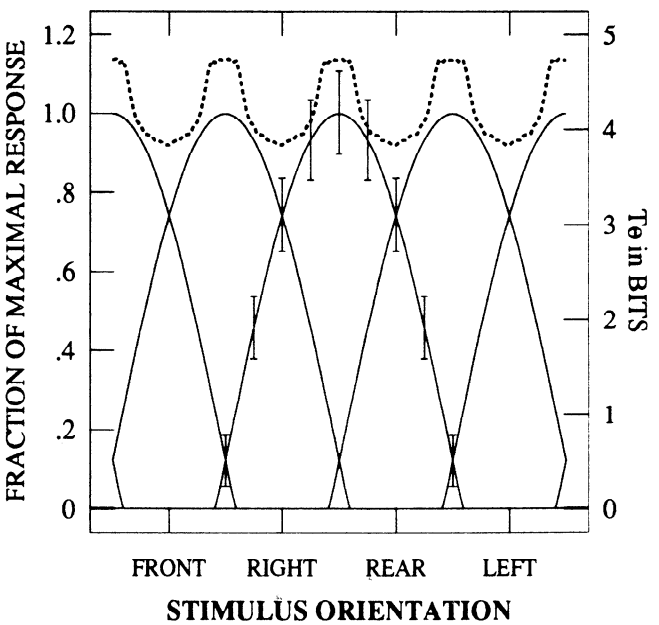

FIG. 2. Mean spike count (MSC) in the responses of the 4 10-2 and 10-3 cells (solid curves) to wind stimuli from different directions, and information transfer of the system (dotted line) in bits for a fixed velocity set at 60% of the saturating velocity. Smooth response curves were obtained by fitting a cosine function to the experimental data. Error bars (shown for only 1 cell) indicate the mean variance in the cells' responses. The mean value of transinformation is 4.2 bits.

[3] Note that in taking the average of over all velocities we are not including any information about the velocity in our transinformation calculation. If one were interested in the encoding of the velocity as well, *Eq. 9* would become $T(\theta, v) = \int \log_2 [p(r|\theta, v)/p(r)]\, p(r|\theta, v)\, dr$. $T(\theta, v)$ would then be integrated over all velocities and directions to obtain the total average transinformation.

termining *stimulus direction*. For this analysis, the calculations of transinformation at each direction were based on simulated responses to several stimuli having identical directions but different velocities, as described earlier. By varying the range of velocities from which each stimulus-response set was selected, different levels of uncertainty about stimulus velocity could be simulated. Three different sets of such calculations were performed, with velocity ranges set as shown in Fig. 3*A*. The stimulus-velocity range for each set was centered on the "standard" stimulus velocity used for the calculations presented in the preceding section. This standard stimulus velocity, indicated in Fig. 3*A* with an asterisk, was the velocity that elicited a response equal to 60% of a cell's "saturating" response in its optimal direction. The limits of the stimulus-velocity range for each of the three calculation sets were adjusted, as shown in Fig. 3*A*, to bracket this standard stimulus velocity by amounts that elicited responses corresponding to ±12, ±24, or ±36% of the MSC at the saturating stimulus velocity around the MSC at the standard velocity.

The results are shown in Fig. 3*B*. In this graph, the average directional transinformation (in bits) across all directions is plotted versus the width of these velocity uncertainty ranges. In other words, with zero uncertainty as to the stimulus velocity, the mean transinformation was 4.22 bits, as calculated in the previous section. As the velocity uncertainty range was increased, the mean directional accuracy decreased. However, the decrease is not as pronounced as might have been expected. Even for velocity ranges for which the MSC ranged by ±36% around the response at the standard velocity (i.e., from 24 to 96% of the cells' saturating responses, which corresponded to a velocity range from 3 to 83% of the saturating stimulus velocity), the mean transinformation decreased by less than one bit to a value of 3.5 bits.

To gain a better intuitive understanding of these plots of directional accuracy, the values for transinformation in *bits* can be translated into values for *angular resolution*, as described in METHODS. According to *Eq. 11*, this calculated range in transinformation between 3.5 and 4.2 bits corresponds to an angular resolution of between 7.7 and 4.7°, respectively, depending on the width of the velocity interval considered significant.

Dependence of transinformation on changes in tuning-curve width

As discussed earlier, the accuracy achievable by a sensory system such as this must be influenced by the relative widths of the parametric tuning curves of the constituent neurons, which in these experiments were measured to be 130°. The dependence of directional accuracy on the tuning-curve widths was assessed by a series of simulations in which the widths (measured at half-maximum amplitude) were varied between 20 and 220°. The variation was achieved by varying a_t in *Eq. 5a* between 0.97 and −1.68, respectively. The simulations were done for two different stimulus-response sets: for the standard stimulus velocity and for a velocity range corresponding to the response interval of ±24% around this standard velocity. All other parameters in the functions were kept constant.

The results of these simulations are plotted in Figs. 4, 5, and 6. In Fig. 4, the simulated tuning curves and corresponding calculations of transinformation versus direction are plotted for each of three specific tuning-curve widths. For each tuning-curve width, the data are presented in two different complementary formats: a polar plot and a Cartesian plot. These graphs illustrate the extent to which modifications of tuning-curve width effect two significant aspects of the local transinformation. First, the variation in the transinformation (plotted as solid lines) at different directions is much greater for the case of the narrow tuning curves. This would imply a significant variation of directional accuracy at different directions for this case. Second,

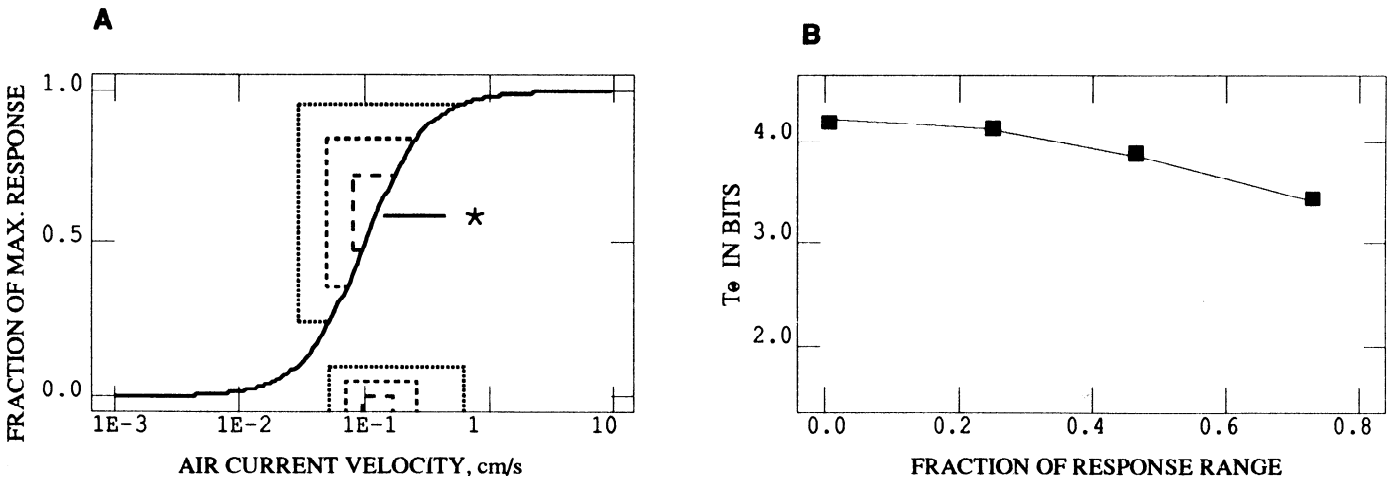

FIG. 3. Typical velocity response curve showing 3 different velocity ranges for which the transinformation calculations were performed (*A*) and plot of the mean transinformation as the velocity range is made wider (*B*). *A*: asterisk shows the point at which the response is at 60% of the saturating response. This point is called the standard velocity and is used to define the 3 progressively larger velocity intervals shown with dotted brackets. The 3 intervals correspond to velocity ranges that elicited responses in intervals centered at value of the response at the standard velocity and with width of 24, 48, and 72% of the total response range. *B*: mean transinformation values for these 3 velocity intervals, and for the case where the velocity was fixed at the standard velocity, are shown plotted against the width of the velocity interval expressed in percent of the total response range.

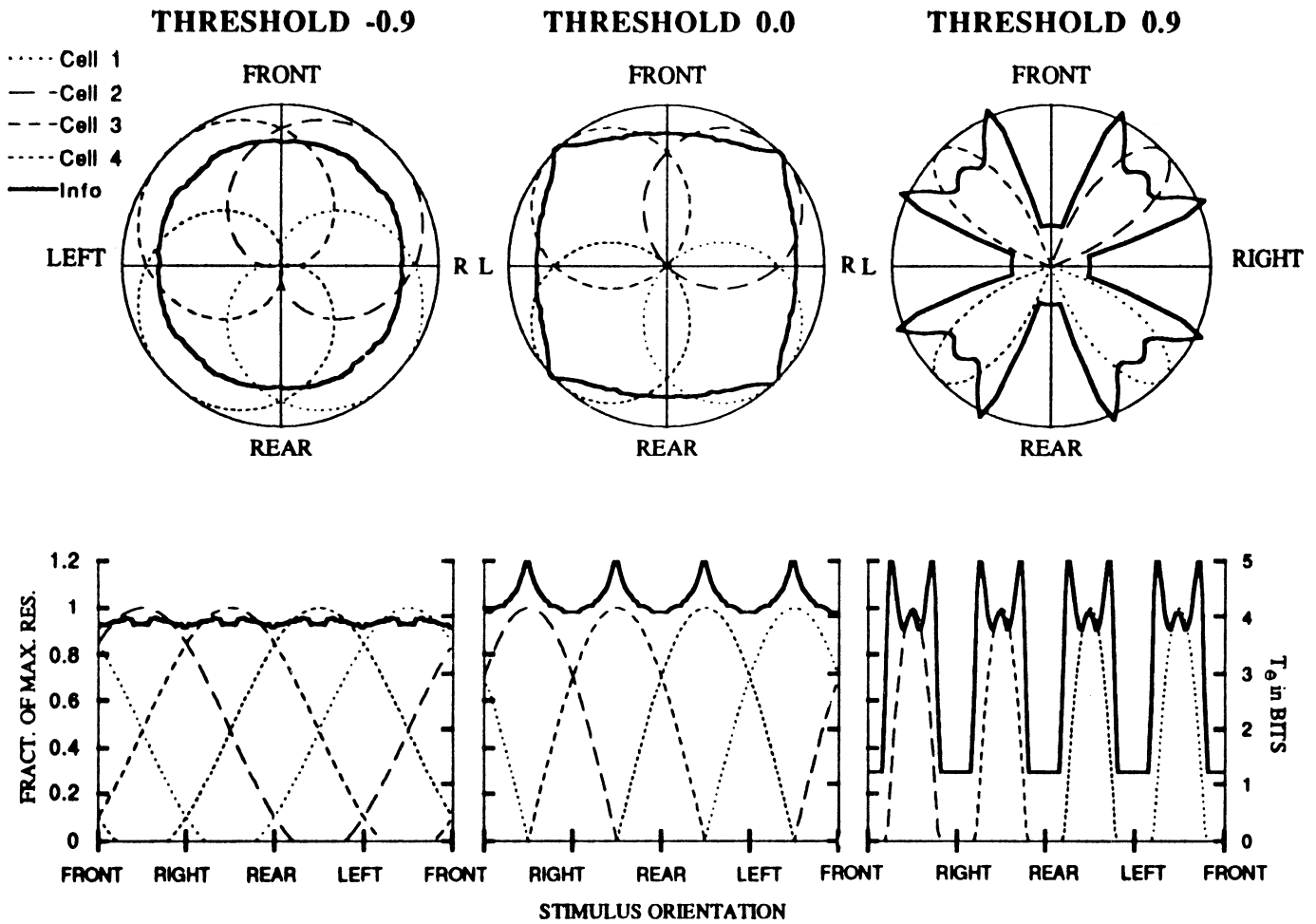

FIG. 4. Simulated response curves (tuning curves) and transinformation curves for different tuning-curve widths (i.e., different values of the threshold parameter a_t). The curves were plotted in polar coordinates in the *top panels* and in cartesian coordinates on the *bottom*. Responses of the 4 cells are shown with thin dashed lines, and the values of transinformation are shown with the thick solid lines. Changing the value of the threshold from negative to positive values narrowed the width of the tuning curves and effectively reduced the amount of overlap. As overlap was reduced, the transinformation spread became larger.

the overall mean transinformation across all directions was significantly lower in the very narrow and very wide tuning-curve sets than in the intermediate width case, which suggests that some optimal width might exist.

The results of many such calculations at a broad range of tuning-curve widths are plotted in Fig. 5 to allow a more systematic analysis of the trends suggested by Fig. 4. The *top* (solid) *curve* in Fig. 5 shows the average transinformation across all directions (in bits) for a wide range of model tuning-curve widths, calculated for the stimulus-response sets at the standard stimulus velocity. The similar curve immediately below that *top curve* shows the mean T_θ for the same range of model tuning-curve widths, but calculated instead for the stimulus-response sets chosen from the intermediate "velocity uncertainty" interval indicated above. In both cases the average transinformation showed a clear global maximum for tuning-curve widths of ~110°. This is close to the point at which $a_t = 0$ (i.e., 120° width) where the tuning curves from three cells begin to overlap. (The simulated tuning curves and calculated T_θ function for that case are shown in the *middle panel* of Fig. 4.) Note that this theoretical optimum tuning-curve width is, indeed, very close to the width of 130° measured for the actual interneurons (Miller et al. 1991).

The *bottom* (dashed) *curve* plotted in Fig. 5 represents another significant aspect of the system response characteristics: the spread (in bits) of the local transinformation value from the mean T_θ as a function of tuning-curve width (i.e., the difference between the minimum and maximum local transinformation values for the different curve widths). These data were obtained by the use of the fixed standard velocity simulation; curves generated from the different velocity uncertainty intervals were very similar. Note that the spread in T_θ was large for the relatively narrow tuning curves. As the tuning-curve widths were increased, the local transinformation curves became smoother, resulting in a decrease in the graph of the spread of T_θ values. The significance of all these results and the "structure" in these curves will be considered in more detail in the discussion.

To determine whether the results were artifacts of the particular model function used to represent the cell's tuning-curve shapes, the calculations that yielded the results plotted in Fig. 5 were repeated with the use of a different model tuning-curve function, as indicated in METHODS. In

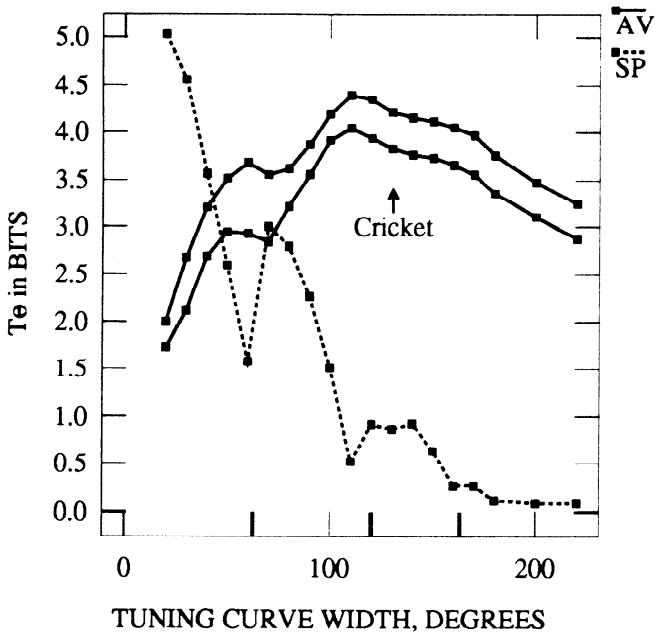

FIG. 5. Mean transinformation values (solid lines) and spread of T_θ values around the mean (dashed line) for different values of tuning-curve width. The top solid line is the mean transinformation calculated for a particular velocity set at 60% of the saturating velocity. The solid curve immediately below shows the mean transinformation calculated for an intermediate velocity interval as defined in Fig. 3. The spread was calculated as the difference between the maximum and minimum transinformation values for all possible directions for the particular velocity case. The arrow points to the transinformation value calculated for the cricket from actual experimental data. The limit of transinformation in the real system is close to the optimal point. The bold tick marks on the x-axis correspond to the critical overlap values between tuning curves of neighboring cells (see text).

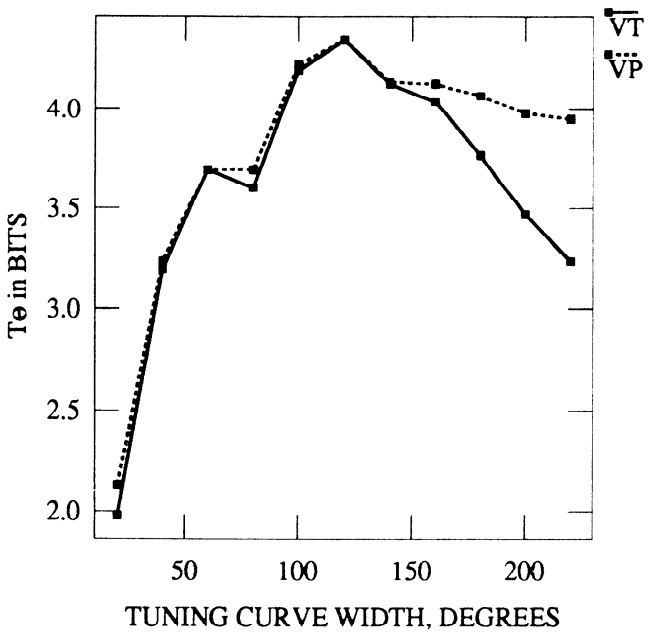

FIG. 6. Average transinformation values vs. tuning-curve widths from 2 simulation sets, where different functions were used to model the tuning curves. The solid line corresponds to simulations in which tuning curves were modeled with *Eq. 5*. The dashed line corresponds to simulations in which tuning curves were modeled with *Eq. 6*.

Fig. 6 the curve plotted with the solid line was calculated with the use of *Eq. 5* to represent the cell tuning curves, and the curve plotted with the dashed curve was calculated with the use of *Eq. 6*. Both simulations were done for the fixed "standard velocity" case. The similarity of these two curves verifies that the results were not artifacts of the particular model function used to represent the cell's tuning curve.

Dependence of transinformation on shifts of the tuning curves

Data presented in the preceding report showed that the points of peak sensitivity of the four cells' tuning curves were spaced equally around the horizontal plane at 90° intervals (Miller et al. 1991). The precise locations and separation of the curves appeared to be constrained very rigidly, as shown by the low variances in those values. Simulations by Heiligenberg (1987) demonstrated that the accuracy of a coarse-coded representation of a discretely sampled sensory continuum such as in the system studied here will depend, to some extent, on the interval spacing of the neural tuning curves. Simulations were carried out to determine the immediate effects on system accuracy that would result from short-term dynamic shifts in the relative spacing of the tuning curves.

In these simulations the mean system accuracy limit was calculated for cases in which each of the four tuning curves were shifted from their mean positions by a random amount, chosen stochastically from a Gaussian distribution with a standard deviation ranging from 0 to 10°. The effect of this additional dynamic shift was to increase the variance of the MSCs at each point. The calculations were done for $a_t = 0$ for the fixed velocity case, and the results are shown in Fig. 7. The curve plotted with a solid line corresponds to the mean transinformation (across all directions) for successively larger values of shift. The dashed curve represents the

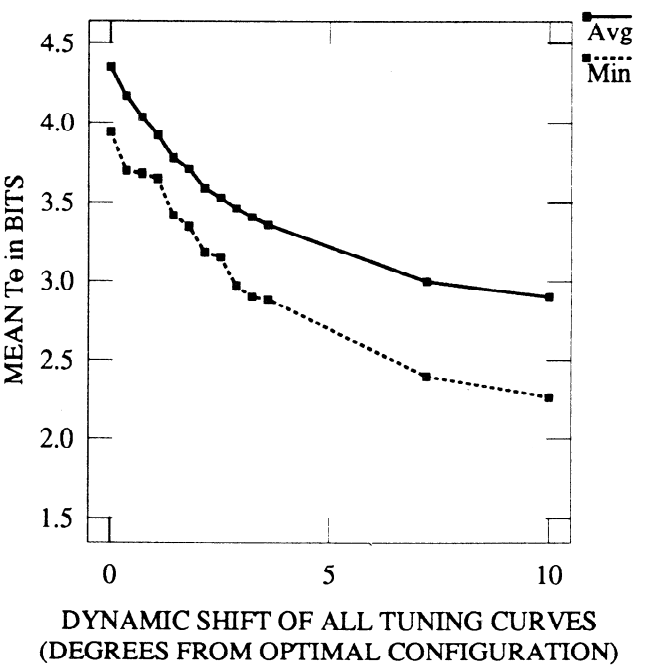

FIG. 7. Average transinformation (solid line) and minimum transinformation (dashed line) as the 4 tuning curves were randomly shifted from their optimal positions. The x-axis shows the standard deviation of this random Gaussian shift.

values of the minimum transinformation (i.e., the lower accuracy limit in the region left "undersampled" by a shifted tuning curve). The decrease of transinformation is quite severe: an entire bit (i.e., one-half of the directional accuracy) is lost for a mean shift of only 3.5°. This suggests that shifts in the directional tuning curves that could not be compensated by "recalibration" of the higher order decoder would drastically reduce system performance.

The possibility was also considered that any long-term static shift in the position of a cell's tuning curve, caused by some accident or aberration during development, could eventually be compensated to some extent by a reoptimization of the higher order neural network serving as the decoder. To determine the net extent to which the system accuracy would degrade after reoptimization to a static shift in a single tuning curve, the transinformation curves were calculated for a series of simulations in which the position of one tuning curve was shifted away from its normal position in 5° increments. A 90° shift corresponded to a superposition of the "test" tuning curve with that of its next-nearest neighbor. Here again, the calculation was done for the fixed standard velocity case. The results of these simulations are shown in Fig. 8.

It is remarkable that even large static shifts of up to 30° affected the overall mean accuracy of the system very little. The minimum accuracy for all directions decreased more rapidly but was still relatively stable to shifts of up to 10°.

Thus, if there existed enough plasticity in the decoding system to allow for such reoptimization to a new configuration, the system would be robust to static shifts in the position of the tuning curves. For comparison, the average and minimal transinformation values for a hypothetical system in which one of the cells was totally eliminated are shown as circles plotted at the 90° shift point. Even with only three functioning cells, the system would be capable of 3.8 bits of accuracy, which is more than would be obtained with four cells at most of the nonoptimal widths (refer back to Fig. 5).

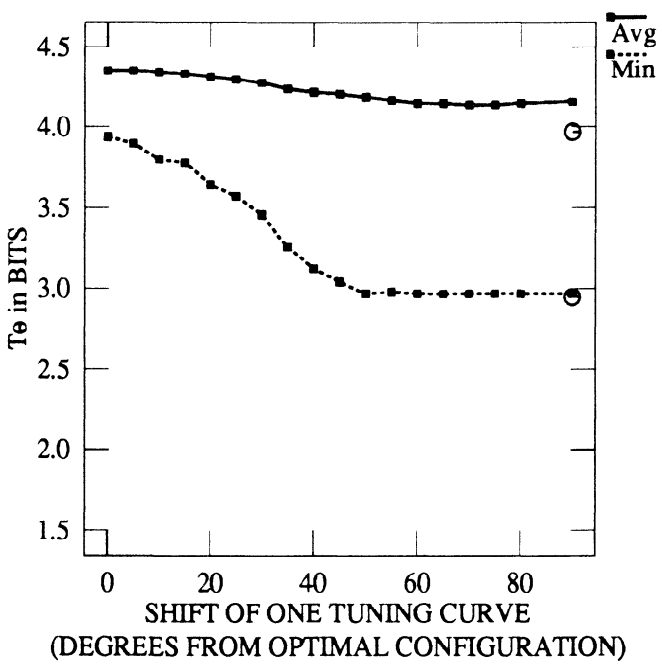

FIG. 8. Average transinformation (——) and minimum transinformation (- - -) as one tuning curve was shifted from its optimal position by values ranging from 0 to 90°. In these simulations, the system was reoptimized to the new configuration. The circles plotted at 90° correspond to the transinformation values for the situation where one of the cells was completely eliminated.

DISCUSSION

In the studies reported here, principles of information theory were used to calculate the transinformation T_θ in the stimulus-evoked responses of four wind-sensitive interneurons of the cricket cercal sensory system. As summarized earlier, T_θ can be thought of as corresponding to the limit of accuracy with which the information about stimulus direction was encoded in the mean spike rates of four interneurons of this system. An optimal decoder constructed from higher order interneurons could theoretically achieve this level of certainty in its estimate of the stimulus direction. The mean value of T_θ was calculated to lie between 3.5 and 4.2 bits, corresponding to a standard error in the system's assessment of stimulus direction lying between 7.7 and 4.7°. This level of accuracy represents a substantial degree of "sensory hyperacuity," because the minimum distinguishable stimulus separation is more than an order of magnitude finer than the widths of (and parametric spacing between) the directional tuning curves of the interneurons themselves.

We reiterate that these calculated values of T_θ depend on the set of assumptions discussed in METHODS. In choosing our set of assumptions, we were "conservative" in several respects, so as to obtain a *lower bound value* for the attainable accuracy limit. First, the values we used for the response variance in our calculations were equivalent to the upper limits of the experimental measurements. Second, all calculations were carried out disregarding the directional information that might have been present in aspects of the *patterns* of the spike trains of individual cells. Third, the *covariance* in the responses of the different cells in the ensemble was not taken into account.

Other implicit assumptions in these calculations might, however, have led to an overestimation of T_θ. First, the response of the cells was obtained by counting spikes over a period of 100 ms after stimulus onset. If the actual "integration time" were significantly less, attainable accuracy might have been less. Studies now in progress indicate, however, that use of 50-ms integration periods yield comparable estimates of T_θ. Second, it was assumed that the shape of the tuning curves was conserved for different velocities. It is possible that the curve shapes would "distort" for velocities near the upper and/or lower limits of the cells' operating ranges and possibly degrade the transmitted information. (Here the "velocity operating range" is the defined as the range for which the response varied proportional to the log of the stimulus velocity. This corresponded to velocities ranging from 0.01 to 0.2 cm/s.) Here again, sample directional plots recorded at various velocities within the ranges considered in our calculations supported the validity of our assumption (discussed in Miller et al. 1991). In any case, the manifestation of such shape distortions near the sensitivity limits would not necessarily decrease the accuracy within the central portion of the range.

After taking all these assumptions into consideration, we obtained an estimate of the directional accuracy, which depended only on the fraction of the velocity operating range over which the calculations were carried out. As we increased this fraction from a single velocity to a velocity interval including almost all of the velocity operating range, the directional accuracy of these four cells decreased from 4.2 to 3.5 bits (or 4.7 to 7.7°). It should not seem surprising that the calculated directional accuracy would depend in such a way on the assumed degree of uncertainty about the stimulus velocity. As the velocity range is increased, the variability in the response for each cell at a particular direction also increases. However, most of this variability is strongly positively correlated between the different cells, because the activities of cells with overlapping tuning curves will tend to follow the same trends in response to a change in stimulus velocity. This strong positive correlation and the fact that the overlapping tuning curves of adjoining cells have opposite slopes, result in different systematic responses to different directions over an ensemble of velocities. This difference allows the system to maintain a relatively large directional resolution for velocity intervals that span most of the velocity operating range of the cells.

Note that by choosing a particular fraction of the total operating range for the calculations, we were effectively assuming a velocity range over which these four cells were "held responsible" for encoding stimulus direction. This is actually somewhat problematic, because other cells in the cercal system might also contribute to directional information in the operating velocity range of the 10-2 and 10-3 cells; furthermore, it is possible that the 10-2 and 10-3 cells might contribute to directional information beyond their saturating velocity. Figure 3 of the preceding paper demonstrates, for example, that there is a range of velocities within which the 10-2 and 10-3 cells are below saturation and the 9-3 are above threshold. A more rigorous approach toward calculating total system accuracy would consider all cells in the cercal system and all velocities relevant to the cricket. The calculated directional accuracy for velocities below the saturating velocity of the 10's would undoubtedly be larger than our lower bound estimate presented here, which ignores other cells.

Studies now in progress will address some of these assumptions by explicitly calculating the directional transinformation on the basis of simultaneous recordings of the spike-train patterns elicited from both the 10-2 and 10-3 class and the 9-2 and 9-3 class. We have confidence in the conservatism of the calculations presented here, however, and we are impressed by the level of implied encoding accuracy.

Correspondence between transinformation and behaviorally observable acuity

In general, the transinformation value is the most rigorous measure of the coding accuracy for any given set of assumptions. Unfortunately, the correspondence between this neural encoding accuracy and the effective sensory acuity that would be observed in "behavioral" or "psychophysics" experiments is difficult to determine in insects. In a recent study, the escape reactions of *Acheta domestica* to wind stimuli were observed under open loop conditions (Stabel et al. 1985). The aim of the study was to describe the change in the cricket's walking behavior in response to speaker-generated tone pulses from different directions. The crickets were observed to turn away from the stimuli, and responses were quantified by plotting the angular velocity of the turns versus stimulus angle. It was demonstrated that the crickets would display significant turns to stimuli with angles <30° lateral from the rear. Unfortunately, the stimulus generation and delivery protocols, the peak stimulus velocity (500 mm/s), and the ambient environment were much different from those in the experiments reported here, and an appropriate estimate of resolution or accuracy from that study is not possible.

Several research groups have carried out behavioral experiments on cockroaches (Camhi and Tom 1978; Camhi et al. 1978; Westin et al. 1977), but the results of these experiments do not adequately address the questions investigated here, for several reasons. First, the cockroach cercal sensory system is substantially different from the homologous system in crickets. (For example, cockroaches have smaller cerci with many fewer filiform hairs.) Second, all behavioral experiments were done in open field conditions, with the use of stimulus velocities much larger and, presumably, much more "noisy" than the ones used in our experiments.

A determination of the extent to which crickets approach the maximum acuity judged possible from an analysis of their information encoding capabilities must await further behavioral experiments, carried out under the same circumstances and with similar protocols to those used for the studies reported here.

Mechanisms for decoding by higher order interneuronal networks

Note that we are not proposing any specific model for the mechanisms by which the information in these ensemble neural responses is decoded by higher order interneuronal or motor circuits. Our calculation of the limits of system accuracy are essentially model independent: we calculated the information encoded in the neural responses in a statistical sense, and any of a wide range of decoders could be derived that would presumably be capable of extracting some or all of this information. Adequate data are not currently available to allow a discrimination between the many conceivable models for information decoding in the cricket cercal sensory system.

Several researchers have, however, proposed candidate models for equivalent higher order decoders in the cercal sensory system of the cockroach. In two recent studies, models were proposed for simple directional decoding algorithms, on the basis of experimentally observed responses of the giant interneurons and of freely behaving cockroaches to wind stimuli (Camhi and Levy 1989; Dowd and Comer 1988). In both studies the stimulus directions were computed on the basis of simple linear summations of the action potentials of the giant interneurons. The statistical variation of the escape responses could be mimicked by adding a certain uncertainty or "noise" on top of the mean simulated responses of the interneurons. Researchers in

both studies were able to find sets of model parameters and "noise levels" that yielded close fits to the behavioral data. However, neither model used noise levels that had been set to correspond to the actual measured variances in cell responses. It must be remembered that the cell response variances are very important determinants of the information available in their mean spike rates; therefore the interpretation of those modeling studies is problematic.

In a third study (Heetderks and Batruni 1982) a specific model for an optimized linear directional discriminator was proposed, in which the principles of linear multivariate statistical analysis were used. The proposed decoding scheme did take into account the variance of the cell responses. By the use of this model, twelve directions could be discriminated with high reliability, on the basis of an analysis of the spike trains of 14 interneurons. Unfortunately, the variance values used in the model were estimated from only five repetitions of the stimulus in each direction, and the covariance in the responses of different cells was calculated on the basis of independent recordings from different animals. Both sets of variance estimates may have therefore been far from accurate.

An adequate understanding of the decoding algorithm used for this direction-sensing task will ultimately require a detailed analysis of the physiology and functional organization of wind-sensitive interneurons in the thoracic ganglia, which are the targets of the cercal interneurons considered in these studies.

Functional optimality of the cricket cercal sensory system

As demonstrated above, the transinformation calculations provided a rigorous assessment of encoding accuracy. Perhaps more significantly, this measure also allowed a quantitative assessment of the dependence of encoding accuracy on several aspects of the system's functional organization. We investigated the optimality of the width and spacing of the directional tuning curves. Our approach was to calculate the transinformation versus direction functions, similar to the function shown in Fig. 2, for a wide range of tuning-curve widths and intercurve spacings.

EFFECT OF TUNING-CURVE WIDTH ON TRANSINFORMATION. Figure 5 shows two interesting characteristics of the transinformation curves calculated for a wide range of model tuning-curve widths. The first significant result was that there exists a single global maximum in the plot of T_θ versus tuning-curve width. This theoretical global maximum occurred at a tuning-curve width of 110°. The second interesting result is that the amplitude of the fluctuation between minimal and maximal T_θ (i.e., the deviation around the mean T_θ) tends to decrease with increasing tuning-curve width. We would suspect that any such fluctuation in T_θ for different directions would essentially degrade the performance of the system as a whole. Here again, we note that there is a local minimum in this graph of T_θ variance versus tuning-curve width at a width of 110°, corresponding exactly to the global maximum for mean T_θ versus tuning-curve width. Thus a consideration of both graphs would suggest an optimal tuning-curve width of 110°. It therefore seems very significant that the experimentally measured tuning-curve widths were, in fact, ~130°.

The existence of an optimal range for tuning-curve widths in coarse-coded systems such as this one is not surprising and, in fact, was predicted in a simulation study carried out by Heiligenberg (1987). The narrowness of this range and the distinct optimum value obtained in our simulations were, however, unexpected. Heiligenberg's simulations suggested that a broad range of tuning-curve widths, between 2 and 20 times the mean interreceptor spacing, would yield equivalent accuracy. Our experimentally observed half-width value of 130° is significantly less than two times the 90° spacing intervals between adjacent tuning-curve optima, and the computed mean accuracy falls to 50% of the optimal value (i.e., mean T_θ decreases by 1 bit) for tuning curves with half-widths <90° or >170°.

In retrospect, the narrowness of this optimal range is entirely reasonable and can be explained in terms of the constraints imposed by the limited dynamic range and intrinsic response variances of the four interneurons. In the simulations using narrower tuning curves, there were ranges of stimulus orientations within which only one or none of the cells were activated to a significant extent. In such cases, large regions of "direction space" had nearly identical systematic ensemble responses, and the transinformation in those regions was consequently very low. This caused the observed degradation of the mean information as plotted in Fig. 5, as well as the severe fluctuations around the mean value. The dynamic range of cells having such narrow tuning curves would not be used to their maximum discriminatory resolution.

For simulations using broader tuning curves, the accuracy would also be expected to degrade. The broader is a cell's directional tuning curve, the lower is that cell's mean response differential at any two different stimulus directions. As the mean response differential between two different stimulus directions approaches the value of the response variance, the stimulus discrimination is severely decreased. In such a situation, the dynamic range of each cell (and therefore of the ensemble as a whole) would be extended too broadly over the stimulus space. Because Heiligenberg did not explicitly consider response variances in his simulations, the resulting compression of the optimality range was not anticipated.

The local minima observed in the dotted curve in Fig. 5 (i.e., T_θ fluctuation vs. tuning-curve width) can be understood in terms of the characteristics of the T_θ versus direction curves (i.e., similar to Fig. 2) near the critical tuning-curve widths where orders of overlap increase. As the widths were systematically increased from the lowest value of 20°, the effective dynamic range of each interneuron was spread over a wider and wider stimulus range, and the region with no response from any of the four cells became smaller and smaller. This resulted in a systematic decrease in the peak T_θ and an increase in minimum T_θ, accounting for the gradual decrease in the variation of T_θ around the mean.

However, at tuning-curve widths >62° (marked with a bold tic on the x-axis), the variation in T_θ jumped to a local maximum. Sixty-two degrees is the critical width at which each cell's tuning curve just overlaps with those of its next nearest neighbors. For tuning curves slightly broader than 62°, the uniqueness of the ensemble response for those

small regions where the tuning curves overlap creates sharp local peaks in the plots of T_θ versus direction, accounting for the jump in the variation around the mean T_θ plotted in Fig. 5.

A similar jump in T_θ variance can be seen in the dotted curve of Fig. 5 for each of the successive higher order "overlap threshold" widths of 120 and 164° (each of which are indicated with a bold tic on the x-axis.) Above 120°, each cell's tuning curve overlaps with those of all three other cells. Above 164°, each cell's tuning curve has "wrapped around" sufficiently to have two regions of overlap with its nearest neighbors. In each case, the incremental overlaps cause local jumps in T_θ values, which increase the variance of T_θ around the mean.

Thus the general trends of these plots, and the fine structure they display, can be understood at a qualitative level on the basis of a consideration of the underlying phenomena.

EFFECTS OF TUNING-CURVE SHIFTS ON TRANSINFORMATION. The overall systematic accuracy depends on a precise relative spacing of the directional tuning curves. Figure 7 demonstrates that any shift in the relative positions of the tuning curves would result in a severe degradation of accuracy, if the higher order decoding circuit could not be reoptimized to the new configuration. This seems remarkable in light of the observation that the position of the cerci are quite variable in freely behaving crickets. Presumably, a shift in cercal angle would result in a corresponding shift of the tuning curves of the left 10-2 and 10-3 cells with respect to those of the right 10-2 and 10-3 cells.

However, the directional tuning curves of these cells[4] were reported in other studies to show little or no change when the cerci were displaced to different positions within their normal range of movements (Rozhkova 1980; Rozhkova and Polishuk 1976). There must exist, therefore, some physiological "constancy mechanism," which translates the coordinate systems attached to the cerci into a coordinate system attached to the cricket's body. The nature of this mechanism is unknown but may involve some form of proprioreceptive feedback from mechanoreceptors located near the base of the cerci.

Fine tuning of the cercal sensory system during development

It is interesting to note that the intra- and interanimal variance in the shapes, widths, and spacing of the tuning curves observed experimentally were extremely small. This suggests that the establishment of these curves during development, and their maintenance during subsequent moults, must be under rigid constraints. Moreover, the excellent correspondence between the observed neural characteristics and the theoretically optimal values increases our confidence in the validity of our general approach and our basic assumptions. Certainly, evolution and natural selection have had a long time to "fine tune" this behaviorally important system toward an optimal configuration.

The physical basis for this functional fine tuning must lie in *1*) the precise anatomic characteristics of the afferent map of wind direction and of the dendritic arbors of the four interneurons within this map, *2*) the relative efficacies of the synapses from the afferents and local interneurons onto these interneurons, and *3*) the electrical properties and spike thresholds of the interneurons. A significant shift in any of these parameters away from their standard values would act to degrade system performance. It will be of substantial interest to determine the normal variance in these parameters in typical crickets, and to determine the relative contributions of genetic preprogramming and activity-dependent plasticity to this functional tuning process.

APPENDIX

Correspondence between transinformation and the standard deviation in the representation of the stimulus

In the preceding report the intrinsic accuracy with which a sensory network could encode a stimulus parameter was calculated in terms of "bits" of information. A more conventional quantity used to represent this accuracy is the standard deviation of an internal representation of the stimulus, which can then be directly associated to the "error" in the response. An average value for this standard deviation can be calculated from the local transinformation values, by assuming a specific probability distribution for the response. The calculation becomes straightforward when one assumes that *1*) the probability distribution is a Gaussian centered at the correct value of the stimulus and *2*) the value of local transinformation is constant for all values of the stimulus parameter and therefore equal to the mean transinformation.

The conditional probability of the response y given a stimulus x at 0 is then written as

$$p(y|x) = \frac{1}{K(\sigma)} e^{-(y^2/2\sigma^2)} \quad \text{where} \quad K(\sigma) = \int_{-\pi}^{+\pi} e^{-(y^2/2\sigma^2)}\, dy$$

The local transinformation is

$$T_y(x) = E\left[\log_2\left(\frac{p(y|x)}{p(y)}\right)\right] = \int_{-\pi}^{+\pi} p(y|x) \log_2\left(\frac{p(y|x)}{p(y)}\right) dy$$

As long as σ is small compared with 2π, $K(\sigma) \cong \sqrt{2\pi}\sigma$, and the local transinformation can be approximated by taking the limits of the integral to plus and minus infinity. In that case and because $p(y) = 1/2\pi$ (from our 2nd assumption)

$$T_y(x) = E\left[\log_2\left(\frac{\sqrt{2\pi}}{\sigma}\right) + \log_2\left[e^{-(y^2/2\sigma^2)}\right]\right]$$

$$T_y(x) = \log_2\left(\frac{\sqrt{2\pi}}{\sigma}\right) - \frac{1}{\ln 2} E\left(\frac{y^2}{2\sigma^2}\right)$$

$$T_y(x) = \log_2\left(\frac{\sqrt{2\pi}}{\sigma}\right) - \frac{1}{2 \ln 2}$$

Since $T_y(x)$ and σ are constant for all x (from our 2nd assumption), we can solve directly for σ in term of the mean transinformation T_y

$$\sigma = \sqrt{\frac{2\pi}{e}}\, 2^{-T_y}$$

where T_y is the transinformation in bits and σ is the standard deviation expressed in radians.

[4] The giant interneurons (GIs) referred to in Rozhkova and Polishuk (1976) and Rozhkova (1980) are identical to the neurons we call 10-2 and 10-3 in this report.

This work was supported by National Institutes of Health Grant R01 DC-00483 to J. P. Miller.

Address for reprint requests: J. P. Miller, Dept. of Molecular and Cell

Biology, Life Sciences Addition, Box 193, University of California, Berkeley, CA 94720.

Received 28 November 1990; accepted in final form 12 June 1991.

REFERENCES

CAMHI, J. M. AND LEVY, A. The code for stimulus direction in a cell assembly in the cockroach. *J. Comp. Physiol. A* 165: 83–97, 1989.

CAMHI, J. M. AND TOM, W. The escape behavior of the cockroach *Periplaneta americana.* I. Turning response to wind puffs. *J. Comp. Physiol.* 128: 193–201, 1978.

CAMHI, J. M., TOM, W., AND VOLMAN, S. The escape behavior of the cockroach *Periplaneta americana.* II. Detection of natural predators by air displacement. *J. Comp. Physiol.* 128: 203–212, 1978.

DAVIS, P. I. AND RABINOWITZ, P. Multiple integration by sampling. In: *Methods of Numerical Integration.* New York: Academic, 1984, p. 384–393.

DOWD, J. P. AND COMER, C. M. The neural basis of orienting behaviour. A computational approach to the escape turn of the cockroach. *Biol. Cybern.* 60: 37–48, 1988.

ECKHORN, R., GRÜSSER, O.-J., KRÖLLER, J., PELLNITZ, K., AND PÖPEL, B. Efficiency of different neuronal codes: information transfer calculations for three different neuronal systems. *Biol. Cybern.* 22: 49–60, 1976.

ECKHORN, R. AND PÖPEL, B. Rigorous and extended application of information theory to the afferent visual system of the cat. I. Basic concepts. *Biol. Cybern.* 16: 191–200, 1974.

ECKHORN, R. AND PÖPEL, B. Rigorous and extended application of information theory to the afferent visual system of the cat. II. Experimental results. *Biol. Cybern.* 17: 7–17, 1975.

FULLER, M. S. AND LOOFT, F. J. An information theoretic analysis of cutaneous receptor responses. *IEEE Trans. Biomed. Eng.* 31.4: 377–383, 1984.

GILLESPIE, D. T. *The Monte Carlo Method of Evaluating Integrals.* China Lake, CA: Naval Weapon Center, 1975. (NWC TP 5714).

HEETDERKS, W. J. AND BATRUNI, R. Multivariate statistical analysis of the response of the cockroach giant interneuron system to wind puffs. *Biol. Cybern.* 43: 1–11, 1982.

HEILIGENBERG, W. Central processing of sensory information in electric fish. *J. Comp. Physiol.* 161: 621–631, 1987.

HINTON, G. E., MCCLELLAND, J. L., AND RUMELHART, D. E. Distributed representations. In: *Parallel Distributed Processing,* edited by D. E. Rumelhart and J. L. McClelland. Cambridge, MA: MIT Press, 1986, vol. 1, p. 77–109.

MACKAY, D. M. AND MCCULLOCH, W. S. The limiting information capacity of a neuronal link. *Bull. Math. Biophys.* 14: 127–135, 1952.

MANSURIPUR, M. *Introduction to Information Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1987, p. 50–51.

MCCLELLAND, J. L., RUMELHART, D. E., AND HINTON, G. E. The appeal of parallel distributed processing. In: *Parallel Distributed Processing,* edited by D. E. Rumelhart and J. L. McClelland. Cambridge, MA: MIT Press, 1986, vol. 1, p. 4–44.

MILLER, J. P., JACOBS, G. A., AND THEUNISSEN, F. E. Representation of sensory information in the cricket cercal sensory system. I. Response properties of the primary interneurons. *J. Neurophysiol.* 66: 1680–1689, 1991.

OPTICAN, L. M. AND RICHMOND, B. J. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. *J. Neurophysiol.* 57: 163–178, 1987.

RAPOPORT, A. AND HORVATH, W. The theoretical channel capacity of a single neuron as determined by various coding systems. *Inf. Control* 3: 335–350, 1960.

ROZHKOVA, G. I. Comparison of the constancy mechanisms in the cercal systems of crickets (*Acheta domesticus* and *Gryllus bimaculatus*) *J. Comp. Physiol. A* 137: 287–296, 1980.

ROZHKOVA, G. I. AND POLISHUK, N. A. Constant representation of signal source location in the cercal system of the cricket. *Biofizika* 21: 725–729, 1976.

RICHMOND, B. J. AND OPTICAN, L. M. Temporal encoding of two-dimensional patterns by single units in the primate primary visual cortex. II. Information transmission. *J. Neurophysiol.* 64: 370–380, 1990.

DE RUYTER VAN STEVENINCK, R. R., AND BIALEK, W. Real time performance of a movement sensitive neuron in the blowfly visual system: coding and information transfer in short sequences. *Proc. R. Soc. Lond. B Biol. Sci.* 234: 379–403, 1988.

SHANNON, C. E. A mathematical theory of communication. *AT&T Bell Lab. Tech. J.* 27: 379–423, 1948.

STABEL, J., WENDLER, G., AND SCHARSTEIN, H. The escape reaction of *Acheta domesticus* under open-loop conditions. In: *Insect Locomotion,* edited by H. Gewecke and G. Wendler. Munich, Federal Republic of Germany: Paul Parey, 1985, p. 79–85.

WESTIN, J., LANGBERG, J. J., AND CAMHI, J. M. Responses of giant interneurons of the cockroach *Periplaneta americana* to wind puffs of different directions and velocities. *J. Comp. Physiol.* 12: 307–324, 1977.