On the interpretation of sensitivity analyses of neural responses

Romain Brette[a)]

Laboratoire de Psychologie de la Perception, Université Paris Descartes, UMR CNRS 8158, Ecole Normale Supérieure, 75005 Paris, France

Responses of auditory neurons vary with many dimensions of acoustical stimuli. As a consequence, there is a difference between sensitivity to a particular dimension (e.g., ITD or level), which is assessed when only that dimension is varied while other dimensions are fixed (yielding tuning curves), and information about that dimension, which requires that all natural variability be considered. In particular, the rate of a neuron can be very sensitive to a dimension while poorly informative about it, if it is also sensitive to other dimensions. One implication is that in a multi-dimensional world, stimulus properties such as ITD are optimally coded with heterogeneous neural populations. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3488311]

I. INTRODUCTION

Neural responses in the auditory pathway vary with ecologically relevant characteristics of auditory stimuli, such as interaural time difference (ITD) or frequency. To understand how neurons might extract this relevant information, many studies have looked at how neuronal responses change when one controlled parameter (ITD, frequency, level) is varied, which defines "tuning curves" with respect to that parameter. If the neuron's response is *sensitive* to the parameter, i.e., the tuning curve has a steep slope, then it is easy to infer the parameter value from the observation of the neural response. Therefore in one-dimensional worlds (e.g. binaural tones with fixed frequency and varying ITD), sensitivity and information are naturally related. In the following, *sensitivity* refers to how much the neural response changes with the parameter (which is related to the slope of the tuning curve), while *information* refers to how accurately one can estimate the parameter from the observed response—or, in the standard information-theoretic sense: how much the uncertainty about the unknown parameter is reduced when the neural response is observed.

With this approach, it was found that the ITD tuning curve of single neurons in the inferior colliculus of guinea pigs is sensitive enough so that the just noticeable difference (JND) in spike count is comparable to psychophysical measurements of ITD discrimination (Skottun *et al.*, 2001). It was also suggested that neurons should be more efficient in encoding ITDs when the slope of their tuning curve is maximized in the relevant ITD range (Harper and McAlpine, 2004). Similar methods led to the conclusion that auditory nerve responses contain significant temporal information about frequency and level up to 10 kHz (Heinz *et al.*, 2001a).

However, in ecologically relevant situations, stimuli are not constrained to predefined one-dimensional sets. In this case, tuning curves represent just one slice of neural responses because neural firing rates depend on many other stimulus characteristics. It follows that sensitivity to a parameter is in general very different from information about that parameter, unless neural responses are invariant to all other unknown stimulus dimensions.

I will first show that information about a parameter is overestimated when only that parameter is allowed to vary (Sec. II). In Sec. III, I will emphasize that the sensitivity of neural responses to a parameter, which is quantified by tuning curves, is equivalent to information in the special case when the parameter is the only source of stimulus variability. Tuning curves do not quantify information in more realistic settings where other stimulus dimensions can vary. Finally, I will show that when neural populations are considered, the optimal code for a sound property such as ITD is a population of neurons with heterogeneous tunings, unless stimuli are constrained to a low-dimensional world (Sec. IV).

II. INFORMATION IN SINGLE-PARAMETER EXPERIMENTS AND IN THE REAL WORLD

To see why information about a parameter is overestimated when only that parameter is allowed to vary, let us consider the fictional example of a bird which eats sliced cakes (Fig. 1). The bird must determine whether the cake is toxic, which occurs if it has more bad slices than good slices. It has a neuron which fires in proportion to the number of bad slices. To understand how this neuron encodes information about cake toxicity, the experimenter presents cakes with three slices, and the number of bad slices is varied between 0 and 3 (first row in Fig. 1). By observing the firing rate of the neuron, the experimenter can infer the number of bad slices and therefore the toxicity of the cake. In terms of information theory, the observation reduces the uncertainty (or entropy) about cake toxicity from 1 bit (good or toxic) to 0 bit (certainty). The mutual information between cake toxicity and neural response is this reduction of uncertainty, which is the maximum of 1 bit in this case.

One may ask how this information depends on the number of slices of the cake. If the same experiment is done with cakes made of just one slice, which is either good or bad (second row in Fig. 1), then the experimenter can also infer cake toxicity from the response of the neuron. Clearly, for
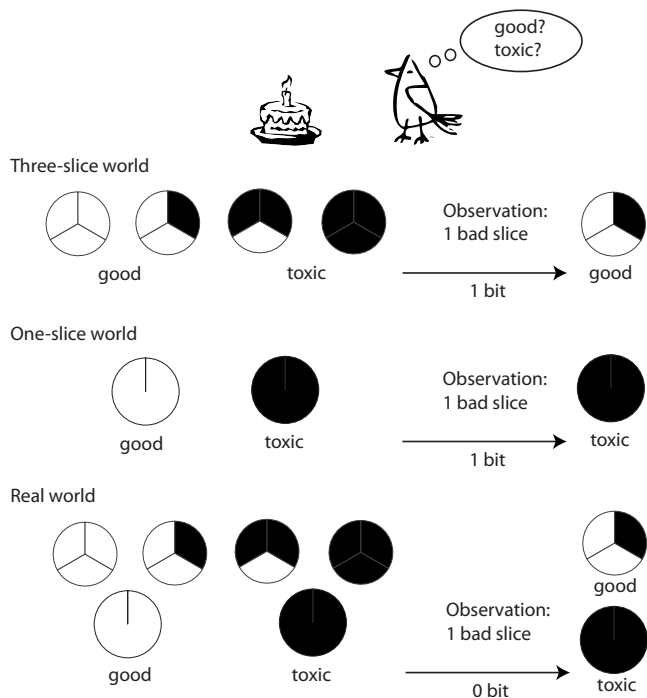
a)Electronic mail: romain.brette@ens.fr

FIG. 1. Overestimation of information in single-parameter experiments. A bird must decide is a cake is toxic, which occurs when the cake has more bad slices (black) than good ones (white). In the first two rows, cakes have a fixed total number of slices and a variable number of bad slices. Observing 1 bad slice reduces uncertainty about toxicity by 1 bit (from 2 alternatives to only 1). In the third row, the total number of bad slices is not fixed. In this case, observing 1 bad slice does not reduce uncertainty at all.

any number of slices in the cake, the information between neural response and cake toxicity is 1 bit, which is optimal. It would seem natural to conclude that this mutual information is 1 bit in general, but this is not true. In the real world, the number of slices in a cake is not fixed and the bird can meet cakes with either 1 or 3 slices (third row in Fig. 1). Suppose the neuron signals the presence of 1 bad slice. Then the cake is toxic if the cake has just one slice, but good if it has 3. Therefore the observation of the neural response does not reduce the uncertainty at all (i.e., 0 bit). Thus in this more realistic setting, the mutual information between neural response and cake toxicity is less than 1 bit. More precisely, if all cakes shown in Fig. 1 are equally likely, then the neural response is unambiguous in 4 out of 6 cases (0 or 2 bad slices), so that the mutual information is 2/3 bits.

This apparent paradox is resolved by observing that the decoding mechanism (or the *optimal estimator)* is different for 1-slice cakes and for 3-slice cakes, and there is no decoding mechanism that works for both cases: 1 bad slice means toxic cake in the former case, but good cake in the latter case. In other words, the information measured in constrained single-parameter experiments is an overestimation because it implicitly contains information in the decoding mechanism that is available only to the experimenter (the total number of slices).

This remark can be mathematically proved in a general setting as follows. Consider that the stimulus can be written as a basis stimulus S with characteristic X, e.g. $(S(t), S(t - X))$ for a binaural stimulus or $XS(t)$ for a sound with varying intensity. The mutual information between the neural response Y (e.g. spike count) and the parameter X is $I(X, Y|S)$ when S is fixed (conditional mutual information) while it is $I(X, Y)$ when S is not fixed (Cover and Thomas, 2006). In Fig. 1, S is the total number of slices, X is the cake toxicity (good or toxic) and Y is the number of bad slices. Then the following inequality holds: $I(X, Y|S) \geq I(X, Y)$. The proof is simple:

$$I(X, Y|S) = H(X|S) - H(X|Y, S) = H(X) - H(X|Y, S)$$
$$\geq H(X) - H(X|Y) = I(X, Y),$$

where the first line is the definition of mutual information (reduction of uncertainty about parameter X when the observation Y is known), the second line expresses the fact that X and S are independent and the third line comes from the fact that uncertainty decreases when S is known (in general, the inequality is strict). Thus information about parameter X is overestimated when stimuli are confined to a one-dimensional set.

III. TUNING CURVES, SENSITIVITY AND INFORMATION

A. A toy example

To understand how neurons might extract information about a continuous parameter (ITD, frequency, level), a typical approach is to measure how neuronal responses change when that parameter is varied, which defines *tuning curves* with respect to that parameter. In such a controlled situation, there is a direct link between sensitivity, as assessed for example by the slope of the tuning curve, and information. To see this, suppose that we recorded the firing rate of a neuron in response to a controlled stimulus, for different values of some parameter x (e.g. ITD). This defines a tuning curve for x, as shown in Fig. 2(a). The value of x can be inferred from the observation of the firing rate simply by inverting the tuning curve: in this case, $x = F/k$, where k is the slope. Uncertainty about x may arise from the fact that the number of spikes may vary for different observations with the same stimulus, i.e., the firing rate F is only the average observed number of spikes over repeated trials with the same stimulus, divided by duration. The uncertainty σ_F about the firing rate F [vertical arrows in Fig. 2(a)] translates to uncertainty σ_x about x [horizontal arrows in Fig. 2(a)]. If the slope of the tuning curve is smaller [Fig. 2(b)], then the uncertainty about x is larger; quantitatively, $\sigma_x = \sigma_F/k$. Thus, the information provided by the rate F about the parameter x is directly related to the sensitivity of the rate to x, i.e., to the slope of the tuning curve.

In fact, there is a precise mathematical relationship between information and sensitivity in this one-dimensional case (only x is varied). The following formula relates the mutual information between the neural response Y (spike count) and a parameter X with a measure of sensitivity called Fisher information (Brunel and Nadal, 1998):
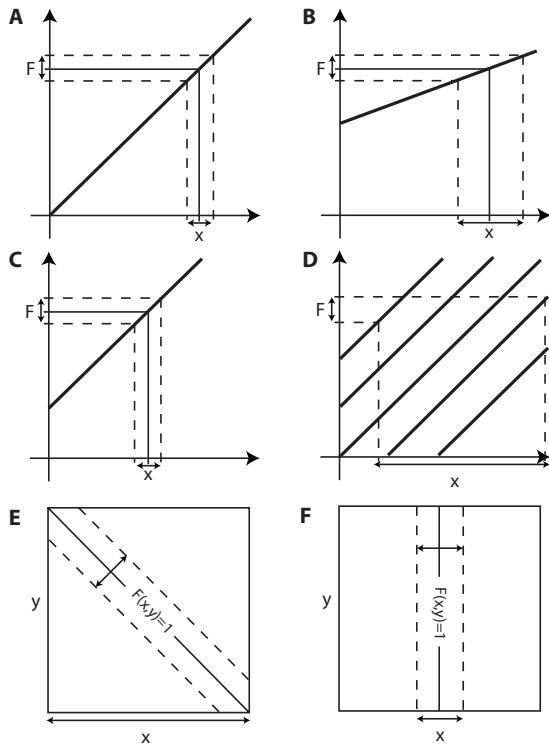
Romain Brette: Sensitivity analysis

FIG. 2. Tuning curves and information. (a) The tuning curve describes the firing rate F as a function of a parameter x. If this relationship is known, the parameter value can be inferred from the observation of F. (b) The estimation error is inversely proportional to the slope of the tuning curve. (c) For a different stimulus, a different tuning curve may be measured with respect to the same parameter x, possibly with the same slope. (d) Suppose the firing rate is in fact $F=x+y$, where y is a hidden variable corresponding to same other property of the stimulus (e.g., level; the lines are tuning curves for different values of y). Then even though x can be inferred from F for any fixed value of y using the tuning curves, it cannot be accurately estimated if y is not fixed. (e) If $F=x+y$ and x and y can both vary between 0 and 1, then the observation of F constrains x and y to a level line (solid diagonal), and uncertainty on F corresponds to uncertainty on the level line (dashed lines). The estimation of x is obtained by projecting the level line on the x axis, which gives a very inaccurate estimation in this case. (f) To obtain an accurate estimation of x, the firing rate must be insensitive to y.

$$I(X,Y) = H(X) - \frac{1}{2}\int p(x)\log\left(\frac{2\pi e}{I_{\text{Fischer}}(x)}\right)dx,$$

where $I_{\text{Fischer}}(x)$ is the Fisher information, defined from the tuning curve (not exactly the derivative of the tuning curve because the variance of spike count depends on the rate F).

Thus, the tuning curve indeed tells us how much information is contained in the neural response about the parameter, in the situation when that parameter is the only source of stimulus variability. However, as we stressed in Sec. II, this is not the same as the information about the parameter when stimuli are not artificially constrained to a restricted set.

In a more realistic situation, information corresponds to sensitivity to the parameter only if neural responses are also insensitive to other dimensions of stimuli. To see this, suppose now that in Fig. 2(a) the neuron is also sensitive to another stimulus characteristic y, and that the tuning curve was measured with $y=0$. Let us assume that the firing rate of that neuron is in fact $F(x,y)=x+y$. If we measure the tuning curve with any another value for y, then we will observe that

the neural response is also sensitive to x [Fig. 2(c)], and therefore the neural response is very informative about x. But, as for the example of the cake-eating bird (Fig. 1), this does not imply that the neural response is very informative about x when y is not constrained: in that case, the uncertainty about y translates to an uncertainty about x [Fig. 2(d)]. It appears that the uncertainty about x is made of two components: the uncertainty about the underlying firing rate F (given the observed spike count) and the uncertainty about the hidden variable y. The tuning curve only assesses the former source of uncertainty, but in Fig. 2(d) the latter is dominant.

Quantitatively, the contribution of the uncertainty about y depends on the sensitivity of the neural response to y. Figure 2(e) shows the set of parameter values (x,y) consistent with the observed firing rate $F=1$ (solid line), where x and y are allowed to vary between 0 and 1. The uncertainty about the rate F is shown by the dashed lines. It appears in this case that any value of x between 0 and 1 is consistent with the observation, and in fact the uncertainty about F does not contribute any additional uncertainty about x. On the other hand, if the neural response was not so sensitive to y, as shown in Fig. 2(f) (here $F(x,y)=x$), then the uncertainty about x would be mainly determined by the uncertainty about F.

In the case of Fig. 2(e), it is possible to estimate the standard deviation of the optimal estimator of x, if we assume that x and y are uniformly distributed between 0 and 1. It is larger than $1/\sqrt{2}$ times that of random guessing. To see this, observe that if $F<1$ (which occurs with probability $1/2$), then x can be any value between 0 and F, with equal probability. Therefore the estimation cannot be better than random guessing between 0 and F, which has a variance $\sigma_{\text{guess}}^2 F^2$ (σ_{guess} is the standard deviation of random guessing between 0 and 1). The expectation of F^2 in the lower triangle $F<1$ in Fig. 2(e) can be calculated: $E[(X+Y)^2|X+Y<1] = 1/2$. Since the case $F>1$ is symmetrical and equally likely, the variance of any estimator is at least $(1/2)\sigma_{\text{guess}}^2$, i.e., the standard deviation is at least $1/\sqrt{2}$ times that of random guessing.

B. ITD tuning curves

To be more specific, let us look at the encoding of ITDs in a single frequency channel, similarly to Harper and McAlpine (2004). The response of binaural neurons is usually described as the cross-correlation of the inputs with a preferred interaural delay d,

$$F_d = \langle S_L(t)S_R(t-d)\rangle$$

(or a variation of this formula) so that, as a function of ITD, the rate of the neuron is:

$$F_d(ITD) = \langle S(t-ITD)S(t-d)\rangle.$$

When the sound S is a pure tone with fixed frequency and level, the tuning curve is a sinusoidal function. If the wavelength of the tone is large enough compared to the range of natural ITDs, then the firing rate is most sensitive to ITD over that range when the inflexion point of the tuning curve
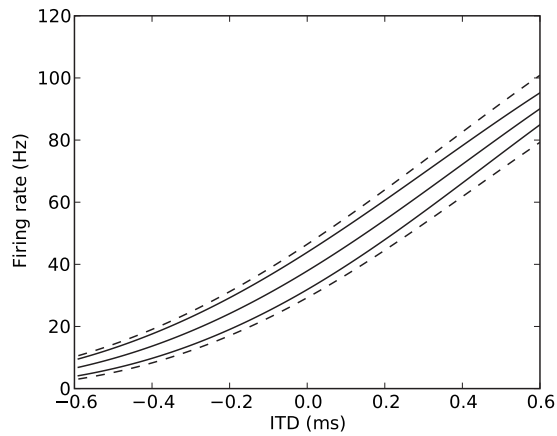
FIG. 3. ITD curve of a cross-correlator neuron with best delay at $1/(2f)$. The stimuli are 500 ms noise bursts filtered by a 4th order gammatone filter with characteristic frequency 200 Hz, half-wave rectified and power normalized (so that the cross-correlation function peaks at 125 Hz). Solid lines represent the average firing rate $\pm$ the standard deviation contributed by sound variability. Dashed lines represent the additional error contributed by spiking variability (assuming Poisson statistics) for a single neuron. When the responses of 100 neurons are pooled, the contribution of spiking variability vanishes, leaving only the error due to sound variability. The resulting ITD estimation error is 100 μs on average (about 15° azimuth for humans).

is at 0 ITD (i.e. when $d = \pi/(2f)$, where f is the frequency). In the case when the only source of variability is ITD this choice indeed maximizes the information about ITD (where uncertainty comes from spiking variability—in fact the optimally informative point is not exactly the inflexion point because information increases with rate but this is irrelevant to the present discussion). The symmetrical choice ($d = -\pi/(2f)$) is also optimal. When many neurons are considered and their firing is independent, then the distribution that maximizes information about ITD is two populations of neurons with best delays near $\pm \pi/(2f)$ (Harper and McAlpine, 2004). ITD estimation can be made arbitrarily precise by increasing the number of neurons—in fact the precision reached with just a few neurons is far better than that found in psychophysical measurements. More generally, when the parameter distribution (i.e., distribution of natural ITDs) is restricted to a limited range, it has been shown in other contexts (evolution theory) that the optimal distribution of tunings is discrete (Sasaki and Ellner, 1995), even though small perturbations can make it continuous (Haccou and Iwasa, 1998) [see also Brunel and Nadal (1998), Sec. 5.2].

However, as we previously stressed, the equivalence between sensitivity and information holds only because all other sources of variability have been neglected. Consider now that stimuli are 500 ms filtered noise bursts (through a gammatone filter). Then the firing rate of the binaural neuron depends not only on the ITD but also on the particular noise token S. This contributes an additional variability in the tuning curve, which is represented by the solid lines in Fig. 3 (to make the estimation easier, the noise bursts were normalized in power). In this case, the standard deviation of the optimal ITD estimator is about 100 μs, even if spiking variability is neglected. The key point is that the source of variability is no longer independent between neurons because it is in the signal, so that pooling neurons with the same tuning will not

reduce the estimation error, which will remain about 100 μs. The contribution of stimulus variability was recently measured in neurons in the inferior colliculus of guinea pigs (Shackleton and Palmer, 2006), with noise bursts of identical power, and was found to be about 20%. This may seem small (although a large part of it was reduced by normalizing the power), and thus neglecting stimulus variability might seem reasonable. The problem is that it does not generalize to populations of neurons, because this part of variability does not decrease when pooling neurons with the same properties. Thus, for a pool of five neurons, stimulus variability is already similar to spiking variability. For 25 neurons, the proportion is reversed. Thus, even though all neurons in the two-channel population model are optimally sensitive to ITDs, they no longer form an optimal code for ITD because the estimation error is lower bounded by the error due to stimulus variability. More generally, optimal codes for the same property (e.g. ITD) are completely different when stimuli are constrained to a one-dimensional set and in more realistic situations.

IV. OPTIMAL POPULATION CODES

A. Optimal population code for ITD: Two-channel model vs. maximal activation model

As explained in Sec. III B, when only the ITD of the sound is allowed to vary, a binaural neuron is optimally informative about ITD if its response is maximally sensitive to it, in the range of natural ITDs. When the sound frequency is small enough, two symmetrical best delays satisfy this condition, such that the tuning curve has maximal slope around 0 ITD, and the optimal code consists of two homogeneous populations of neurons with opposite best delays. I will now show that when sound variability is not neglected, a maximal activation model (e.g. Jeffress model) performs better than the two-channel model, which is therefore not optimal.

Let us consider two infinitely large populations of binaural neurons tuned at opposite best delays (e.g. $\pm \pi/(2f)$, but the precise value is irrelevant). We wish to estimate the ITD of the sound source from the activation of these neurons. Since we consider infinitely many neurons, variability in spike count vanishes and we may consider only the two firing rates x_L and x_R. Estimation in the two-channel model is generally described as a weighted average of the two symmetrical delays $\pm \Delta$:

$$ITD_{\text{estimated}} = \frac{x_L \Delta - x_R \Delta}{x_L + x_R}.$$

This is not the most accurate estimator of ITD given the rates x_L and x_R (it is biased unless the tuning curve is linear). But I will show that no single estimator can provide a perfect estimation of ITD from these two values. Suppose such an estimator exists, which maps (x_L, x_R) to ITD. Changing sound level with a constant ITD leaves the ratio x_R/x_L unchanged, which implies that the estimator is in fact a function of that ratio (it is indeed the case of the estimator above). The ITD can then be estimated from that interaural ratio only if there is a univocal mapping from ITD to the interaural ratio, i.e., if that ratio does not depend on the
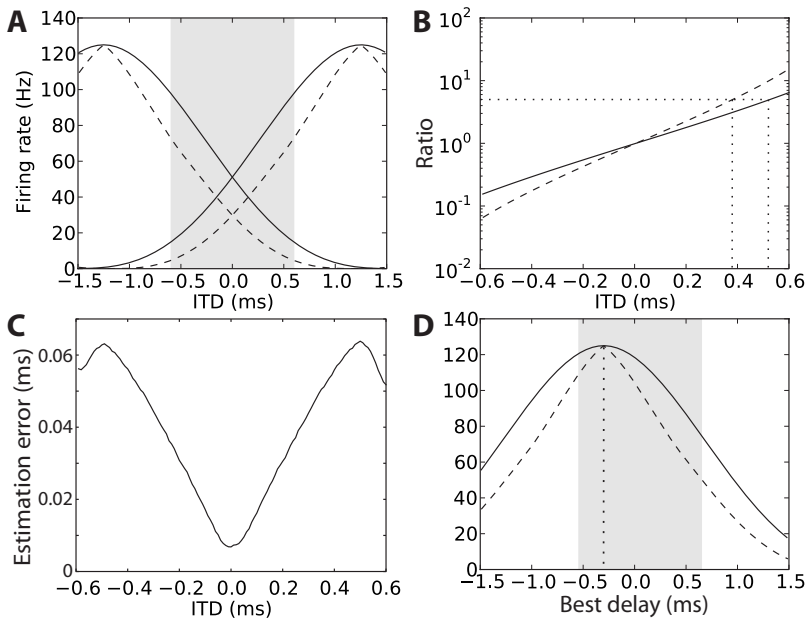
FIG. 4. ITD estimation [(a)–(c)] in the two-channel model and (d) in the maximal activation model. The sounds are gammatone filtered noise bursts, as in Fig. 3. (a) In the two-channel model, the rate of two populations of neurons with symmetrical best delays are compared. Their ratio is level-independent, but the individual tuning curves depend on the noise token (solid and dashed lines correspond to two different tokens). (b) As a result, the ratio between the two population rates depend on the noise token. Therefore a given observed ratio can correspond to different ITD estimates, depending on the noise token, which is unknown (dotted lines). (c) This contributes an estimation error when the ITD is estimated by inverting the mapping $ITD \mapsto$ mean interaural firing rate ratio. (d) On the other hand, if the population consists of neurons with heterogeneous best delays, the maximally activated neuron for a given ITD is always the same for all noise tokens (solid and dashed curves correspond to two different tokens; note that here the x axis is neuron best delay and the ITD is fixed). Thus the ITD can be accurately inferred from the observation of the firing rates.

sound but only on the ITD. While this is true if the sound is fixed (i.e., pure tone with fixed frequency), it is not if any dimension is allowed to vary. For example, consider two tones with different frequencies f_1 and f_2 within the channel bandwidth: the mapping $ITD \mapsto x_R/x_L$ is periodic with respect to the input (not preferred) frequency, and therefore a given ITD is mapped to different interaural ratios depending on input frequency (except for a finite number of intersection points). For noise bursts, the ratio x_R/x_L depends on the specific noise token except for 0 ITD where the ratio is always 1. This is shown in Figs. 4(a) and 4(b), with the same model as in Fig. 3. For two different tokens of noise (solid and dashed lines), the tuning curves for the two populations are similar but quantitatively different, even if they are normalized in level [Fig. 4(a)]. As a result, the tuning curve for the ratio of activities in the two channels also depends on the noise token [Fig. 4(b)], except for 0 ITD where it is always 1. This contributes an uncertainty about the ITD that is due to the uncertainty about the stimulus [Fig. 4(b), dotted lines], even though spiking variability is negligible, in the same way as in our example in Fig. 3. This implies that estimating the ITD from the ratio without previous knowledge of the sound will necessarily yield errors except near 0 ITD, as is shown in Fig. 4(c).

On the other hand, consider a heterogeneous population of neurons with different best delays, sampling the whole range of possible ITDs. Then when a sound is presented at a given ITD, neurons respond differently as a function of their best delay d [Fig. 4(d), solid line]. If a different token of noise is presented at the same ITD, the pattern of activation is different (dashed line), but the same neuron is maximally activated, which unambiguously indicates the ITD even though the sound is unknown (dotted vertical line). Indeed, for any sound S, the maximally activated crosscorrelator neuron is the one with best delay $d=ITD$, as shown by the following calculations:

$$\langle (S(t-ITD) - S(t-d))^2 \rangle \geq 0,$$

$$2\langle S^2 \rangle - 2\langle S(t-ITD)S(t-d) \rangle \geq 0,$$

$$F_{ITD}(ITD) \geq F_d(ITD),$$

where the first line simply expresses the positivity of squares (angle brackets mean temporal average), the second line comes from the development of the squared expression, using the fact that $\langle S(t)^2 \rangle = \langle S(t-d)^2 \rangle$ (time shift), and the third line is just the definition of the cross-correlation function $(F_d(ITD) = \langle S(t-ITD)S(t-d) \rangle)$. Thus the mapping from ITD to the place of maximal activation is one-to-one and does not depend on the sound S, which implies that the estimation error of a maximal activation model converges to zero when the number of neurons increases.

More precisely, consider N best delays regularly sampled in the physiological range of ITDs and N populations of M neurons tuned at these best delays. The ITD estimation is the best delay of the maximally activated population. Spike count variability vanishes when M is large enough and the pattern of population activation approximates the function $d \mapsto F_d(ITD)$ (neural activity vs. best delay for a given ITD) when N is large enough, so that the maximally activated population indeed signals the correct ITD of the sound with a large number of neurons.

Since we found an estimator (the maximal activation model) which performs better than the two-channel model, that model is not an optimal code for ITD when stimulus variability is not neglected. Unfortunately, this does not tell us what the optimal one is. It seems unlikely that this maximal activation model (or "peak coding") is optimal, because it discards the information in all channels but the ones around the peak. But better estimators can be devised, using for example the fact that the cross-correlation function is symmetrical with respect to the ITD, independently of the sound [e.g., with a weighted population average (Fitzpatrick et al., 1997)]. In general, it is very difficult to answer the question of optimality when there are many dimensions, but we can prove that the optimal distribution of best delays

must be heterogeneous, that is, it must include an infinite number of different best delays (or, at least as many dimensions as stimulus dimensions).

B. Optimal codes in multidimensional worlds

Consider a stimulus $S=(X,\theta)$, where θ is the parameter that we want to estimate and X is an n-dimensional random variable. For example, θ could be the ITD, and X could be the frequency and level of a tone. For white noise, X would be infinite-dimensional. Suppose a population of m neurons encodes this stimulus in their firing rate through a set of functions $f_i(X,\theta)$. For a fixed X and variable θ, $f_i(X,\theta)$ represents the tuning curve of neuron i. To simplify, we assume that the responses have the same functional form but a different preferred parameter value θ_i: $f_i(X,\theta)=f(X,\theta-\theta_i)$. For the example of ITD processing, θ_i would be the best delay of neuron i. How should we choose these preferred values θ_i in order to optimally encode θ in the firing rate of the neurons? The idea is simple: to raise any ambiguity, the mapping from (X,θ) to $(f(X,\theta-\theta_1),\ldots,f(X,\theta-\theta_m))$ should be one-to-one, which requires that at least $n+1$ of the m values $\theta_1,\ldots,\theta_m$ are different.

This reasoning still holds when spiking variability is considered. Suppose neurons fire randomly, so that the observed number of spike counts is $Y_i=G_i(f(X,\theta-\theta_i))$, where G_i are independent random Markov processes with the same properties (e.g. Poisson processes). We want to estimate θ from the m observed spike counts. From the data processing inequality, it follows that the mutual information between θ and the spike counts is smaller than the information between θ and the firing rates: $I(\theta,\{Y_i\})\le I(\theta,\{f(X,\theta-\theta_i)\})$. This upper bound depends only on the set of distinct values θ_i, rather than on the number of neurons m. If there are fewer than $n+1$ distinct values, then the function which maps (X,θ) to the rates $\{f(X,\theta-\theta_i)\}$ is not invertible, and therefore the information is strictly smaller than the maximum: $I(\theta,\{f(X,\theta-\theta_i)\})<H(\theta)$. Therefore the information between θ and the spike counts cannot approach the maximum, even if there are many neurons in the population.

On the other hand, one can easily construct an estimator which can be as precise as desired from a population of $m=p(n+1)$ neurons, with $n+1$ different preferred parameter values. The population is divided into $n+1$ subgroups of p neurons, where all neurons in each subgroup have the same preferred value θ_i. The total spike count of each subgroup is used to estimate the underlying firing rate $\{f(X,\theta-\theta_i)\}$. This estimation can be made as precise as desired by having more neurons. Then from the $n+1$ firing rates, one can deduce the value of X and θ.

We constructed this estimator for the sake of the proof, but of course it does not mean that it is the optimal one. However, it proves that an estimator with fewer than $n+1$ preferred values is not optimal, and therefore that the optimal estimator must include at least $n+1$ preferred values. Thus, an optimal population code in a multidimensional world is necessarily heterogeneous, even though only one dimension is to be estimated.

V. DISCUSSION

Because neuron responses depend on many aspects of auditory stimuli, sensitivity should not be equated with information about a particular property. The reasoning can be summarized as follows:

(1) Tuning curves measure the sensitivity of neural responses to a given parameter, which is equivalent to information about that parameter when possible stimuli only differ in that dimension.
(2) When stimuli are allowed to vary in other dimensions, information can drop drastically, if neural responses are also sensitive to them. Therefore sensitivity analyses overestimate information in realistic settings.
(3) This drop in information is due to the signal (stimulus variability) rather than the noise (spiking variability). Therefore pooling many neurons with the same tuning does not eliminate this source of uncertainty. Thus when stimuli are variable, optimally encoding a property requires a population of neurons with heterogeneous tunings.

I will now review a few results based on sensitivity analysis in the light of these remarks. It was found in computational models that the auditory nerve responses contain temporal information about tone frequency and level that is several orders of magnitude better than that implied in psychophysical measurements, when a single dimension is varied (Heinz *et al.*, 2001a). This conclusion is correct because these psychophysical measurements correspond to the situation where a listener is previously trained to a set of tones that only vary in frequency, and tested with the same set. However it does not necessarily generalize to other tasks that may be more ecologically relevant, for example determining the pitch of a periodic sound [but see Heinz *et al.* (2001b) for level variability].

It was found that just noticeable differences (JNDs) obtained from tuning curves of ITD-sensitive neurons were comparable with psychophysically measured JNDs even if only a single neuron is considered (Skottun *et al.*, 2001). The tuning curves were obtained with pure tones with varying ITD. As argued in Sec. III B, the information about ITD is much lower if stimuli are not constrained as in the experiment. Yet psychophysical measurements have shown that behavioral performance is also very good in more realistic settings where sounds are fresh noise tokens (Mossop and Culling, 1998), and the quantity derived from the tuning curve is a poor indicator of information in this case. In contrast, a previous study estimated that the responses of 40 thalamic neurons need to be pooled to reach behavioral accuracy (Fitzpatrick *et al.*, 1997). To obtain this number, the authors used a simple estimator of ITD: the center of gravity of the response profile (firing rate vs. best delay). In contrast with the two-channel model, the error of this estimator converges to zero in the limit of an unbounded range of best delays (because of the symmetry of the cross-correlation function). However it might not be optimal (Snippe, 1996), therefore the estimation is an upper bound on the optimal estimation.

For low frequency sounds in small animals, it was claimed that a two-population model optimally encodes ITD (Harper and McAlpine, 2004). Although this is correct for the considered situation, i.e., tones of fixed frequency and level and variable ITD, it does not generalize to the situation where the sound is not previously known. In that setting, the optimal distribution of best delays should in fact be heterogeneous rather than discrete (Sec. IV B)—but not necessarily uniformly distributed in the natural range of ITDs. This is not in contradiction with the observed distribution of best delays in mammals [see e.g. Yin and Chan (1990) for cats and Pecka et al. (2008), supplementary Fig. 2, for gerbils]: even though best delays are not uniformly distributed, their distribution is far from discrete (even within a given frequency channel).

The idea that neurons should be maximally sensitive to ITD ("slope coding") relies on an implicit assumption: that neurons provide independent information about the stimulus. But this is not true with realistic stimuli. This has consequences for coding (Sec. IV A), but perhaps more importantly for decoding. To see this, consider two neurons with different preferred ITDs θ_A and θ_B. Suppose they both fire at 100 Hz in response to an unknown low-frequency sound at an unknown ITD. Each of these two responses provides a poor estimation of ITD, and the pooled estimate is only slightly better. But the joint observation of the two responses provides a much better estimate: if the neurons fire at the same rate while they have different tunings, then the ITD is in fact $(1/2)(\theta_A + \theta_B)$. Thus, pooling is not an efficient way to decode heterogeneous codes. Instead, it is precisely the ITD-dependent correlations that provide information.

We have only considered two elementary aspects of realistic auditory environments: ITD and signal variability. But biological organisms must face much more complex situations, which could potentially have consequences on the neural codes they should use. For example, in low frequencies, there is not a single ITD but rather a specific frequency-dependent pattern of ITDs for any given location, because of diffraction by the head (Kuhn, 1977). Perhaps more importantly, binaural signals are always corrupted by background noise, and an "optimal code" should be robust to this noise. Taking this noise into account has consequences in terms of coding. For example, in the two-channel model, uncorrelated binaural noise cannot be distinguished from a sound source at 0 ITD (all neurons fire at the same rate), while these two stimuli can clearly be distinguished if tunings are heterogeneous. In other words, two channels can convey information about ITD and level, but at least one additional channel is required to convey information about the reliability of the estimations.

This discussion raises the question of the relevance of theoretical approaches based on optimality principles. Low frequency binaural neurons in birds have more heterogeneous tunings than those in small mammals, even when the head size and preferred frequencies are similar (Carr and Köppl, 2007; Wagner et al., 2007), which suggests that at least one of these best delay distributions is not optimal. Yet, it seems reasonable to expect that a neural structure that has evolved over millions of years should be optimal for its func-tion. If we hypothesize that these distributions are optimal, then there are two alternatives: (1) either the binaural cues are different, or (2) the two distributions are not optimal for the same task. One acoustical specificity of small mammals such as guinea pigs is that, unlike birds, their ears are very close to the ground. This implies that for most sound sources, the direct sound is shortly followed by a reflected sound, which is not very attenuated in low frequencies. If the delay is short enough, it cannot be suppressed by the auditory system [via the precedence effect (Litovsky et al., 1999)], and therefore both ITDs and ILDs would be impacted. Another difference might be that most relevant sound sources lie on the horizontal plane for small mammals, but in a three-dimensional space for birds. The other alternative is that the distributions are optimal for different tasks. For example, to steer towards a sound source in a two-dimensional world, it is sufficient to know whether the source is on the left or on the right. For this simple categorization task, the two-channel model is sound-invariant, and it is in fact optimal [see Bonnasse-Gahot and Nadal (2008)]. Thus, two symmetrical broad channels might provide a motor code [as suggested by Hancock (2007)].

Finally, should codes be optimal in biological systems? First, from the perspective of a biological organism, what should be optimal is not coding *per se*, but computation, that is, coding together with decoding. If the optimal code has to be decoded by a very complex system, then it is not very useful. More generally, if no other constraint is considered (e.g. decoding), then optimal coding is achieved when no processing is done at all (i.e., keeping the two monaural signals unprocessed). Second, various quantities could be optimized: estimation accuracy, but also number of neurons, wiring length and more generally energy consumption. If two neurons were sufficient to locate a sound source with reasonable accuracy, then an extended optimality principle, including energy consumption, would predict that binaural structures should have no more than two neurons (or perhaps a few more, if cell death is taken into account). I would then suggest that, while optimality is an interesting theoretical principle, it is most useful when it is applied to the difficult tasks that biological organisms must face, rather than to artificial situations which are not challenging for them.

Bonnasse-Gahot, L., and Nadal, J. (**2008**). "Neural coding of categories: Information efficiency and optimal population codes," J. Comput. Neurosci. **25**, 169–187.

Brunel, N., and Nadal, J. (**1998**). "Mutual information, fisher information, and population coding," Neural Comput. **10**, 1731–1757.

Carr, C., and Köppl, C. (**2007**). "Strategies for encoding ITD in the chicken nucleus laminaris," *Hearing from Sensory Processing to Perception* (Springer, Berlin), pp. 417–424.

Cover, T. M., and Thomas, J. A. (**2006**). *Elements of Information Theory*, 2nd ed. (Wiley, New York), pp. 1–776.

Fitzpatrick, D. C., Batra, R., Stanford, T. R., and Kuwada, S. (**1997**). "A neuronal population code for sound localization," Nature (London) **388**,

871–874.

Haccou, P., and Iwasa, Y. (**1998**). "Robustness of optimal mixed strategies," J. Math. Biol. **36**, 485–496.

Hancock, K. E. (**2007**). "A physiologically-based population rate code for interaural time differences (ITDs) predicts bandwidth-dependent lateralization," *Hearing from Sensory Processing to Perception* (Springer, Berlin), pp. 389–397.

Harper, N. S., and McAlpine, D. (**2004**). "Optimal neural population coding of an auditory spatial cue," Nature (London) **430**, 682–686.

Heinz, M. G., Colburn, H. S., and Carney, L. H. (**2001a**). "Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve," Neural Comput. **13**, 2273–2316.

Heinz, M. G., Colburn, H. S., and Carney, L. H. (**2001b**). "Evaluating auditory performance limits: II. One-parameter discrimination with random-level variation," Neural Comput. **13**, 2317–2338.

Kuhn, G. F. (**1977**). "Model for the interaural time differences in the azimuthal plane," J. Acoust. Soc. Am. **62**, 157–167.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (**1999**). "The precedence effect," J. Acoust. Soc. Am. **106**, 1633–1654.

Mossop, J. E., and Culling, J. F. (**1998**). "Lateralization of large interaural delays," J. Acoust. Soc. Am. **104**, 1574–1579.

Pecka, M., Brand, A., Behrend, O., and Grothe, B. (**2008**). "Interaural time difference processing in the mammalian medial superior olive: The role of glycinergic inhibition," J. Neurosci. **28**, 6914–6925.

Sasaki, A., and Ellner, S. (**1995**). "The evolutionarily stable phenotype distribution in a random environment," Evolution (Lawrence, Kans.) **49**, 337–350.

Shackleton, T. M., and Palmer, A. R. (**2006**). "Contributions of intrinsic neural and stimulus variance to binaural sensitivity," J. Assoc. Res. Otolaryngol. **7**, 425–442.

Skottun, B. C., Shackleton, T. M., Arnott, R. H., and Palmer, A. R. (**2001**). "The ability of inferior colliculus neurons to signal differences in interaural delay," Proc. Natl. Acad. Sci. U.S.A. **98**, 14050–14054.

Snippe, H. P. (**1996**). "Parameter extraction from population codes: A critical assessment," Neural Comput. **8**, 511–529.

Wagner, H., Asadollahi, A., Bremen, P., Endler, F., Vonderschen, K., and von Campen-hausen, M. (**2007**). "Distribution of interaural time difference in the barn owl's inferior colliculus in the low- and high-frequency ranges," J. Neurosci. **27**, 4191–4200.

Yin, T. C., and Chan, J. C. (**1990**). "Interaural time sensitivity in medial superior olive of cat," J. Neurophysiol. **64**, 465–488.