

A quantitative analysis of information about past and present stimuli encoded by spikes of A1 neurons

Stefan Klampfl, Stephen V. David, Pingbo Yin, Shihab A. Shamma and Wolfgang Maass
J Neurophysiol 108:1366-1380, 2012. First published 13 June 2012;
doi: 10.1152/jn.00935.2011

You might find this additional info useful...

This article cites 45 articles, 21 of which you can access for free at:
<http://jn.physiology.org/content/108/5/1366.full#ref-list-1>

Updated information and services including high resolution figures, can be found at:
<http://jn.physiology.org/content/108/5/1366.full>

Additional material and information about *Journal of Neurophysiology* can be found at:
<http://www.the-aps.org/publications/jn>

This information is current as of September 11, 2012.

A quantitative analysis of information about past and present stimuli encoded by spikes of A1 neurons

Stefan Klampfl,^{1,4} Stephen V. David,^{2,3} Pingbo Yin,² Shihab A. Shamma,² and Wolfgang Maass¹

¹Institute for Theoretical Computer Science, Graz University of Technology, Graz, Austria; ²Institute for Systems Research, University of Maryland, College Park, Maryland; ³Oregon Hearing Research Center, Oregon Health and Science University, Portland, Oregon; and ⁴Know-Center GmbH, Graz, Austria

Submitted 17 October 2011; accepted in final form 8 June 2012

Klampfl S, David SV, Yin P, Shamma SA, Maass W. A quantitative analysis of information about past and present stimuli encoded by spikes of A1 neurons. *J Neurophysiol* 108: 1366–1380, 2012. First published June 13, 2012; doi:10.1152/jn.00935.2011.—To process the rich temporal structure of their acoustic environment, organisms have to integrate information over time into an appropriate neural response. Previous studies have addressed the modulation of responses of auditory neurons to a current sound in dependence of the immediate stimulation history, but a quantitative analysis of this important computational process has been missing. In this study, we analyzed temporal integration of information in the spike output of 122 single neurons in primary auditory cortex (A1) of four awake ferrets in response to random tone sequences. We quantified the information contained in the responses about both current and preceding sounds in two ways: by estimating directly the mutual information between stimulus and response, and by training linear classifiers to decode information about the stimulus from the neural response. We found that 1) many neurons conveyed a significant amount of information not only about the current tone but also simultaneously about the previous tone, 2) the neural response to tone sequences was a nonlinear combination of responses to the tones in isolation, and 3) nevertheless, much of the information about current and previous tones could be extracted by linear decoders. Furthermore, our analysis of these experimental data shows that methods from information theory and the application of standard machine learning methods for extracting specific information yield quite similar results.

auditory cortex; information theory; temporal integration; data analysis; awake ferrets

A NUMBER OF STUDIES in the auditory cortex have demonstrated that neurons respond differently to a certain sound, depending on which sensory events preceded this sound by several 100 ms. For example, it has been shown in monkeys (Bartlett and Wang 2005; Brosch et al. 1999; Malone et al. 2002; Yin et al. 2008), cats (Brosch and Schreiner 2000; McKenna et al. 1989), and rats (Kilgard and Merzenich 1999) that the responses to a given tone can change if another tone has been played immediately before. For different configurations of frequency, intensity, and temporal separation, the response to this second tone can be facilitated or suppressed. In fact, the effect of a preceding tone can have many different time scales, ranging from the level of the refractory period of a cortical neuron to short-term forward masking and longer term suppression or facilitation. Furthermore, it has been shown that neurons in

songbirds respond preferentially to the bird's own song but weakly to a different sequence of song syllables (Doupe 1997; Lewicki and Arthur 1996; Margoliash and Fortune 1992). However, a quantitative analysis of these effects has been missing.

These effects could be viewed as mere imperfections of an ideal neural code, where neurons try to represent the impinging stimulus in an invariant manner but where this ideal representation is perturbed by the stimulus history in very complicated ways. A more positive functional role of these effects has been proposed by the “liquid computing model” (Buonomano and Maass 2009; Maass et al. 2002). This model proposes that, although these perturbations of neural codes through stimulus history are very complicated and vary from neuron to neuron, they may together play an important computational role: they enable neural readouts from the neural circuits involved to decode the stimulus history in addition to the current stimulus. This additional information obviously would be useful to downstream networks for higher level sensory perception and decision making. According to the underlying theory, diversity and nonlinearity of neural responses are necessary and sufficient for the enhanced decoding ability (Maass et al. 2002). This model has attracted substantial interest in the computational as well as the experimental neuroscience community (e.g., Bernacchia et al. 2011; Nikolic et al. 2009; Sussillo and Abbott 2009). However, experimental evidence in sensory systems so far for this computing model has only been reported for anesthetized cats (Nikolic et al. 2009), and it is not exactly clear how anesthesia affects sensory information processing. One goal of this study was to test the predictions of this model for neural responses in the primary auditory cortex of awake animals. We evaluated how well a linear readout (perceptron) could decode the identity of the preceding tone from the neural response to the current tone. We also tested another prediction of the liquid computing model: that the diversity of neural responses also supports nonlinear computations on spatial or temporal components of a complex stimulus. More precisely, the model predicts that the results of such nonlinear computations can be approximated by suitably trained linear readouts.

It is also of interest to understand how much information is altogether contained in neural responses, irrespective of how much a linear readout can extract. The most principled and rigorous way to analyze the information contained in neural responses is to directly use methods from information theory (Cover and Thomas 1991). However, it is well-known that the direct estimation of the mutual information between stimulus and response suffers from a severe bias problem due to the

Address for reprint requests and other correspondence: S. Klampfl, Institute for Theoretical Computer Science, Graz Univ. of Technology, Inffeldgasse 16b/I, A-8010 Graz, Austria (e-mail: klampfl@igi.tugraz.at).

limited number of available trials (Miller 1955; Panzeri and Treves 1996). The method proposed in by Panzeri et al. (2007) also produces reliable information estimates for quite limited sample sizes and has been successfully validated in a variety of analysis studies, for example, in the rat barrel cortex (Arabzadeh et al. 2004, 2006) and in the monkey visual cortex (Montemurro et al. 2008). We used this method to quantify the temporal integration capability of single A1 neurons by estimating the amount of information they convey at a particular point in time simultaneously about currently and previously played tones. We also compared this estimated amount of total information with the amount that can be extracted by linear readouts.

Altogether, our findings show that the responses of A1 neurons are consistent with the predictions of the liquid computing model. Our results provide the first test of these predictions in the auditory system and the first test in the sensory system of awake animals.

MATERIALS AND METHODS

Experimental Procedures

Auditory responses were recorded extracellularly from single neurons in primary auditory cortex (A1) of four awake, passively listening ferrets. All experimental procedures conformed to standards of the National Institutes of Health and were approved by the University of Maryland Animal Care and Use Committee. Details of the surgical and neurophysiology procedures are described elsewhere (David et al. 2009) and briefly summarized here. Animals were implanted with a stainless steel head post to permit stable recordings. Single-unit activity was recorded from four independently movable high-impedance (2–4 M Ω) electrodes in a sound-attenuating chamber. Spiking events were extracted from the continuous signal using principal components analysis and *k*-means clustering. In total, 122 single A1 neurons were isolated from 23 multichannel recordings.

Stimuli were presented from digital recordings using custom Matlab software (The MathWorks, Natick, MA). Digitally synthesized sounds were transformed to analog (National Instruments, Austin, TX), equalized to achieve flat gain (Rane, Mukilteo, WA), amplified to a calibrated level (Rane, Mukilteo, WA), and attenuated (Hewlett Packard, Palo Alto, CA) to the desired sound level. These signals were presented through an earphone (Etymotics, Elk Grove Village, IL) contralateral to the recording site. Before each experiment, the equalizer was calibrated according to the acoustical properties of the earphone insertion. In each experiment, stimuli were presented at a fixed sound level (65 dB SPL).

Tone Sequence Stimuli

During each recording, random tone sequences were presented as stimuli to the passively listening animal. Individual trials consisted of 100 tones, each of which had a duration of 150 ms; thus the duration of each sequence was 15 s. Table 1 gives detailed information about each recording: the number of recorded neurons, the number of tone sequences presented to the animal, and the different tone frequencies used. The frequency step between two consecutive tones in the sequences was either half an octave up or down. The direction of tone change after each tone was randomly chosen, except for the maximum (minimum) frequency within a recording, where the next lower (higher) tone followed with 100% probability. The first tone on each trial was selected from a uniform distribution, which ensured that each frequency appeared approximately the same number of times within the tone sequences of one recording (except for the 2 extreme frequencies, which appeared about one-half as many times). Furthermore, each frequency appeared about equally often before the next higher and the next lower possible frequency. Tone sequences have

Table 1. Information about the 23 recordings made from the 4 ferrets

Animal	Recording	No. of Neurons	No. of Sequences	Start Frequency, Hz	No. of Frequencies
<i>Ferret 1</i>	ele093b05	4	54	1,000	9
<i>Ferret 2</i>	per001c06	3	35	2,125	7
	per002d03	6	35	1,775	7
	per003b08	3	35	1,775	7
	per006a08	8	35	2,510	7
	per007b10	4	35	1,202	7
	per011b05	2	35	2,943	11
	per011c05	1	35	2,943	11
	per018c10	7	35	2,050	7
	per019a05	9	35	2,298	7
	per026a09	6	42	1,727	7
	per027a07	10	42	354	7
	per027b06	10	42	354	7
	per028a09	6	42	1,945	7
	per031a08	6	42	1,768	7
per031a09	6	54	1,250	9	
per034a09	7	54	1,250	9	
<i>Ferret 3</i>	sas009b13	5	54	800	9
	sas028b07	4	54	700	9
	sas029a07	2	54	250	9
	sas031a07	6	54	625	9
<i>Ferret 4</i>	sag002e19	4	54	350	9
	sag005a03	3	54	725	9

No. of neurons indicates the number of simultaneously recorded neurons; no. of sequences denotes the number of tone sequences (each with 100 tones) used as stimuli. For each recording, the sequences consist of a series (no. of frequencies) of frequency values increasing from a starting value (start frequency). The quotient between 2 neighboring frequency values was $\sqrt{2} = 1.41$.

been used previously in several studies of A1, although most focused on either two frequencies (Ulanovsky et al. 2003, 2004) or one varying frequency paired with a fixed base frequency (Brosch et al. 1999; Brosch and Schreiner 2000). These studies reported response characteristics that were similar to those we found (see Fig. 1).

Methods of Information Estimation

We employed several methods of quantifying the information contained in the responses about current and preceding tones. First, we estimated directly the mutual information between the response of single neurons and the current, previous, and preceding tones. We used a fine temporal discretization of the responses to rigorously capture information encoded in both the rate and timing of spikes. Second, we trained linear classifiers on the activity of the ensemble of all simultaneously recorded neurons. This measures the amount of information that can be decoded by a hypothetical neuron that receives input from these recorded neurons. Finally, we recalculated the mutual information for the firing rates of the ensemble of simultaneously recorded neurons to allow a direct comparison with the amount of information that can be extracted by linear decoders.

Direct estimation of mutual information. Information contained in the neural response can be analyzed by estimating directly the mutual information between stimulus and response as a general measure from information theory (Cover and Thomas 1991; Shannon 1948). The mutual information between stimuli *S* (in our case, tone frequencies or direction of tone changes) and evoked responses *R* is given by

$$I(S;R) = H(R) - H(R|S) = \sum_{r,s} P(s)P(r|s) \log \frac{P(r|s)}{P(r)}. \quad (1)$$

where *S* and *R* are random variables characterized by probability distributions *P*(*s*) and *P*(*r*), respectively. The conditional entropy

(or noise entropy) $H(R|S)$ describes the variability of the responses for a fixed stimulus s [expressed by the conditional distribution $P(r|s)$]. If this variability is small compared with the overall variability of responses $H(R)$, the response conveys a large amount of information about the stimulus. Mutual information thus quantifies the reduction of uncertainty about the stimulus that can be gained from observation of a single response trial (Rieke et al. 1997).

Calculation of mutual information requires accurate estimation of the response probabilities $P(r)$ and $P(r|s)$ from the finite experimental data, which suffers from a sampling problem: estimating the probability distribution from a finite number of observations leads to a systematic underestimation of the entropy of this distribution (Miller 1955) and consequently to an upward bias of mutual information (Panzeri and Treves 1996). To overcome this bias problem, we used a shuffling-based estimator (Panzeri et al. 2007), which eliminates the bias by subtracting the noise entropy of a randomly shuffled data set. Furthermore, an additional bias correction technique was applied that uses quadratic extrapolation of the individual entropy measures to infinite sample sizes (Panzeri et al. 2007). However, the bias of the information estimation cannot be fully eliminated for arbitrarily small sample sizes.

To evaluate the information contained in the responses about a particular feature of the stimulus sequences (e.g., the current tone, or the direction of previous tone change), we viewed all available instances of this feature as our set of stimuli, and the spike trains in response to all occurrences of one particular instance within the tone sequences of one recording as our set of responses for this stimulus. For example, to calculate the mutual information $I(T;Y)$ between the response Y and the current tone T , the set of stimuli consisted of all available tone frequencies used in one recording, and each of these stimuli was associated with a set that is composed of the neural responses during all occurrences of the respective frequency.

To measure the information contained in the responses about the previously played tone, we took into account the special temporal structure of the stimulus sequences. Because each tone is followed by either the tone with the next higher or the next lower frequency, there is already substantial mutual information between two consecutive tones, so a high value of the mutual information between the response and the previously played tone might actually reflect information about the current tone because it is not independently chosen from the previous tone. We therefore calculated the mutual information about the direction of the tone step given by the previous and current tone.

The total information that is contained in the response Y at time t about the tone pair (T_1, T_2) can be written as follows, according to the chain rule of information theory (Cover and Thomas 1991):

$$I(\Delta, T_2; Y_t) = I(T_2; Y_t) + I(\Delta; Y_t | T_2), \quad (2)$$

where Δ is a binary variable indicating whether the tone pair (T_1, T_2) is an up or down step in frequency. Note that the tone pair (T_1, T_2) is completely determined by the tuple (Δ, T_2) . The first term, $I(T_2; Y_t)$, is then simply the information about the current tone (during T_2), and the second term, $I(\Delta; Y_t | T_2)$, measures for a given value of the current tone T_2 the additional information that is conveyed about the direction of the tone change from T_1 to T_2 .

This approach can be extended to investigate whether the neural response also contains information about tone changes farther back in the sequence. Note that a sequence of n successive tones (T_1, \dots, T_n) is completely determined by the value of the last tone, T_n , and a sequence of $n - 1$ binary variables $(\Delta_1, \dots, \Delta_{n-1})$ indicating the directions (up or down) of the $n - 1$ tone steps of the sequence (T_1, \dots, T_n) . We can write analogously to Eq. 2,

$$\begin{aligned} I(\Delta_1, \dots, \Delta_{n-1}, T_n; Y_t) &= I(T_n; Y_t) + I(\Delta_{n-1}; Y_t | T_n) \\ &\quad + I(\Delta_{n-2}; Y_t | T_n, \Delta_{n-1}) + \dots \\ &\quad + I(\Delta_i; Y_t | T_n, \Delta_{n-1}, \dots, \Delta_2) \\ &= I(T_n; Y_t) + \sum_{i=1}^{n-1} I(\Delta_i; Y_t | T_n, \Delta_{n-1}, \dots, \Delta_{i+1}). \end{aligned} \quad (3)$$

That is, the information can be written as a sum of contributions of individual tone changes, conditioned on later tones. Note that for $n = 2$, Eq. 2 follows as a special case.

Reliable information estimates require a suitable discretization of the response space. We discretized each spike train of a neuron as a binary word with bin size of 5 ms, i.e., a value of 1 in one bin indicates the presence of at least one spike in the corresponding 5-ms period. At time t following stimulus onset, we calculated the mutual information between the stimulus and the response of a single neuron during a time window of 20 ms (4 time bins) preceding t . Because of the overlapping time windows and the temporal integration capability of the auditory system information, values calculated at different values of t are not independent of each other. We chose a time window of 20 ms because this is in the order of typical durations of excitatory postsynaptic potentials (EPSPs) and reflects the relevant time period during which cortical neurons integrate incoming information arriving at their synapses. We further discretized this 20-ms time window into binary bins to capture information encoded in both the overall rate and the exact timing of spikes. The bin size of 5 ms is in the order of typical refractory mechanisms. Different time windows and bin sizes might lead to different estimated information values.

We also used a different discretization scheme when we compared mutual information with classification performance. We used the spike count, limited to values from 0 to 4, in a time window of 20 ms preceding a particular time point t to represent the response of a neuron at time t . In contrast to the binary code described earlier, where we calculated the information conveyed by individual neurons, we used the vector composed of the spike counts of all simultaneously recorded neurons as our response. The same vector can be used as a classifier input, which makes the comparison between mutual information and classifier performance possible.

With this discretization of the response into vectors of spike counts, the size of the response space becomes potentially 5^n , where n is the number of simultaneously recorded neurons. This increases the risk of bad information estimates caused by undersampling of the response space. To determine the reliability of the measured information values, we evaluated the mutual information on a random subset of one-half of all available trials. If the mutual information value estimated on this truncated data set changed by $<10\%$ of its original value, a reliable estimate was reported. As software, we used the Python implementation described by Ince et al. (2009).

Analysis of information using linear classifiers. Information in the neural response about the stimulus can also be extracted by classifiers, with spike trains of simultaneously recorded neurons as input patterns and the stimulus identity as the target label (see e.g., Nikolic et al. 2009, for an application of this method). We used (binary) linear classifiers, more precisely, Support Vector Machines (SVMs) with a linear kernel (see e.g., Ben-Hur et al. 2008; Bishop 2006; Schölkopf and Smola 2002), to decode one particular bit of information about the stimulus from the responses. These linear classifiers approximate the computational operation of a hypothetical neuron that reads out information from all simultaneously recorded neurons. The performance of a linear classifier can be viewed as a lower bound on the information contained in the responses about this particular bit, the amount of information that is accessible to a linear readout neuron.

All response spike trains were low-pass filtered by an exponential filter with time constant $\tau = 20$ ms, which mimicked the time course of typical EPSPs. At particular points in time, the values of these analog traces of multiple simultaneously recorded neurons of one particular recording form multidimensional input vectors that serve as

input patterns to the classifier. Performance of the classifiers was always evaluated with 10-fold cross-validation. The parameter C of the linear SVM, which determines the trade-off between minimizing training errors and generalization, was chosen to be 100; however, we found that the achieved performance was not very sensitive to this parameter. Furthermore, we tried SVMs with different kernels (polynomial kernels of different degrees and radial basis function kernels with different widths), but no significant performance increase was detected. This means that in this case, linear classifiers performed as well as most nonlinear classifiers, and no significant additional information can be extracted using these nonlinear classifiers.

With this method we also tested whether the neural response provides a nonlinear combination of input components. This is the case if the linear classifier is able to reproduce the exclusive-or (XOR) computation of two bits in the input sequence. This computation yields result "1" if these two bits are different and "0" if they are the same. In this XOR experiment (see Fig. 6), we measured the performance as the point-biserial correlation coefficient between the binary target variable and the continuous linear combination of the low-pass-filtered input spike trains learned by the classifier (i.e., the output of the classifier before the threshold operation). This is necessary to ensure that any significant classification performance on this nonlinear XOR combination really can be attributed to a nonlinearity implicitly provided by the neural responses. It can be shown that any such correlation coefficient significantly greater than zero indicates nonlinear transformations in the neural processes itself (see Nikolic et al. 2009 for a formal proof).

When we compared mutual information with classification performance, we did not use the low-pass-filtered spike trains of all simultaneously recorded neurons as input to the classifier; rather, we used the vector composed of the spike counts of these neurons during the last 20 ms (the same value as the time constant τ). This discretization allowed the application of both a mutual information estimator and a classifier to the same data. To convert the classification performance (%correct) into an information value (bit), we calculated directly the mutual information (1) between the stimulus and the classifier prediction. As software for solving the SVM optimization problem, we used the LIBSVM package (Chang and Lin 2001) included in the PyML toolbox for Python (<http://pyml.sourceforge.net/>).

Data Preprocessing and Statistical Analysis

To calculate the peristimulus time histograms (PSTH) for a neuron in response to a given frequency, we collected the spike times during all occurrences of this frequency during 150-ms tone periods. The bin size of the PSTHs was chosen to be 1 ms; for display purposes only, the PSTHs were smoothed with a 10-ms Hamming window. To measure correlations between two variables, we used the standard Pearson correlation coefficient, which yields an r and a P value. The r values denote the correlation coefficient itself, and P values are the probability that these correlations are produced by random chance. Thus a low P value indicates a significant correlation. If r values are reported without a P value, $P < 0.0005$.

To assess the significance of the obtained mutual information values, we performed a label-shuffling test: for each neuron, we generated 1,000 different shuffled data sets by randomizing the stimulus identity (label) for each response. This random assignment of labels to whole spike trains of 150-ms duration ensured that all higher order information contained within these spike trains was maintained. Information was then calculated from these shuffled instances, yielding a distribution of shuffled information values. For a particular point in time, a significant information value ($P < 0.05$) was reported if the "true" information value, which was estimated from the correct stimulus labels, was larger than the mean plus two times the standard deviation of this distribution of shuffled information values. For the information about previously played tones, we compared the average information value across different tones (or tone sequences, see Eqs.

2 and 3) with the distribution of the corresponding averages of the shuffled information values. The significance of the whole response of a given neuron was assessed in the following way. During the 150-ms tone period, we selected 7 time points [20 ms, 40 ms, . . . , 140 ms] and checked whether at least at one of these time points the measured information was significant according to the method described above. If this was the case, this neuron was scored as conveying significant information. The 20-ms interval was chosen as the size of the overlapping time windows to account for the fact that temporally adjacent information values are highly dependent. Furthermore, we employed a standard Bonferroni correction to the alpha value to correct the bias from making seven comparisons.

For assessing the significance of performance values of linear classifiers, we employed the same shuffling procedure. At a particular point in time, a significant classification performance ($P < 0.05$) was reported if it was larger than the mean plus two times the standard deviation of this distribution of shuffled performance values. The gray-shaded regions in Figs. 5A and 6 denote the distribution of shuffled performance values, indicating nonsignificant deviations from chance level.

RESULTS

Neural responses to Tone Sequences in Primary Auditory Cortex

To measure the temporal integration of sensory information in neural responses of A1, we recorded the activity of 122 neurons isolated from 23 multichannel recordings in 4 awake, passively listening ferrets. The stimuli were sequences of tones of 150-ms duration. The frequency step between two consecutive tones was always randomly half an octave up or down, and the sequences were designed to present tones at each frequency approximately the same number of times.

Figure 1A shows a snapshot of such a stimulus sequence used in one recording, overlaid with the associated responses of four simultaneously recorded neurons. The PSTH response for a given frequency was computed by averaging the spike trains following the onset of all tones at that frequency (see examples in Fig. 1B). Individual neurons responded to different tone frequencies in various ways. For instance, *neuron 4* had a very sparse response compared with the other units. Other neurons, such as *neuron 3*, typically responded with a strong transient burst to a tone onset or change, whereas the responses of still other neurons, such as those of *neuron 2*, were sustained across the tone duration. Furthermore, especially neurons with a strong transient response tended to be more sharply tuned to specific frequencies or frequency ranges (see also Fig. 2B).

As had been shown previously (Brosch and Schreiner 2000; Ulanovsky et al. 2004), the responses of individual neurons to individual frequencies differed, depending on the tone frequency that had been played immediately before. Figure 2A shows PSTHs of four sample neurons plotted as in Fig. 1B but conditioned on the direction of the preceding tone step. For some neurons and frequencies, there was a substantial difference in the firing rate in response to the same stimulus frequency. This difference was particularly large for neurons responding with a strong transient burst. These context-dependent responses give rise to conditional tuning curves, which plot mean firing rate as a function of tone frequency separately for both up and down steps in frequency from the preceding to the current tone (Fig. 2B). A downstream decoder could extract information about preceding tones from the response of neu-

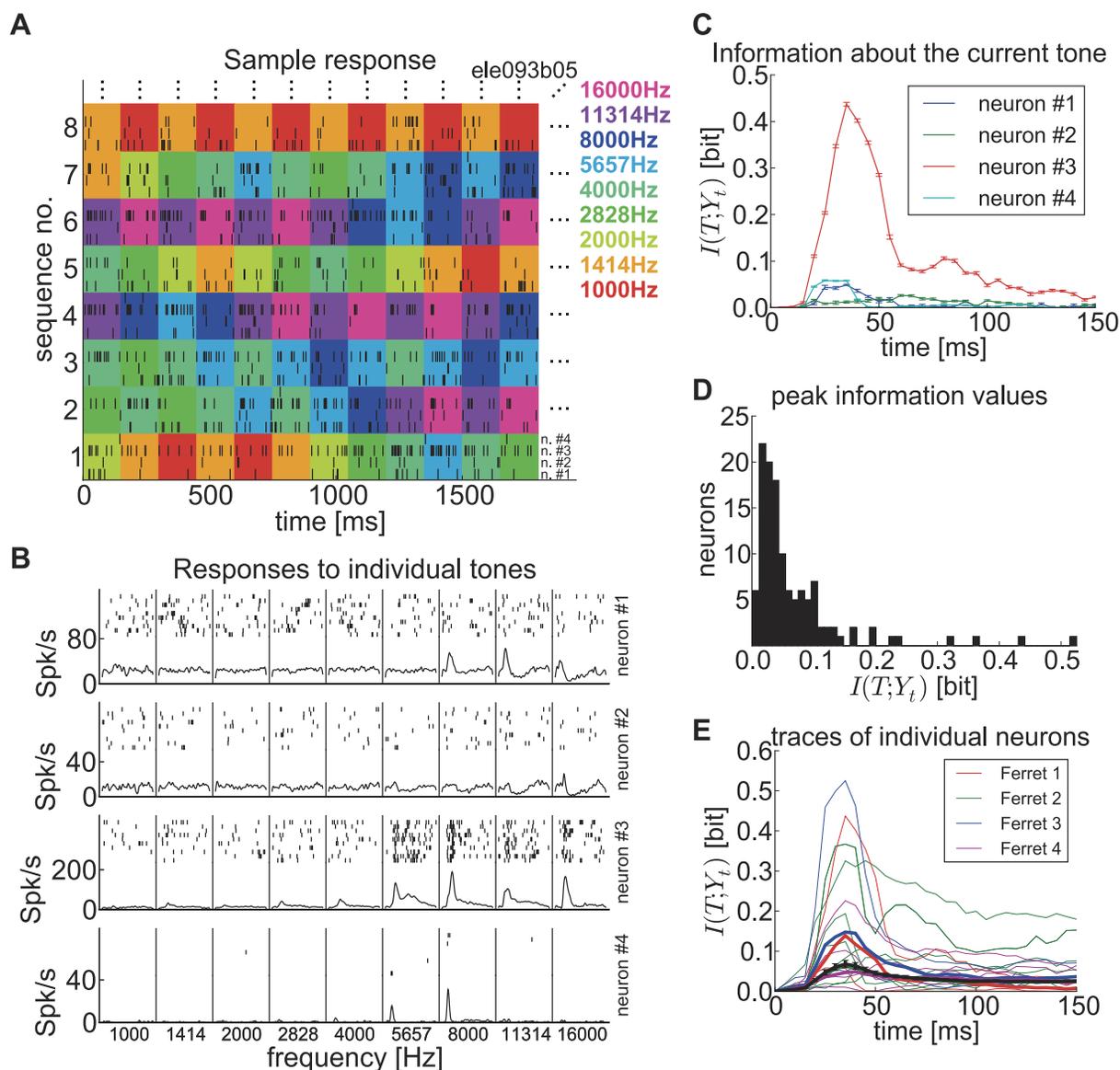


Fig. 1. Most neurons convey a significant amount of information about the current tone. *A*: sample stimulus sequences and neural responses for the first 12 tones of the first 8 sequences played during 1 particular recording (*ferret 1*). In total, the stimuli for this recording consisted of 54 sequences of 100 tones each. The background color indicates the frequency of the current tone, as denoted at *right*. In this experiment, 4 neurons were simultaneously recorded. Black lines show the spike times of the 4 different neurons (from *bottom* to *top* within each row, *neurons 1–4*). *B*: frequency tuning of *neurons 1–4* in *A*. The *top* of each row shows 10 individual spike trains of 1 neuron in response to each frequency during tone presentation periods (150 ms) of 10 randomly chosen occurrences of that frequency within the stimulus sequences. The *bottom* of each row shows the peristimulus time histograms (PSTHs) for that neuron (bin size, 1 ms). Note the different scales on the *y*-axis. *C*: time courses of mutual information (MI) between the responses of *neurons 1–4* and the currently played tone for the recording shown in *A* and *B*. MI values are estimated every 5 ms from the spike train in the 20-ms time window preceding the time indicated on the *x*-axis. Error bars denote the standard error of the mean (SE) of the MI estimator. *D*: histogram of peak MI values across the tone duration (150 ms) of all 122 neurons. *E*: MI traces (as in *C*) for 16 sample neurons across all 4 animals, including the neurons with maximal information for each animal. The other 12 neurons were randomly selected. Thick lines represent the average information trace for each animal; the black line shows the average across all neurons from all animals.

rons with the same tuning for the current tone but a differential modulation by the preceding tone.

This investigation of the responses of A1 neurons already provides qualitative insight into the temporal integration capabilities of the auditory system. It suggests that not only information about the currently played tone, but also about whether the previous tone was higher or lower, might be encoded in differences in both the mean firing rates across the whole tone duration and the timing of the spikes relative to the tone onset. Although this has already been demonstrated qualitatively through a number of previous findings, we provide in this

article a novel quantitative analysis of these effects using recently developed methods of information estimation (Nikolic et al. 2009; Panzeri et al. 2007).

Direct Estimation of Mutual Information

The most principled and rigorous way to quantify the information contained in neural responses is to directly estimate the mutual information between stimulus and response. Here, we used a recently proposed method (Panzeri et al. 2007) to estimate information conveyed by individual neurons about current and preceding tones.

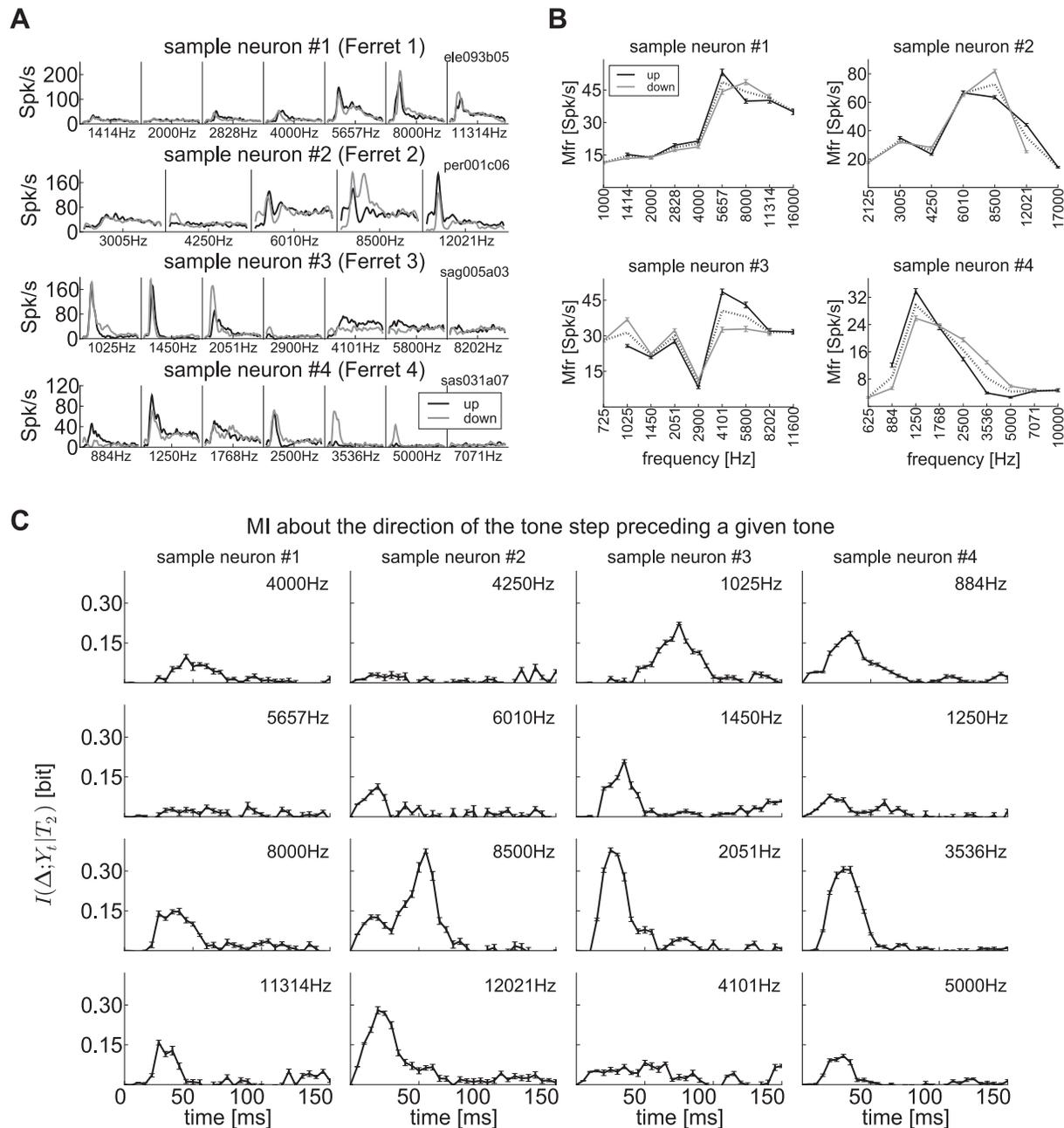


Fig. 2. Many neurons convey significant information about the direction of the preceding tone step. **A**: the PSTHs of 4 sample neurons (1 for each ferret; bin size, 1 ms) to each tone frequency are substantially different, depending on whether the frequency of the previously played tone was higher (down step, gray) or lower (up step, black). **B**: conditional tuning curves for the sample neurons in **A** show this difference in the mean firing rate of these neurons during 150-ms periods in response to all available frequencies, depending on whether the next lower or the next higher frequency has been played immediately before. The black dotted line shows the overall tuning curve calculated from all responses independently of the previous tone. Error bars denote SE. **C**: time course of mutual information $I(\Delta; Y_t | T_2)$ about the direction of the tone change (Δ) preceding a given frequency during the current tone T_2 . Columns correspond to different recordings; rows correspond to different values of the current tone. MI values are estimated every 5 ms from the spike train in the 20-ms time window preceding the time indicated on the x-axis. Error bars denote SE of the MI estimator. Spk/s, spikes/s; Mfr, mean firing rate.

Most neurons convey a significant amount of information about the currently played tone. We first investigated how much information is contained in the neural responses about the tone frequency which is currently played. To address this question, we measured the information $I(T; Y_t)$ conveyed by the responses of individual neurons Y at time t about the frequency of the current tone T . For all available frequencies used in one recording, we viewed the response spike trains of one neuron during all occurrences of that frequency as individual trials in

response to the same stimulus (see Fig. 1B). From these stimulus-response associations, we then calculated the mutual information every 5 ms throughout the tone duration of 150 ms.

Figure 1C shows the time courses of information conveyed by the individual neurons shown in Fig. 1, A and B. In this recording, neuron 3 conveyed the largest amount of information about the current tone frequency. As shown in the responses to individual frequencies in Fig. 1B, this neuron was the most selective because it responded strongly and reliably to

higher frequencies. The transient response after tone onset is reflected in the time course of the mutual information in that it reaches a relatively high peak at about 40 ms after onset, decreases afterwards, but remains significant during the second half of the tone (significance assessed by a label-shuffling test, $P < 0.05$; see MATERIALS AND METHODS). *Neurons 1* and *4* of this particular recording conveyed much less information, but the amount was still significant between 20 and 40 ms, which can be explained by their transient responses to some frequencies. *Neuron 2*, on the other hand, which had a rather nonselective response for all frequencies, did not convey a significant amount of information, even though it responded with a higher firing rate than *neuron 4*. The information transmitted by other neurons from different recordings across different animals varied in time course and magnitude (see Fig. 1E). These examples show that different neurons conveyed information in various ways.

Figure 1D summarizes the peak information value across the entire set of A1 neurons in our study. From the 122 neurons recorded, the responses of 75 neurons conveyed significant information about the current tone (the response of a given neuron was scored as conveying significant information if at least at 1 time point every 20 ms during the tone interval of 150 ms, a significant mutual information value was measured; significance assessed by a label-shuffling test with standard Bonferroni correction, $P < 0.05/7$; see MATERIALS AND METHODS). The most informative neuron conveyed a peak information value of 0.53 bit (measured in *ferret 3*), but this value lay below 0.1 bit for most of the neurons. The average peak information across all neurons was 0.067 bit and was not significantly different between animals (*ferret 1*: 0.143 bit; *ferret 2*: 0.060 bit; *ferret 3*: 0.124 bit; *ferret 4*: 0.066 bit). Although the average information conveyed by single neurons was substantially less than the theoretical maximum (2.75, 3.12, or 3.42 bits, depending on whether 7, 9, or 11 frequency values were used in the recording), this information can accumulate over large numbers of neurons in A1 to accurately encode the stimulus.

Neurons that conveyed large amounts of information also tended to have high firing rate responses. The information conveyed by individual neurons and their average firing rate were correlated. The value of Pearson's correlation coefficient

between mutual information and mean firing rate was significantly positive ($r = 0.69$, $P < 0.0005$) for the recording shown in Fig. 1, A–C. The overall correlation coefficient for all responses was $r = 0.50$ ($P < 0.0005$).

The same neurons simultaneously carry information about the previous tone. To look for evidence of stimulus integration over time, we next measured the information contained in the responses of individual neurons about the previously played tone. A high mutual information between the current response and the previous tone indicates that temporal integration is strong. We measured the information value $I(\Delta; Y_t | T_2)$ between the response Y at time t and the direction of the tone change (up or down, indicated by the binary variable Δ) preceding the current tone T_2 (Eq. 2 in MATERIALS AND METHODS). Because each tone was followed by either the next higher or the next lower tone, all information about the previous tone, given the current tone, was captured by this single-bit value.

Figure 2 illustrates the calculation of information between the response and the direction of the previous tone change Δ in detail for four neurons. These neurons differed substantially in their responses to some frequencies, depending on whether the higher or lower tone was played immediately before (Fig. 2A). This difference was particularly large in the transient responses. The conditional tuning curves in Fig. 2B show that these differences in the mean firing rates for a certain neuron were most prominent for the range of preferred frequencies of that neuron. Figure 2C shows the time course of information about the frequency change $I(\Delta; Y_t | T_2)$ for each of these four sample neurons for different values of T_2 . It can be seen that neurons encoded this information in various ways. The most information was contained during the transient response, but it could sometimes persist for the duration of the tone. The amount of information conveyed also varied with the current tone T_2 .

Figure 3A compares the average peak information about current and previous tones across the entire set of A1 neurons. From all the 122 neurons recorded, the responses of 27 neurons conveyed significant information about the direction of the previous tone change Δ (the response of a given neuron was scored as conveying significant information if at least at 1 time point every 20 ms during the tone interval of 150 ms, the

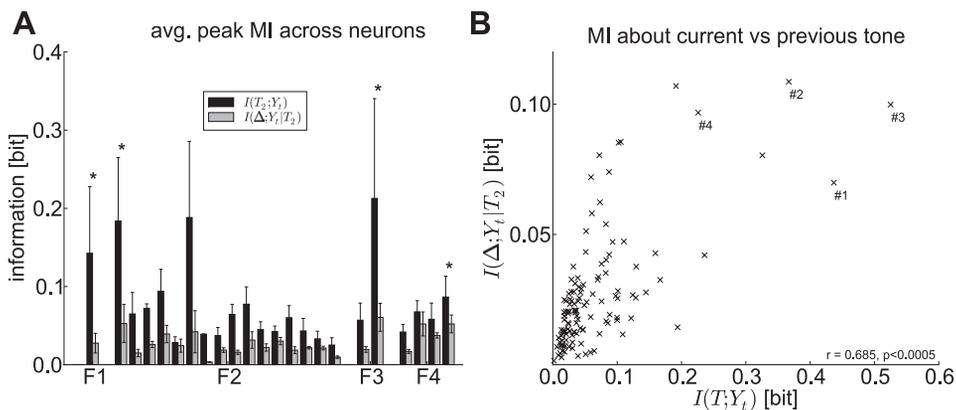


Fig. 3. Neurons simultaneously convey information about the current and previous tone. *A*: comparison of peak MI values measured about the current tone (as in Fig. 1C; black bars) and about the direction of the previous tone change (as in Fig. 2C; gray bars). For each of the 23 recordings, the average across neurons (and tones) is shown. On average, the ratio of these information values is about 3:1. Error bars denote SE; recordings are grouped by animals (F1–F4, *ferret 1–4*). Asterisks indicate recordings that include the sample neurons shown in Fig. 1. *B*: neurons simultaneously transmit information about current and previous tone. The scatter plot compares the peak MI values about current and previous tones for each of the 122 neurons. Data for sample neurons 1–4 in Fig. 1 are labeled (#1–#4).

average information across tones T_2 was significant; significance assessed by a label-shuffling test with standard Bonferroni correction, $P < 0.05/7$; see MATERIALS AND METHODS). On average, the ratio between the information values was 74:26, i.e., the information value $I(T_2; Y_i)$ was about three times larger than $I(\Delta; Y_i | T_2)$. The maximal peak information value measured about Δ was 0.158 bit in *ferret 1*, 0.391 bit in *ferret 2*, 0.378 bit in *ferret 3*, and 0.306 bit in *ferret 4*. The average peak information was 0.045 bit (*ferret 1*: 0.045 bit; *ferret 2*: 0.040 bit; *ferret 3*: 0.062 bit; *ferret 4*: 0.058 bit). These values were not significantly different between animals. Note that the maximum possible value for the information about Δ is 1 bit, whereas the maximum information about the current tone, as measured in Fig. 1, is the entropy of the distribution of tone frequencies in the respective recording (2.75, 3.12, or 3.42 bits, depending on whether 7, 9, or 11 frequency values were used in the recording).

Most of the information about the change from the previous tone can be explained by the firing rate differences in response to the two possible predecessors of a given tone (see examples in Fig. 2, A and B). There was a significant correlation between the peak value of the mutual information $I(\Delta; Y_i | T_2)$ and differences in the mean firing rate across the whole tone duration (Fig. 2B). The overall correlation coefficient was $r = 0.729$, and this correlation was also significant for individual animals (*ferret 1*: $r = 0.483$, $P = 0.009$; *ferret 2*: $r = 0.798$; *ferret 3*: $r = 0.491$; *ferret 4*: $r = 0.768$; all $P < 0.0005$).

Information values were often high for frequencies where the difference in response to the previous tone was large (e.g., sample *neuron 2*, 8,500 Hz), but also in cases where this difference was not particularly large (e.g., sample *neuron 3*, 2,051 Hz). Vice versa, a large response difference does not necessarily imply a larger amount of information (e.g., sample *neuron 4*, 1,250 Hz, or sample *neuron 3*, 4,101 Hz). This indicates that some, but not all, of the available information about the previous tone is encoded in the mean firing rates of neurons. Also, the timing of the spikes relative to the tone onset is important. For example, sample *neuron 3* conveyed a large amount of information about the tone preceding 2,051 Hz because it responded with a stronger transient to a down step (Fig. 2A), even though the mean firing rate across the whole tone duration was similar in both cases.

Figure 3B shows that the peak value of the information conveyed about the current tone, $I(T_2; Y_i)$, and the peak value of the information about Δ given the current tone, $I(\Delta; Y_i | T_2)$, are significantly correlated. This means that neurons that transmitted a large amount of information about the current tone simultaneously also tended to convey a considerable amount of information about the direction of the tone change that had led to this current tone. This indicates that there are no neurons which “specialize” on either current or previous tones, but rather that the auditory system really integrates information about previously arrived stimuli into the current responses of individual neurons in A1.

Temporal integration of information about earlier tones. To investigate whether information is integrated also about tone changes farther back in the sequence, we extended the approach in the previous section. We used the chain rule of information theory to evaluate the information about tones an increasing number of time steps in the past. More precisely, we calculated information values $I(\Delta_i; Y_i | T_n, \Delta_{n-1}, \dots, \Delta_{i+1})$ be-

tween the response Y at time t and the binary variable Δ_i , specifying whether the direction of the tone step i steps before the current tone T_n was up or down, conditioned on all the subsequent tones up to the current tone T_n (see Eq. 3 in MATERIALS AND METHODS). Note that for $n = 2$ and $i = 1$, this term is equal to the information about the direction of the preceding tone change, $I(\Delta; Y_i | T_2)$, considered in the previous section.

In Figure 4 we compared the average of these information values for one to four tone steps back across neurons and across sequences $T_n, \Delta_{n-1}, \dots, \Delta_{i+1}$. We included only those neurons that conveyed a significant amount of information about the direction of the previous tone change ($n = 27/122$). For these 27 neurons, the average information values were higher than the average across all 122 neurons: The average information about the current tone was 0.115 bit, and the average information about the direction of the tone change preceding the current tone was 0.049 bit. The responses of 25, 12, and 6 neurons conveyed significant information about the tone change Δ_2, Δ_3 , and Δ_4 , respectively (at least at 1 time point during the tone interval of 150 ms, the average information across sequences $T_n, \Delta_{n-1}, \dots, \Delta_{i+1}$ was significant). It can be seen that the information decreases for increasing number of tone steps considered. After two tone steps, the information saturates at a nonsignificant low residual value. This would indicate that the response does not contain any information about the tone step direction more than two tones in the past. However, we are exploring a large stimulus space because we have to average over all possible sequences of length i . Thus the number of available samples for mutual information estimation decreases by one-half for each additional tone step considered. Therefore the information values for many backward time steps are less reliable, and we cannot exactly measure the influence of earlier tones. Moreover, because the information bias cannot be eliminated for arbitrarily small sample sizes, the absolute level of information might be biased for longer time lags.

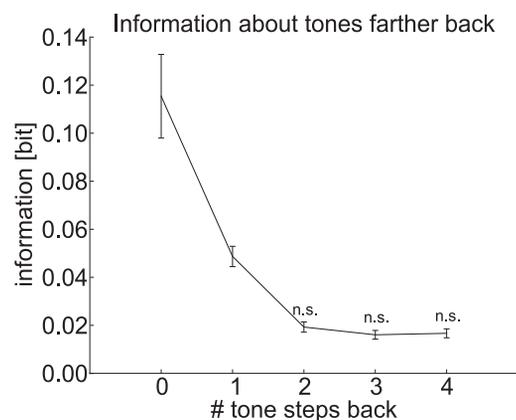


Fig. 4. No additional information can be gained about the direction of the tone step more than 2 tones back. MI values are calculated between the response and the direction of the tone change the specified number of time steps in the past. The average across neurons and sequences (see text) of the peak information during a 150-ms tone duration is shown. Labels on the abscissa denote the number of time steps back (0: information about the current tone; 1: information about the previous tone, 2: two steps back, etc.). Error bars denote SE; n.s., nonsignificant deviation from chance level, estimated by a label-shuffling test ($P > 0.05$).

Information Analysis Using Linear Classifiers

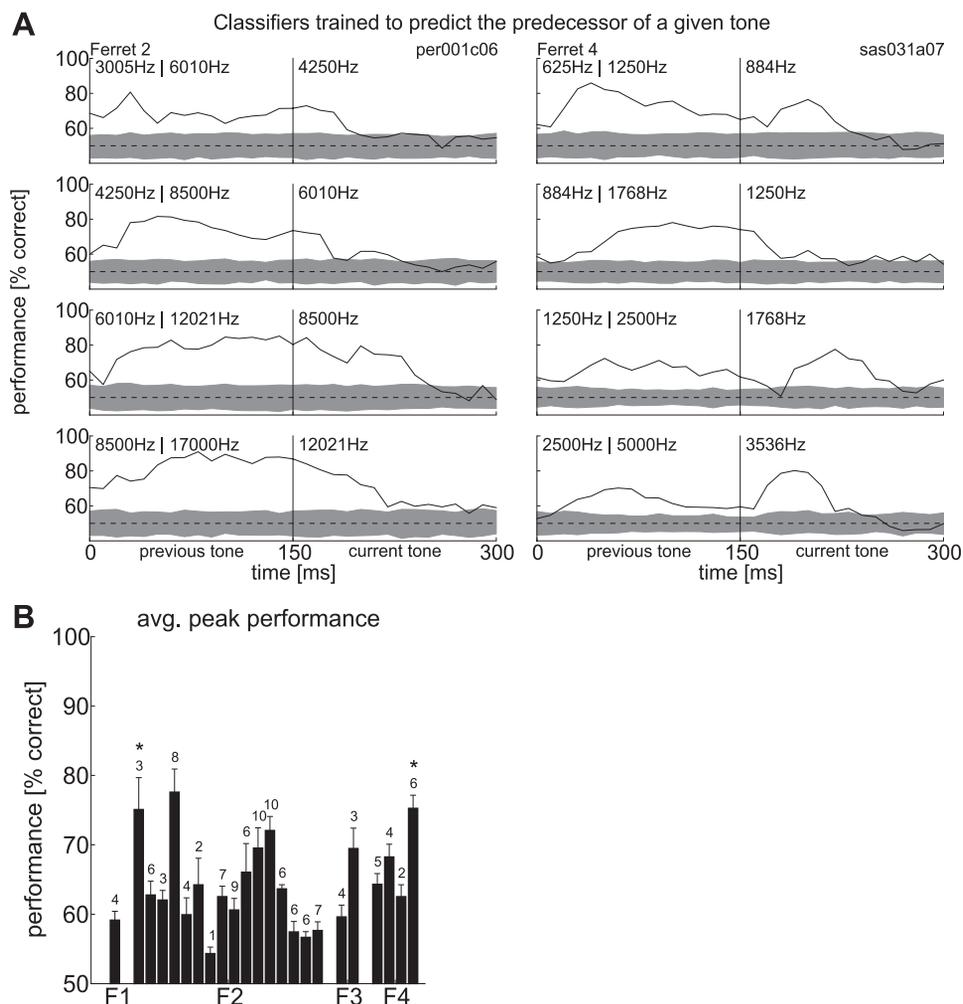
An alternative way to investigate the temporal integration of information is to measure the amount of information about current and past stimuli that can be extracted by a neuron that reads the current activity of all simultaneously recorded neurons. To analyze this information, we trained (binary) linear classifiers, SVMs with a linear kernel (see MATERIALS AND METHODS), on the spike trains of simultaneously recorded neurons to decode information about current and previous tones (see Nikolic et al. 2009 for another application of this method). We also used linear classifiers to analyze the nonlinear processing capability of the auditory system and to investigate the fraction of the total mutual information that was linearly decodable.

Information about the current and previous tone can be extracted by linear decoders. First, we investigated the performance of these classifiers on the current tone. We selected two specific frequency values and collected the responses to all occurrences of these two frequencies. We low-pass filtered the spike responses ($\tau = 20$ ms), and to get as good an estimate of the linearly decodable information as possible, we trained a different classifier every 10 ms to discriminate between the two possible current tones. To train a linear classifier for a specific point in time is “fair” from the perspective of neural modeling if other neurons can trigger the activation of a corresponding

readout neuron at a specific point in time relative to stimulus onset (Jin et al. 2009). For most tone pairs, performance was significantly above baseline level of 50% during the whole duration of 150 ms, across different recordings and animals. Seventeen of 23 recording sites performed significantly above chance (on average across all tone pairs). Peak classification performance (up to 90%) was often achieved within the first 50 ms, most probably due to the discriminative initial bursting behavior of some neurons. Performance typically decreased for later time points. The average peak performance across all tone pairs and experiments was 59.75% (*ferret 1*: 65.23%; *ferret 2*: 58.38%; *ferret 3*: 62.46%; *ferret 4*: 61.21%). In principle, one could also investigate the performance of multiclass classifiers on all available tone frequencies. However, these multiclass classifiers are essentially a combination of different binary linear classifiers, which either distinguish between one frequency and the rest (one-versus-all strategy) or between every pair of frequencies (one-versus-one strategy). Hence, they can only extract information that would also be extracted by standard binary linear classifiers.

We then used linear classifiers to investigate the temporal integration of information, i.e., the information contained in the responses about the previously played tone. Figure 5A shows examples of performance of linear classifiers trained on all simultaneously recorded neurons to discriminate between the two possible predecessors of the current tone, i.e., to extract

Fig. 5. Linear classifiers are able to discriminate between the 2 possible predecessors of a given tone. **A**: performance of linear classifiers trained at time points every 10-ms during the 300-ms interval of 2 consecutive tones T_1 , T_2 to discriminate between the 2 possible predecessors T_1 (previous tone) of a given tone T_2 (current tone) (i.e., to predict Δ) for 2 example recordings from 2 different animals. Columns correspond to different recordings; rows correspond to different values of T_2 . Performance is evaluated by 10-fold cross-validation. The dashed black line indicates baseline performance of 50%. The gray shading around the baseline performance denotes the region of nonsignificant deviations from chance level, estimated by a label-shuffling test ($P > 0.05$). **B**: average peak performance values. For each of the 23 recordings, the peak classification performance during the duration of the second tone T_2 is shown (average across frequencies). Numbers denote the dimensionality of the classifier input (i.e., the number of simultaneously recorded neurons). Error bars denote SE; recordings are grouped by animals F1–F4. Asterisks indicate the recordings shown in **A**.



information about Δ in Eq. 2. We trained a different classifier every 10 ms to discriminate between the two different frequencies of the first tone T_1 in the pair (T_1, T_2) for a given tone T_2 . In Fig. 5A, performance is also shown during the preceding tones. In this way we could analyze the information extracted by classifiers about which of two different frequencies was currently played and how this information is maintained during the next tone. One sees that in most cases performance stays above chance level for some time after the switch of the tone frequency before it drops, indicating that the information is maintained about the stimulus after the tone has switched. In other cases, performance reached a second (albeit smaller) peak after the tone switch. This is most probably to differences in the transient responses of some neurons.

Figure 5B summarizes the average peak classifier performance for each data set (averaged over tones T_2 of the maximum performance in extracting information about T_1 during the period of T_2). Classifiers were able to extract a considerable amount of information about the previous tone across different frequencies for most recordings. Eighteen of 23 recording sites performed significantly above chance (on average across all tones T_2). The maximal measured performance achieved was 88.17% (in *ferret 2*, Fig. 5A, left). The average peak performance was 64.92% (*ferret 1*: 59.18%; *ferret 2*: 64.40%; *ferret 3*: 63.86%; *ferret 4*: 68.92%). Note, however, that direct comparisons of the classification performance between different recordings should be viewed with caution, because the performance values depend on several factors, such as the number of neurons simultaneously recorded (ranging from 1 to 10, as shown in Fig. 5B) or the total number of spikes in the recording. Similar to the mutual information values about the previous tone, peak classification performance was also strongly correlated with the absolute firing rate differences of individual neurons (Fig. 2B). The overall correlation coefficient was $r = 0.394$ (*ferret 1*: $r = 0.483$, $P = 0.192$; *ferret 2*: $r = 0.411$, $P < 0.0005$; *ferret 3*: $r = 0.459$, $P = 0.001$; *ferret 4*: $r = 0.272$, $P = 0.003$).

These performance values of linear classifiers on the previous tone are on average similar or even slightly higher than the performance values on the current tone. Note that this finding does not contradict the previously reported 3:1 ratio of mutual information about current vs. previous tone, because the mutual information about the current tone is evaluated using all available frequencies at the same time, whereas for the classifier, only pairs of frequencies are considered. To obtain a fair comparison, we also estimated the mutual information about the current tone for each pair of tones. The average peak information was 0.044 bit (*ferret 1*: 0.044 bit; *ferret 2*: 0.070 bit; *ferret 3*: 0.039 bit; *ferret 4*: 0.078 bit). These information values are similar to the values of $I(\Delta; Y_i | T_2)$ reported previously and confirm the approximate 1:1 ratio of the classification performance. These findings demonstrate the prominent temporal integration of information present in the responses of A1 neurons.

Nonlinear superposition of information. Furthermore, we analyzed whether the neural response provides a nonlinear superposition of information about sequentially arriving stimuli. This property is beneficial for information processing because it boosts the computational power of a neuron that reads this neural response (like a kernel in the terminology of machine learning). In this way a linear neuron is effectively

able to compute a nonlinear function of the input. Such nonlinear superposition can be proved if a linear classifier is able to reproduce the XOR computation of two bits in the input sequence. This simple binary computation yields result “1” if these two bits are different and “0” if they are the same, but it cannot be solved by any linear model. If a linear decoder is able to predict the resulting bit of information from the responses, the neural system itself has to provide the necessary nonlinear combination. Nikolic et al. (2009) performed such an analysis for neural responses of the primary visual cortex of anesthetized cat and showed that these responses exhibited a significant nonlinear combination of sequentially presented stimuli in the visual field. In our case, we evaluated the XOR performance of a classifier for our particular stimulus sequences in the following way: We collected responses to subsequences ABA, ABC, CBA, and CBC of three tones A, B, and C [note that a tone B is both preceded and followed by either tone A (lower) or C (higher)] and checked whether a linear classifier was able to predict the bit that resulted from the XOR combination of whether the first and the third tone in these subsequences was equal or not for a fixed frequency of the intervening tone.

To ensure that any significant classification performance on this nonlinear XOR combination can really be attributed to a nonlinearity implicitly provided by the neural responses and not to the nonlinear classification threshold, we measured the performance as the point-biserial correlation coefficient between the binary target variable and the continuous linear combination of the low-pass-filtered input spike trains learned by the classifier (Nikolic et al. 2009) (i.e., the output of the classifier before the threshold operation). It can be shown that any such correlation coefficient significantly greater than zero indicates nonlinear transformations in the neural processes itself (see Nikolic et al. 2009 for a formal proof). Figure 6 shows that there was a significant performance increase above chance level for a considerable period during the third tone in the sequence for at least three different recordings of two different animals, and for some values of tone B. The performance of most recordings, however, was only slightly above the significance level for only a limited number of time points, and often only for a single tone. Note that this finding provides evidence for both nonlinear superposition and temporal integration of information because a past stimulus is involved in the nonlinear computation.

Linear decoders are able to extract a substantial fraction of the total information. Finally, we addressed the question how much of the total information about the stimulus is accessible to linear decoders, and thus to hypothetical readout neurons of A1 responses. For that we compared the direct estimation of mutual information with the achieved performance of a linear classifier. For a direct comparison, we used the vector composed of these spike counts (20-ms sliding window) for all simultaneously recorded neurons as input both to the linear classifier and for mutual information estimation. Note that, in contrast to the previous mutual information measurements, we calculated here the combined information conveyed by all simultaneously recorded neurons, instead of individual neurons. The mutual information between the classifier prediction and the actual stimulus could then be compared with the mutual information directly estimated from the vector of spike counts.

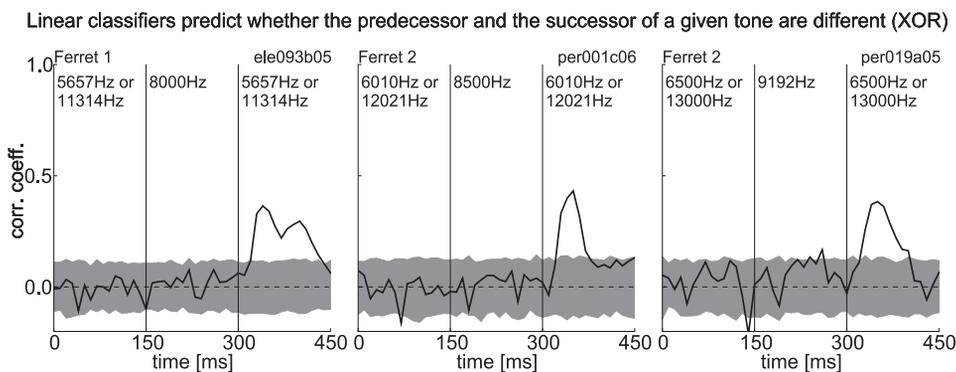
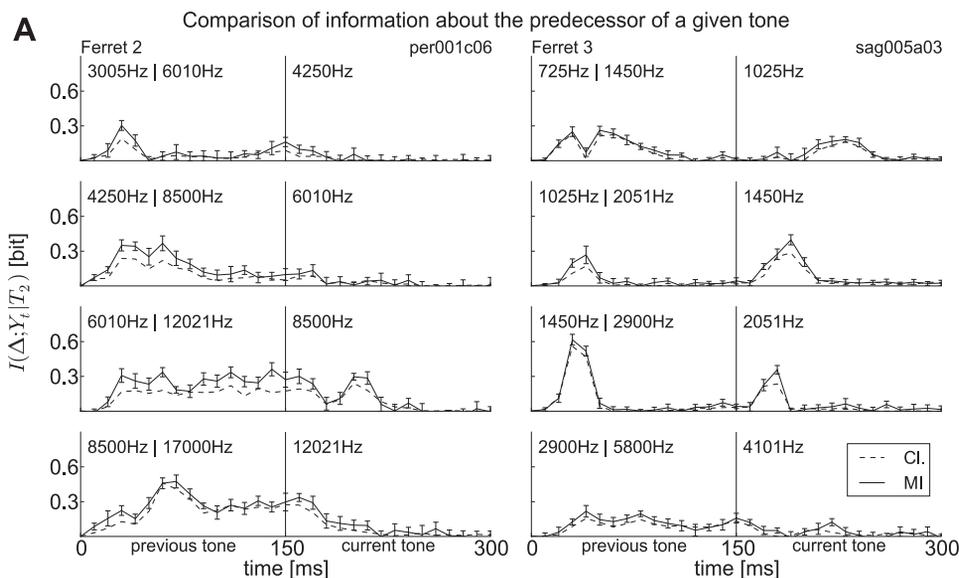


Fig. 6. Linear classifiers reveal a nonlinear superposition of information. Performance of linear classifiers trained every 10 ms to predict a nonlinear exclusion-or (XOR) computation on the input sequence: whether the first and the third tone of a 3-tone sequence were equal or not for a fixed frequency of the intervening tone. Shown is the performance from 1 frequency of 3 different recordings from 2 different ferrets. The performance is assessed by the point-biserial correlation coefficient between the binary target variable and the continuous readout “depolarization.” The dashed black line indicates baseline correlation of 0. The gray shading around the baseline performance denotes the region of nonsignificant deviations from chance level, estimated by a label-shuffling test ($P > 0.05$).

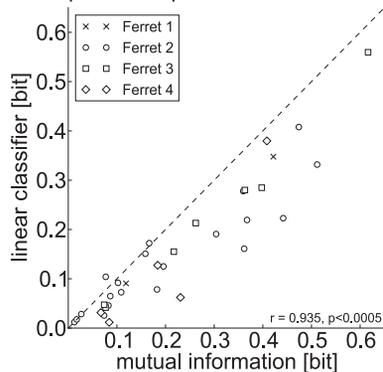
Figure 7A shows such a comparison for the information between the response and the direction of the preceding tone change (as in Figs. 2 and 5). Linear classifiers are trained to predict from the spike count information in response to two successive tones, T_1 and T_2 , which of the two possible predecessors T_1 of a given tone T_2 has been played as the first tone. This information is compared with the direct mutual information estimation. It can be seen that this amount of total information is

higher than the information of the classifiers throughout the whole duration of the tone pair, which is obvious because a linear classifier can extract only a subset of the total amount of information. What is interesting is that most of this available information is accessible to a linear classifier. This also holds for the information about the tone currently played, where the information is evaluated between the response and which of two selected frequencies are currently played.

Fig. 7. Most of the total information can be extracted by a linear classifier. A: time courses of information estimated directly (MI; solid lines) and through the performance of a linear classifier (CL; dashed lines) conveyed by all simultaneously recorded neurons about the direction of the tone step preceding a given tone T_2 . Most of the time the classifier trace stays close below the MI trace. Information values are estimated every 10 ms during the 300-ms interval of 2 consecutive tones T_1 (previous tone), T_2 (current tone) for recordings from 2 different animals using the sequence of spikes in the 20-ms time window preceding the time indicated on the x-axis. Error bars denote SE of the MI estimator. Columns correspond to different recordings; rows correspond to different values of T_2 . B: each point in the scatter plot compares MI calculated directly from the neural response and from the output of the linear classifier for experiments, where MI could be reliably estimated. Information values are taken at the time point of maximum information throughout the 300-ms duration shown in A. Points that lie closely under the diagonal (dashed line) denote recordings where much of the total information was accessible to linear classifiers (median ratio 0.689, Pearson correlation $r = 0.935$).



B Comparison at peak information value



In Fig. 7B we compare the mutual information analysis and the training of linear classifiers for many recordings across different animals. The graph shows that linear classifiers extract a fair portion of, but in general not all, information that is available in the spiking activity. For some recordings, however, both methods extract roughly the same amount of information. The median ratio between both information values was 0.689, i.e., about 69% of the information is accessible to linear classifiers. The correlation between both information values was $r = 0.935$, i.e., for neural responses that contained much information, linear classifiers also tended to decode a lot of information, and this effect was consistent across recordings. However, several recordings did not allow a direct comparison because no reliable estimate for the combined information of all neurons could be generated. This is because the response space is undersampled, either because there were too few occurrences of a given frequency (pair) or because the response space was too large (too many simultaneously recorded neurons). Figure 7B shows only mutual information values for such recordings and frequencies where a reliable mutual information estimation was possible. The reliability of an estimate was determined by evaluating the mutual information on a random subset of one-half of all available trials. If the mutual information value changed by less than 10% of its value, a reliable estimate was reported. This suggests that A1 provides an effective generic preprocessing of the temporal structure of acoustic stimuli and provides a suitable neural response that makes a large portion (but not all) of this information available to higher areas that can easily read out this information, e.g., via a simple linear neuron.

DISCUSSION

In this study we investigated the well-known ability of the auditory system to integrate the rich temporal structure of the acoustic environment into a neural response that facilitates further processing of the stimulus information by higher areas. In contrast to previous studies, we addressed this question in a qualitative manner. More precisely, we measured the amount of temporal integration present in responses of neurons in the primary auditory cortex of awake ferrets by measuring the information about both past and present stimuli.

Quantitative Analysis of the Temporal Integration of Information About Current and Preceding Sounds

We compared two methods for quantifying the impact of the immediate history of auditory stimulation onto the neural responses in primary auditory cortex: the direct estimation of mutual information between the response and stimulus parameters (Panzeri et al. 2007) and the training of linear discriminators to decode stimulus parameters (Nikolic et al. 2009). To the best of our knowledge, this is the first study that directly compares the analysis of multichannel spike data via mutual information with an analysis using linear classifiers. Our stimuli, which consisted of tone sequences changing by a fixed step up or down, captured dynamic features typical of many naturally occurring sounds for which both the current and preceding frequency are important for accurate perception (Singh and Theunissen 2003). The fact that we observe information about tone history encoded in the neural responses suggests that neurons in A1 would produce similar representations also for more complex spectrotemporal dynamics, which would enable the efficient discrimination of natural sounds, as well. Our use

of a wide range of stimulus frequencies allowed us to use information metrics effectively and to build on previous studies of auditory context (Asari and Zador 2009; Bartlett and Wang 2005; Brosch et al. 1999; Brosch and Schreiner 2000; Doupe 1997; Kilgard and Merzenich 1999; Lewicki and Arthur 1996; Malone et al. 2002; Margoliash and Fortune 1992; McKenna et al. 1989; Ulanovsky et al. 2004; Yin et al. 2008).

It is possible that sampling stimulus frequencies at half-octave steps misses neurons that are more narrowly tuned. For example, *neuron 4* in Fig. 1, A and B, might have responded more strongly if we had tested a wider range of stimuli such as noise or natural sounds. Technically, because the neural responses are discretized into binary words, their strength makes no difference for the information analysis. We employed a significance test, which removes spurious contributions from weakly responding or nonselective responses (such as *neuron 2* in Fig. 1, A and B). Despite the large variability in the average responses of different neurons, the direct estimation of mutual information revealed that many neurons simultaneously conveyed information about both the current sound (61%) and the immediately preceding sound (22%).

Although most neurons transmitted a significant amount of information, the absolute value of the maximal information varied substantially and tended to be lower in neurons with very sparse responses (DeWeese et al. 2003, 2005; Hromádka et al. 2008). Moreover, we found that the same neurons that conveyed a large amount of information about the current tone also conveyed a considerable amount of information about the previous tone. This supports the hypothesis that auditory neurons integrate information about preceding sounds into their current responses, rather than encoding these pieces of information separately between different populations.

A large amount of the measured information about both the current and the previous tone was encoded in the mean firing rates of neurons, rather than in their fine spike timing. Neurons that showed a clear tuning to frequency typically conveyed a larger amount of information. However, coarse timing of neural responses was also crucial. As in many previous studies, we observed both transient and sustained responses (Hromádka and Zador 2009; Lu et al. 2001; Wang et al. 2005). Some neurons showed a strong transient response to a tone change that is beneficial for information transmission. Moreover, responses that differed in the strength of their transients for different stimuli often contained a large amount of information, even if the mean firing rate across the whole tone duration was similar for these stimuli. Recent evidence suggests that the role of precise spike timing is more prominent for stimuli varying on a faster time scale and therefore depends on the particular stimulus dynamics (Kayser et al. 2010).

A Large Amount of the Information About Preceding Sounds Can Be Extracted by Linear Decoders

Our comparison of information and classifier methods revealed that the information contained in the responses about current and preceding tones is largely accessible to linear classifiers. This suggests that the primary auditory cortex provides a neural response that facilitates the task of later processing stages to instantaneously read out information about the complex temporal structure of the acoustic stimuli. Linear decoders showed a significant performance above chance level

when trained to discriminate between the two possible predecessors of a given tone. In general, the amount of information that could be decoded in a linear manner was lower than the value obtained by direct estimation of mutual information, indicating that linear classifiers were able to capture not all available information contained in the spiking activity of A1 neurons. Still, in many cases these classifiers extracted a considerable portion of the total information. Such efficiency of linear decoding mechanisms has been previously reported, e.g., in the fly visual system (Rieke et al. 1997).

Moreover, information about current and previous tones did not occur by a simple linear superposition of information. Instead, the neural response appears to be a nonlinear combination of sequentially arriving inputs, as revealed by the successful XOR classification. The XOR computation is a very simple function that tells whether the current tone and the tone two steps back are the same for a fixed frequency of the intervening tone. However, this problem cannot be solved by a linear model. The fact that linear decoders are able to predict the resulting bit of information from the responses suggests that the neural system itself provides the necessary nonlinear combination. Evidence for such nonlinear interactions has recently been reported (Sadagopan and Wang 2009), where A1 neurons were found to be sensitive to nonlinear combinations of spectral and temporal stimulus properties.

However, neither method revealed that A1 neurons encode significant information about the direction of the previous tone step, independent of the frequency of the current tone. Detection of frequency changes independent of base frequency may be important for some behaviors (Yin et al. 2010), but the information must be extracted from stimuli in other cortical areas.

Mechanisms for Integration of Stimulus History Into Neuronal Responses

A number of possible mechanisms have already been studied that could allow information about previously played tones to persist in the responses to the current tone. First, this persistence could be implemented in an earlier auditory area responsible for some kind of echoic memory (Anderson et al. 2009). Second, the A1 neurons themselves could “remember” the influence of previous inputs in their biophysical state, e.g., by mechanisms of short-term plasticity of their synapses (David et al. 2009; Elhilali et al. 2004; Wehr and Zador 2005; Zucker and Regehr 2002). Third, the integration of stimulus history into the neuronal responses could be implemented by lateral (possibly inhibitory) inputs from other neurons. Information about a preceding tone could be contained in the activity of inhibitory inputs, which then shape the subsequent responses to the current tone. There are many different possible pathways, and it could also be possible that these processes involve inputs from higher cortical areas. Although we do not propose a specific mechanism that achieves the observed representation, our study provides the first quantitative analysis of these aspects in the responses of A1 neurons.

Neuronal adaptation is a related mechanism that has been extensively studied in the auditory context (see e.g., Condon and Weinberger 1991; Malone and Semple 2001; Malone et al. 2002; Ulanovsky et al. 2003, 2004). One effect of neuronal adaptation is to maximize information transmission by matching the neural code to the stimulus statistics (Fairhall et al.

2001). Ulanovsky et al. (2004) studied such stimulus-specific adaptation (SSA) for stimulus sequences of pure tones. This study found evidence that A1 responses are influenced by tones up to four or five steps in the past. Although we found a similar effect in our data, we had limited statistical power to measure information about stimuli more than two steps back in time because we sampled frequencies over a large range of the tuning curve of the neurons to estimate mutual information. Given the previous reports of SSA in A1 and other contextual effects in V1 (Nikolic et al. 2009), such long-lasting dependence of information on preceding stimuli is likely.

On the other hand, the influence of earlier stimuli on the current neural responses is probably affected by anesthesia, e.g., it might change the impact of inhibitory neurons so that they are not able to reset cortical circuits after a stimulus as they might do in awake animals. In contrast to the related studies (Nikolic et al. 2009; Ulanovsky et al. 2004), we presently report results from awake animals, and it is possible that the long-lasting persistence of information reported there is at least partly an effect of anesthesia.

Novel Experimental Evidence for the Liquid Computing Model

Liquid computing (Buonomano and Maass 2009; Maass et al. 2002) has emerged as a framework for understanding computations in biological networks of neurons consisting of diverse types of neurons and synapses with different time constants. This model proposes two fundamental operations of neural circuits: to provide 1) analog fading memory to accumulate incoming information over time in the current neural activity and 2) a nonlinear projection into a space of typically higher dimension than the input space. With this generic preprocessing, even simple static linear neurons that “read” only the neural response at one point in time are able to extract information about the stimulus that is nontrivially spread out in time. Our analysis revealed significant evidence for both of these operations: sequentially arriving stimulus information is integrated over time and superimposed in a nonlinear manner onto the neural responses at one point in time. Already at this early stage of sensory processing, the neural system transforms the auditory information in a way that eases the extraction of information by later processing stages. This is further supported by our finding that linear classifiers are able to extract a considerable amount of the total mutual information between stimulus and response.

Our results provide the first test of the predictions of the liquid computing model in the auditory system and the first test in the sensory system of awake animals. In a different context, evidence has also been reported in the prefrontal, cingulate, and parietal cortex of awake monkeys (Bernacchia et al. 2011). Recent observations in the primary visual cortex of anesthetized cats (Nikolic et al. 2009) were similar to ours, but the effects of anesthesia on the effects reported there are unclear. Moreover, the data from that study did not permit the direct application of standard information measures because the relevant information was spread over too many neurons. Our findings suggest that for many neural systems, the computationally very efficient analysis of information with the use of linear classifiers might provide almost as good an estimate for the actual information as the direct estimation of information between stim-

ulus and response, which can only be applied to recordings from fewer neurons because of undersampling problems. Our analysis also suggests a new perspective in the experimental analysis of neuronal coding in sensory systems, which had so far focused on codes for the currently present stimulus.

ACKNOWLEDGMENTS

We thank Stefan Häusler and Stefano Panzeri for stimulating comments and discussions.

GRANTS

This work was supported by National Institute for Deafness and Other Communication Disorders Grants R01 DC005779 and F32 DC008453 and by European Union (ORGANIC) Project FP7-231267.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

Author contributions: S.K., S.V.D., P.Y., S.A.S., and W.M. conception and design of research; S.K. and P.Y. analyzed data; S.K., S.V.D., P.Y., S.A.S., and W.M. interpreted results of experiments; S.K. prepared figures; S.K. drafted manuscript; S.K., S.V.D., and W.M. edited and revised manuscript; S.K., S.V.D., P.Y., S.A.S., and W.M. approved final version of manuscript; S.V.D. and P.Y. performed experiments.

REFERENCES

- Anderson LA, Christianson GB, Linden JF. Stimulus-specific adaptation occurs in the auditory thalamus. *J Neurosci* 29: 7359–7363, 2009.
- Arabzadeh E, Panzeri S, Diamond ME. Whisker vibration information carried by rat barrel cortex neurons. *J Neurosci* 24: 6011–6020, 2004.
- Arabzadeh E, Panzeri S, Diamond ME. Deciphering the spike train of a sensory neuron: counts and temporal patterns in the rat whisker pathway. *J Neurosci* 26: 9216–9226, 2006.
- Asari H, Zador AM. Long-lasting context dependence constrains neural encoding models in rodent auditory cortex. *J Neurophysiol* 102: 2638–2656, 2009.
- Bartlett EL, Wang X. Long-lasting modulation by stimulus context in primate auditory cortex. *J Neurophysiol* 94: 83–104, 2005.
- Ben-Hur A, Ong CS, Sonnenburg S, Schölkopf B, Rätsch G. Support vector machines and kernels for computational biology. *PLoS Comput Biol* 4: e1000173, 2008.
- Bernacchia A, Seo H, Lee D, Wang XJ. A reservoir of time constants for memory traces in cortical neurons. *Nat Neurosci* 14: 366–372, 2011.
- Bishop CM. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. New York: Springer, 2006.
- Brosch M, Schreiner CE. Sequence sensitivity of neurons in cat primary auditory cortex. *Cereb Cortex* 10: 1155–1167, 2000.
- Brosch M, Schulz A, Scheich H. Processing of sound sequences in macaque auditory cortex: response enhancement. *J Neurophysiol* 82: 1542–1559, 1999.
- Buonomano D, Maass W. State-dependent computations: spatiotemporal processing in cortical networks. *Nat Rev Neurosci* 10: 113–125, 2009.
- Chang CC, Lin CJ. *LIBSVM: a Library for Support Vector Machines* (Online). Chang CC, Lin CJ. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Condon CD, Weinberger NM. Habituation produces frequency-specific plasticity of receptive fields in the auditory cortex. *Behav Neurosci* 105: 416–430, 1991.
- Cover TM, Thomas JA. *Elements of Information Theory*. New York: Wiley, 1991.
- David SV, Mesgarani N, Fritz JB, Shamma SA. Rapid synaptic depression explains nonlinear modulation of spectro-temporal tuning in primary auditory cortex by natural stimuli. *J Neurosci* 29: 3374–3386, 2009.
- DeWeese MR, Hromádka T, Zador AM. Reliability and representational bandwidth in the auditory cortex. *Neuron* 48: 479–488, 2005.
- DeWeese MR, Wehr M, Zador AM. Binary spiking in auditory cortex. *J Neurosci* 23: 7940–7949, 2003.
- Doupe AJ. Song- and order-selective neurons in the songbird anterior fore-brain and their emergence during vocal development. *J Neurosci* 17: 1147–1167, 1997.
- Elhilali M, Fritz JB, Klein DJ, Simon JZ, Shamma SA. Dynamics of precise spike timing in primary auditory cortex. *J Neurosci* 24: 1159–1172, 2004.
- Fairhall AL, Lewen GD, Bialek W, de Ruyter van Steveninck RR. Efficiency and ambiguity in an adaptive neural code. *Nature* 412: 787–792, 2001.
- Hromádka T, DeWeese MR, Zador AM. Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS Biol* 6: e16+, 2008.
- Hromádka T, Zador AM. Representations in auditory cortex. *Curr Opin Neurobiol* 19: 430–433, 2009.
- Ince RA, Petersen RS, Swan DC, Panzeri S. Python for information theoretic analysis of neural data. *Front Neuroinform* 3: 4, 2009.
- Jin DZ, Fujii N, Graybiel AM. Neural representation of time in cortico-basal ganglia circuits. *Proc Natl Acad Sci USA* 106: 19156–19161, 2009.
- Kayser C, Logothetis NK, Panzeri S. Millisecond encoding precision of auditory cortex neurons. *Proc Natl Acad Sci USA* 107: 16976–16981, 2010.
- Kilgard MP, Merzenich MM. Distributed representation of spectral and temporal information in rat primary auditory cortex. *Hear Res* 134: 16–28, 1999.
- Lewicki MS, Arthur BJ. Hierarchical organization of auditory temporal context sensitivity. *J Neurosci* 16: 6987–6998, 1996.
- Lu T, Liang L, Wang X. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat Neurosci* 4: 1131–1138, 2001.
- Maass W, Natschläger T, Markram H. Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput* 14: 2531–2560, 2002.
- Malone BJ, Scott BH, Semple MN. Context-dependent adaptive coding of interaural phase disparity in the auditory cortex of awake macaques. *J Neurosci* 22: 4625–4638, 2002.
- Malone BJ, Semple MN. Effects of auditory stimulus context on the representation of frequency in the gerbil inferior colliculus. *J Neurophysiol* 86: 1113–1130, 2001.
- Margoliash D, Fortune ES. Temporal and harmonic combination-sensitive neurons in the zebra finch's hvc. *J Neurosci* 12: 4309–4326, 1992.
- McKenna TM, Weinberger NM, Diamond DM. Responses of single auditory cortical neurons to tone sequences. *Brain Res* 481: 142–153, 1989.
- Miller GA. Note on the bias of information estimates. In: *Information Theory in Psychology: Problems and Methods*, edited by Quastler H. Glencoe, IL: Free Press, 1955, p. 95–100.
- Montemurro MA, Rasch MJ, Murayama Y, Logothetis NK, Panzeri S. Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr Biol* 18: 375–380, 2008.
- Nikolic D, Häusler S, Singer W, Maass W. Distributed fading memory for stimulus properties in the primary visual cortex. *PLoS Biol* 7: 1–19, 2009.
- Panzeri S, Senatore R, Montemurro MA, Petersen RS. Correcting for the sampling bias problem in spike train information measures. *J Neurophysiol* 98: 1064–1072, 2007.
- Panzeri S, Treves A. Analytical estimates of limited sampling biases in different information measures. *Network* 7: 87–107, 1996.
- Rieke F, Warland D, de Ruyter van Steveninck R, Bialek W. *Spikes—Exploring the Neural Code*. Cambridge, MA: MIT Press, 1997.
- Sadagopan S, Wang X. Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex. *J Neurosci* 29: 11192–11202, 2009.
- Schölkopf B, Smola AJ. *Learning with Kernels*. Cambridge, MA: MIT Press, 2002.
- Shannon CE. A mathematical theory of communication. *AT&T Bell Labs Tech J* 27: 379–423, 1948.
- Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoust Soc Am* 114: 3394–3411, 2003.
- Sussillo D, Abbott LD. Generating coherent patterns of activity from chaotic neural networks. *Neuron* 63: 544–557, 2009.
- Ulanovsky N, Las L, Farkas D, Nelken I. Multiple time scales of adaptation in auditory cortex neurons. *J Neurosci* 24: 10440–10453, 2004.
- Ulanovsky N, Las L, Nelken I. Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6: 391–398, 2003.

- Wang X, Lu T, Snider RK, Liang L.** Sustained firing in auditory cortex evoked by preferred stimuli. *Nature* 435: 341–346, 2005.
- Wehr M, Zador AM.** Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron* 47: 437–445, 2005.
- Yin P, Fritz JB, Shamma SA.** Do ferrets perceive relative pitch? *J Acoust Soc Am* 127: 1673–1680, 2010.
- Yin P, Mishkin M, Sutter M, Fritz JB.** Early stages of melody processing: stimulus-sequence and task-dependent neuronal activity in monkey auditory cortical fields a1 and r. *J Neurophysiol* 100: 3009–3029, 2008.
- Zucker RS, Regehr WG.** Short-term synaptic plasticity. *Annu Rev Physiol* 64: 355–405, 2002.

