



University  
of Glasgow

Brasselet, R., Panzeri, S., Logothetis, N.K., and Kayser, C. (2012)  
Neurons with stereotyped and rapid responses provide a reference frame  
for relative temporal coding in primate auditory cortex. *Journal of  
Neuroscience*, 32 (9). pp. 2998-3008. ISSN 0270-6474

<http://eprints.gla.ac.uk/69319>

Deposited on: 07 September 2012

# Neurons with Stereotyped and Rapid Responses Provide a Reference Frame for Relative Temporal Coding in Primate Auditory Cortex

Romain Brasselet,<sup>1,2</sup> Stefano Panzeri,<sup>2,3</sup> Nikos K. Logothetis,<sup>1,2,4</sup> and Christoph Kayser<sup>1,2</sup>

<sup>1</sup>Max Planck Institute for Biological Cybernetics, 72076 Tübingen, Germany, <sup>2</sup>Bernstein Center for Computational Neuroscience, 72076 Tübingen, Germany, <sup>3</sup>Robotics, Brain and Cognitive Sciences Department and Neuroscience and Cognitive Systems Centre, Italian Institute of Technology, 16163 Genova, Italy, and <sup>4</sup>Division of Imaging Science and Biomedical Engineering, University of Manchester, Manchester M13 9PT, United Kingdom

The precise timing of spikes of cortical neurons relative to stimulus onset carries substantial sensory information. To access this information the sensory systems would need to maintain an internal temporal reference that reflects the precise stimulus timing. Whether and how sensory systems implement such reference frames to decode time-dependent responses, however, remains debated. Studying the encoding of naturalistic sounds in primate (*Macaca mulatta*) auditory cortex we here investigate potential intrinsic references for decoding temporally precise information. Within the population of recorded neurons, we found one subset responding with stereotyped fast latencies that varied little across trials or stimuli, while the remaining neurons had stimulus-modulated responses with longer and variable latencies. Computational analysis demonstrated that the neurons with stereotyped short latencies constitute an effective temporal reference for relative coding. Using the response onset of a simultaneously recorded stereotyped neuron allowed decoding most of the stimulus information carried by onset latencies and the full spike train of stimulus-modulated neurons. Computational modeling showed that few tens of such stereotyped reference neurons suffice to recover nearly all information that would be available when decoding the same responses relative to the actual stimulus onset. These findings reveal an explicit neural signature of an intrinsic reference for decoding temporal response patterns in the auditory cortex of alert animals. Furthermore, they highlight a role for apparently unselective neurons as an early saliency signal that provides a temporal reference for extracting stimulus information from other neurons.

## Introduction

Dynamic natural stimuli are represented by time-varying patterns of neural activity in sensory cortices (Bair and Koch, 1996; Rieke et al., 1999; Victor, 2000). In auditory cortex, for example, ethological stimuli such as vocalizations or speech are encoded by precisely timed responses of individual neurons on the few tens of millisecond scale (Nelken et al., 2005; Chechik et al., 2006; Russ et al., 2008). Importantly, this information is considerably reduced when integrating the same responses over much longer time scales, e.g., when considering spike counts in windows of several tens of milliseconds (Schnupp et al., 2006; Engineer et al., 2008; Kayser et al., 2010). These results, together with the high information provided by response onset latencies about the sound frequency of tones or noises (Nelken et al., 2005; Bizley et al., 2010), support the importance of time-dependent responses

for auditory information processing. Still, it remains unclear how precisely timed responses are decoded within or across cortical populations (Soteropoulos and Baker, 2009; Sharpee et al., 2011).

Temporal response patterns are typically analyzed by aligning spikes and sensory events using a laboratory-based clock that registers the timing of individual stimuli and of neural events with supreme accuracy (Victor, 2000; Panzeri et al., 2010). The nervous system, however, does not have access to this artificial reference (Gollisch and Meister, 2008; Shusterman et al., 2011). It is therefore unclear how it succeeds in interpreting time-varying responses. When the sampling of the sensory stimulus is not initiated by an active movement or when the stimulus appears suddenly and unpredictably, sensory cortices cannot access independent estimates of stimulus timing based on a motor efference copy or some intrinsic stimulus regularity. In such cases, it has been suggested that the nervous systems must interpret time-varying responses using an intrinsically available reference, such as by encoding information in the relative timing of responses (deCharms and Zador, 2000; Furukawa et al., 2000; Chase and Young, 2007; Gollisch and Meister, 2008). Indeed, previous studies in anesthetized animals supported this hypothesis by reporting that auditory neurons in colliculus or cortex still carry information about spatial sound features when their responses are decoded using relative, rather than absolute, onset latencies

Received Oct. 27, 2011; revised Dec. 15, 2011; accepted Jan. 11, 2012.

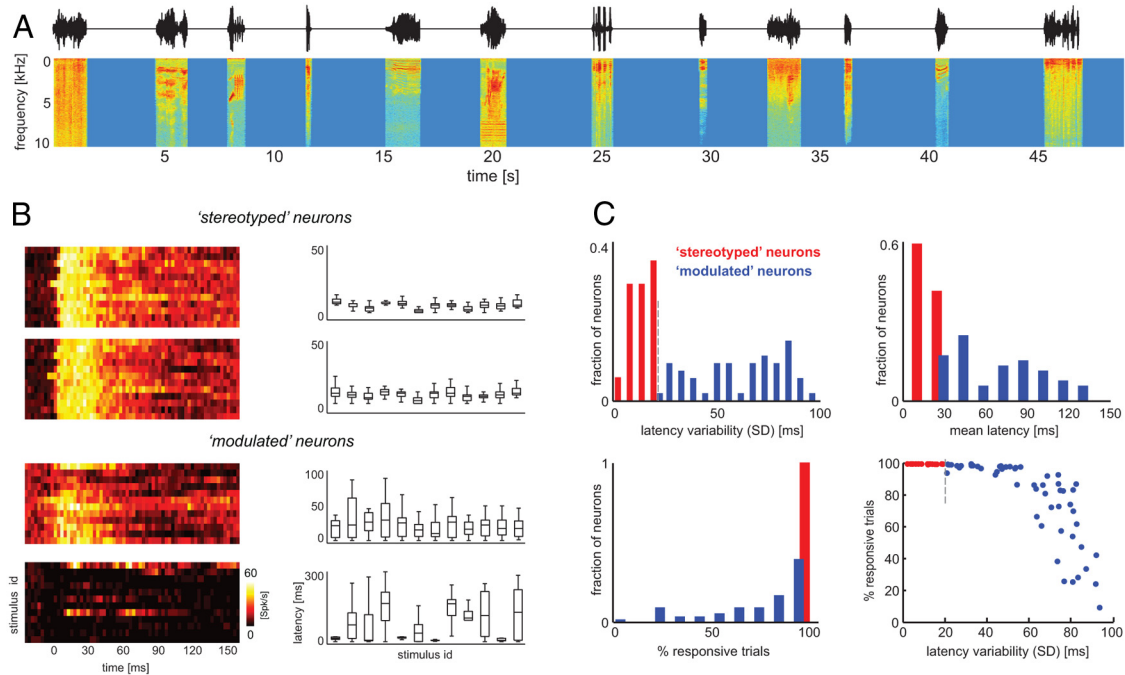
Author contributions: S.P. and C.K. designed research; R.B. and C.K. performed research; N.K.L. contributed unpublished reagents/analytic tools; R.B., S.P., N.K.L., and C.K. analyzed data; R.B., S.P., N.K.L., and C.K. wrote the paper.

This work was supported by the Max Planck Society, by the Neural Computation project of Italian Institute of Technology, by the Compagnia di San Paolo, and by the Bernstein Center for Computational Neuroscience Tübingen, funded by the German Federal Ministry of Education and Research (FKZ: 01GQ1002).

Correspondence should be addressed to Christoph Kayser, Max Planck Institute for Biological Cybernetics, Speemannstrasse 38, 72076 Tübingen, Germany. E-mail: christoph.kayser@tuebingen.mpg.de.

DOI:10.1523/JNEUROSCI.5435-11.2012

Copyright © 2012 the authors 0270-6474/12/322998-11\$15.00/0



**Figure 1.** Example data and neural response characteristics. **A**, Sound wave and spectrogram for the acoustic stimulus sequence on one example trial. Twelve natural sounds were presented in pseudo-random sequence and with random intersound intervals (blue periods) on each trial. The top shows the sound wave, the bottom the spectral representation (red colors indicate sound high amplitude). **B**, Left, Trial-averaged response time course (high response amplitude corresponds to bright colors) to all 12 stimuli (vertical axis) for two stereotyped and two modulated neurons. Time  $t = 0$  corresponds to sound onset. Right, Distribution of response onset latencies across trials for the same neurons and for each stimulus. Box plots display the median (central bar), 25th and 75th percentiles (box), and the full data range (bar). The stereotyped neurons responded to each stimulus, with short latencies and low trial-by-trial variability of the latency. **C**, Response characteristics for the entire population. These characteristics are the trial-by-trial variability of onset latency, the mean latency (averaged across trials and stimuli), and the fraction of nonresponsive trials (across all stimuli). The distributions are shown separately for stereotyped (red) and modulated (blue) neurons. These were (by definition) separated using a single criterion: a threshold applied to the latency variability (gray dashed line on latency variability), but segregate well also for the other response characteristics. Importantly, stereotyped neurons excel by short mean latency, low-latency variability, and responsiveness on most trials.

(Furukawa et al., 2000; Zohar et al., 2011). However, to understand how such relative coding schemes could be implemented as a general principle in cortex, it is necessary to identify an explicit reference signal that is sufficiently robust to allow the extraction of information about complex stimulus features, such as the identity of naturalistic sounds, even in the alert animal and without external predictive clues about stimulus timing.

Here we investigate the viability of a relative coding scheme in the auditory cortex of awake primates. We recorded neural responses to naturalistic sounds using a paradigm where predictive cues about stimulus appearance were minimized. We found that a subset of neurons responded rapidly, with highly reproducible latency and to all tested sounds. Using computational analysis we confirmed that these stereotyped neurons constitute an ideal reference for accessing the information carried by onset latencies and full spike trains of other, stimulus selective neurons.

## Materials and Methods

### Recording procedures, sensory stimuli, and data extraction

All procedures were approved by the local authorities (Regierungspräsidium Tübingen) and were in full compliance with the guidelines of the European Community (EUVD 86/609/EEC). Neural activity was recorded from the auditory cortex of two adult male rhesus monkeys (*Macaca mulatta*) using procedures detailed in previous studies (Kayser et al., 2009, 2010). Briefly, responses were recorded using multiple microelectrodes (1–6 M $\Omega$  impedance, 750  $\mu$ m spacing), high-pass filtered (4 Hz, digital two pole Butterworth filter), amplified (Alpha Omega system) and digitized at 20.83 kHz. Recordings were performed in a dark and anechoic booth while the animals were passively listening to the acoustic stimuli. Recording sites were located in primary auditory cortex

(field A1), as confirmed by frequency maps constructed for each animal and the responsiveness for tone versus band-passed stimuli (Kayser et al., 2009). Spike-sorted activity was extracted using commercial spike-sorting software (Plexon Offline Sorter) after high-pass filtering the raw signal at 500 Hz (third-order Butterworth filter). For the present study only units with high signal-to-noise ratio ( $>8$ ) and  $<2\%$  of spikes with interspike intervals shorter than 2 ms were included.

Acoustic stimuli (average 65 dB SPL) consisted of 12 naturalistic sounds (3 vocalizations of conspecifics, 9 vocalizations or noises of other animals, 0.3–1.5 s duration, 8 ms cosine-ramp) and were delivered from two calibrated free field speakers (JBL Professional) at 70 cm distance positioned  $45^\circ$  to the left and right of the head. Individual sounds were presented in pseudo-random order and were separated by silent inter-stimulus intervals of random duration (between 1.8 and 5 s; cf. Fig. 1A). The randomization of the interstimulus time was made to ensure that the stimulus paradigm does not provide any information about the timing of presentation of individual stimuli, does not create an expectation of presentation of the next stimulus within a restricted time interval, and cannot entrain auditory cortical rhythms to a temporal regularity in the stimulus sequence (cf. Lakatos et al., 2008; Jaramillo and Zador, 2011). All sounds were presented many times (on average about 50 repeats of the same stimulus, range 39–70 repeats). For the present purpose only the responses during the first 300 ms of each sound were analyzed.

### Quantification of single-trial response latencies and of response selectivity

Response latencies were defined for each individual stimulus presentation (trial) as follows. Spike trains were convolved with an exponentially decaying filter (3 ms time constant) to create an analog signal. The response onset was defined as the first point during stimulus presentation where this signal exceeded the 95% percentile of the distribution of its values attained during the prestimulus baseline period. On trials for

which no threshold crossing occurred (during the first 300 ms of stimulus) the neuron was considered nonresponsive. These single-trial latencies were later used for referencing spike times of other neurons recorded simultaneously in the same trial (see below). To characterize the statistical properties of response latencies we computed the following quantities from the distribution of single-trial latencies of each neuron: (1) the mean latency (across trials and stimuli), (2) the fraction of responsive trials in which a well defined latency could be detected, (3) the latency variability, computed as the SD of the latency over all responsive trials to each stimulus, and then averaged across stimuli. The selectivity of each neuron was determined using the “50% of maximal response” criterion (Tian et al., 2001; Remedios et al., 2009): for each stimulus we computed the average firing rate during the 300 ms window. From this we computed the fraction of stimuli for which this firing rate was larger than half the maximal of all these 12 responses.

#### Classification of neurons based on latency variability

These response characteristics varied considerably across cells (Fig. 1C). However, a subset of neurons excelled by very low trial-by-trial latency variability and by responsiveness on nearly every trial. For subsequent analysis we hence defined two groups of neurons using a single criterion: a threshold applied to the latency variability (threshold = 19.5 ms, dashed gray line in lower right, Fig. 1C). The group of neurons with latency variability smaller than this criterion was termed “stereotyped latency neurons” (briefly “stereotyped neurons”), as these neurons had very low latency variability (by definition) and short mean latencies. The remaining neurons were termed “modulated latency neurons” (briefly “modulated neurons”) because their response latency was (by definition) highly variable across trials, was longer, and was modulated by sound identity.

#### Calculation of stimulus information

We calculated the information carried by the responses of modulated neurons about which stimulus was currently presented (stimulus identity). We performed this calculation for each of  $n = 48$  modulated neurons for which we recorded at least one other modulated and one stereotyped neuron at the same time (usually on a different electrode). This allowed us to compare the information carried by the same single-trial responses when referenced to (1) the actual physical stimulus onset time, (2) the response onset of a stereotyped neuron, and (3) the response onset of a modulated neuron. We performed this analysis by considering the stimulus information carried by different putative neural codes, (1) the full spike pattern within a specific time window, and (2) the response onset latency. To perform the actual calculations of stimulus information we used different analytical techniques to obtain direct estimates of the stimulus information carried by spike sequences within short time windows, direct estimates of the stimulus information carried by onset latencies and we used a stimulus decoding approach to quantify the stimulus information provided by responses in progressively longer time windows. The results obtained from the different analytical methods and different putative neural codes were well consistent.

#### Information carried by spike trains estimated using the direct approach

**Information relative to the physical stimulus onset time.** We estimated the stimulus information carried by spike trains at a given poststimulus time by dividing these using sliding windows of length  $T$  ( $T$  ranging from 20 to 40 ms). Within each window we quantified the response  $r$  as binary 5-letter word: the time window was divided into five 4, 6 or 8 ms bins, and the letter (1/0) associated with each bin indicated the presence/absence of spike(s) within the respective bin (Strong et al., 1998; Kayser et al., 2009). Information between stimuli and the so defined neural responses was computed for each window  $T$  using Shannon’s formula:

$$I(S; R) = \sum_{r,s} P(s) P(r|s) \log_2 \frac{P(r|s)}{P(r)}, \quad (1)$$

with  $P(s)$  the probability of stimulus  $s$ ,  $P(r|s)$  the probability of the response  $r$  given presentation of stimulus  $s$ , and  $P(r)$  the probability of response  $r$  across all trials to any stimulus. To correct for the sampling bias, we used the so-called shuffling procedure to compute information from high-dimensional codes (Montemurro et al., 2007; Panzeri et al., 2007), combined with the quadratic extrapolation procedure (Strong et al., 1998). These calculations were performed using the “information break-down toolbox” for Matlab (<http://www.ibtb.org>) (Magri et al., 2009). For subsequent analysis, we obtained the “average” information per neuron, defined as the average of the stimulus information over the first 300 ms of stimulus presentation.

**Information relative to an intrinsic reference frame.** In the above calculation each single-trial spike train was aligned to the actual stimulus onset on the respective trial. To quantify the information carried by same spike trains when referenced relative to an internal reference, we proceeded as follows. Each spike train was realigned to one of two internal reference frames: (1) the response onset latency of a simultaneously recorded stereotyped neuron, or (2) the response onset latency of a simultaneously recorded modulated neuron. For alignment each single-trial spike train was shifted according to the single-trial latency of the respective reference neuron such that the response onset time of the reference neuron was considered as time  $t = 0$  (cf. Fig. 2). For trials on which the reference neuron was unresponsive, the actual single-trial spike train of the considered neuron was replaced by a random section of the neuron’s response extracted from the prestimulus baseline period, mimicking the fact that there is no well defined reference point for alignment in this case. After performing these shifts, the stimulus information in the re-referenced spike trains was calculated as above. For quantitative analysis we compared the information carried by each modulated neurons response when aligned relative to a stereotyped reference and when aligned to other modulated reference neurons (averaged over all modulated neurons available as reference for this neuron).

**Response features contributing to stimulus information.** The information provided by the spike train within window  $T$  could be carried by either stimulus variations of the time-dependent firing rate or by correlations between spike patterns (expressed, for example, as combinations of spikes in different bins occurring with a distribution that cannot be explained only in terms of firing rate variations) (Perkel and Bullock, 1968; Panzeri et al., 2010). We obtained an estimate of the contribution of the time-dependent firing rate using the construct of  $I_{\text{PSTH}}$  (Montemurro et al., 2007). This reflects the information carried by a hypothetical neuron with the same time-dependent firing rate as the considered one, but whose spike train has no additional correlations between spike times other than those arising from the time-dependent rate:

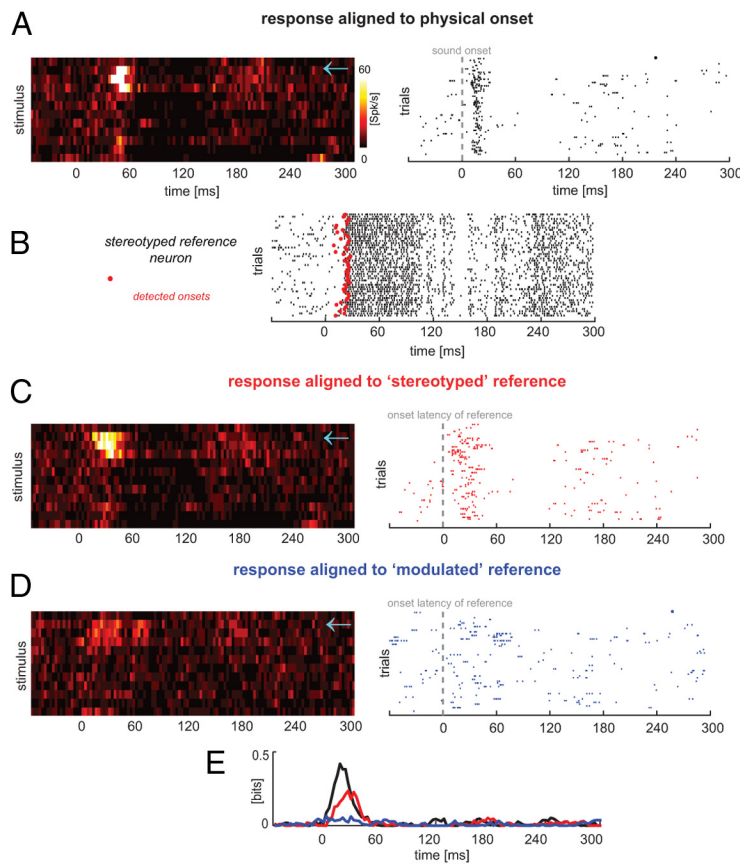
$$I_{\text{PSTH}} = \sum_{r,s} P(s) P_{\text{ind}}(r|s) \log_2 \frac{P_{\text{ind}}(r|s)}{P_{\text{ind}}(r)}. \quad (2)$$

Here  $P_{\text{ind}}(r|s)$  is the probability of obtaining response  $r$  to stimulus  $s$  of a Poisson neuron with the same time-dependent rate as the one measured experimentally, and  $P_{\text{ind}}(r)$  is the average across stimuli of  $P_{\text{ind}}(r|s)$  weighted by  $P(s)$ .  $I_{\text{PSTH}}$  was computed with the same bias corrections as used in Equation 1. The comparison between  $I_{\text{PSTH}}$  and  $I(S;R)$  provides a direct assessment of the additional information carried by spike pattern correlations. In agreement with previously published results (Kayser et al., 2010) we found the contributions of correlations to be small.

#### Information carried by spike trains estimated using a decoding approach

By using direct information estimates the above analysis includes an assessment of arbitrary and possibly nonlinear stimulus–response relations (Quian Quiroga and Panzeri, 2009; Panzeri et al., 2010). However, given the technical difficulty of obtaining reliable direct information estimates using high-dimensional response variables this approach is limited to responses in short windows (i.e., few time bins). We performed a separate analysis using a framework of stimulus decoding that al-





**Figure 2.** Example data from one modulated example neuron showing responses aligned using different reference frames. **A**, Responses of the example neuron when each single-trial response was aligned to the actual stimulus onset time. Trial-averaged responses for each stimulus are shown in the color-coded panel on the left. The single-trial spike rasters for one stimulus (indicated by blue arrow in left panel) are shown in the panel on the right. Time  $t = 0$  corresponds to stimulus onset. **B**, Single-trial spike rasters for one stimulus of the stereotyped reference neuron which is used to align the response of the example neuron in **C**. The detected single-trial response onsets of the stereotyped reference neuron are indicated by the red dots. Time  $t = 0$  corresponds to the stereotyped reference neuron. Trial averaged responses for each stimulus are shown on the left and single-trial spike rasters for one stimulus are shown on the right. Note that  $t = 0$  corresponds to the reference time point, i.e., the onset latency of the reference neuron on each trial. As a result, the apparent response latency differs from that visible when the same responses are aligned to the actual stimulus onset (cf. **A**). **D**, Response of the modulated example neuron when each single-trial response was aligned relative to a modulated reference neuron. When using this as reference, the stimulus-evoked response becomes much weaker, both in the trial averaged responses (left) and the spike raster (right). **E**, Time course of the stimulus information in spike trains on a short time scale for this example neuron calculated using each reference condition. Information was calculated using the direct approach and using sliding windows of 4 ms bins.

lowed us to consider longer time windows and to obtain estimates of cumulative stimulus information.

Specifically, we used a linear discriminant decoder in conjunction with a leave-one-out cross-validation procedure (Nelken and Chechik, 2007; Russ et al., 2008; Kayser et al., 2010). For each individual trial of a given stimulus ( $s_i$ ), this proceeded as follows. (1) The average responses to all other 11 stimuli were computed by averaging the responses of all repeats of the respective stimuli. (2) For the current stimulus ( $s_i$ ) the mean response was computed by averaging across all trials, excluding the current “test” trial. The thereby obtained average responses represent the “codebook.” (3) The Euclidean distance (over time points) was computed between the response on the test trial and the average responses in the codebook, and the test trial was decoded as that stimulus yielding the minimal distance to the test response. This procedure was repeated for each trial of each of the 12 stimuli providing the total percentage of correctly decoded trials and the confusion matrix. The values on a given row  $s$  and column  $d$  of the confusion matrix  $Q(d|s)$  represent the fraction of trials on which the presented stimulus  $s$  was decoded to be stimulus  $d$ . If decoding were perfect, the values in  $Q$  would be one on the

diagonal and zero otherwise. A measure of mutual information between stimulus and response can be derived from the confusion matrix using the following formula (Victor and Purpura, 1996; Quiñones Quiroga and Panzeri, 2009):

$$I(S; D) = \sum_{d,s} P(s)Q(d|s) \log_2 \frac{Q(d|s)}{Q(d)} \tag{3}$$

While even for an optimal decoder the information in the confusion matrix (Eq. 3) may be less than the total information available in the response, its computation is more data-robust than the one based on the direct method. Therefore Eq. (3) offers the possibility to characterize selectivity of spike trains over longer time windows. We used this method to evaluate the cumulative information available after observing the response in windows of increasing duration from stimulus onset (see Fig. 4E). Specifically, we computed the information for response epochs of different duration following stimulus onset, using data chunks sampled from 100 ms before stimulus onset up to time point  $t$ , and systematically varying  $t$  from 0 up to 300 ms following stimulus onset.

As for the above described direct method, we computed information in the confusion matrix of the decoder using different referencing frames, by aligning spike trains to the actual stimulus onset or to the latency of other simultaneously recorded neurons exactly as described in Information relative to an intrinsic reference frame, above.

*Information carried by response onset latencies*

Information in the response onset latency was calculated by binning the latency values for each neuron into 6 equi-populated bins, and by calculating the mutual information between these values and the stimulus identity using direct estimates of Shannon mutual information (Eq. 1). The information in the onset latency was also computed either when latency was measured relative to the stimulus time or relative to the latency of another neuron, as described in above sections.

*Dependence of information upon the effective temporal precision of the reference frame*

The amount of information that could be recovered from responses aligned to other reference neurons rather than stimulus onset was reduced mostly because the reference neuron provides an imperfect representation of stimulus timing. To better interpret the information loss caused by using an intrinsic reference, it is useful to compare this to the information loss induced by directly degrading the precision with which single-trial responses are aligned to stimulus onset (see Fig. 5B). We estimated this decrease in information parametrically by adding independent Gaussian random jitter with zero mean and SD  $J$  to the actual stimulus onset time of each trial. We obtained the stimulus information as a function  $J$  for each neuron by averaging 10 repeats for each value of  $J$ , and we expressed the fraction of the information at each value of  $J$  relative to the information without added temporal jitter ( $J = 0$ ). The result was averaged across neurons and confidence intervals for the population average were obtained using a jackknife procedure: we calculated a distribution of values by systematically discarding individual neurons

and derived the confidence interval associated with this jackknife distribution (Sokal and Rohlf, 1995). For later use we fit the resulting curve using an exponential (see Fig. 5B, blue line):  $y = 0.35 + 0.65 * \exp[-J/21]$ .

#### Using a population of neurons as intrinsic reference

We performed additional computational analysis to quantify whether and how much more information could be extracted when using a population of  $N$  stereotyped neurons as reference, as opposed to using a single reference neuron. To this end we assumed that the response latencies of reference neurons are distributed according to a multivariate Gaussian distribution, and that the latency of the aggregate population response is obtained as the mean latency across all neurons. Following previous theoretical work the variability of the population latency ( $\sigma_{\text{pop}}$ ) can be derived analytically as follows, given a population of neurons with known trial-to-trial latency variability  $\sigma$  and known latency covariance  $c$  between pairs of neurons (Abbott and Dayan, 1999):

$$\sigma_{\text{pop}}^2 = \frac{\sigma^2}{N}[1 + c(N - 1)]. \quad (4)$$

The resulting values of  $\sigma_{\text{pop}}$  are shown in Figure 4C (see below) for a range of values for  $\sigma$  and  $c$ . To obtain a self-consistent estimate of  $\sigma_{\text{pop}}$  with a population of  $N$  reference neurons, we estimated the parameters  $\sigma$  and  $c$  from the actual data as follows. We derived the latency variability across neurons using the dependency of the recovered information on the precision with which the stimulus onset is registered (see Fig. 5B): using a single stereotyped neuron as reference, one can recover on average 86% of the information available relative to actual stimulus onset, which corresponds to an effective precision of  $\sigma_{\text{eff}} = 5.1$  ms (see Fig. 5B, orange line). The latency covariance was estimated from the observed covariance of 8 pairs of simultaneously recorded stereotyped neurons across 12 stimuli as the median (robust estimate) across these 96 ( $8 \times 12$ ) values ( $c = 0.042$ ). To convert the variability of the population latency ( $\sigma_{\text{pop}}$ ) into estimates of information loss, we again exploited the dependency of stimulus information on the temporal precision of the reference system: the exponential fit to Figure 5B (see below) was used to relate  $\sigma_{\text{pop}}$  to the fraction of preserved information. Finally, the inversion of this relationship and the inversion of Eq. 4 allowed us to determine the minimal population size  $N$  required to reach an information value that exceeds 95% of the information available when referencing responses relative to the physical stimulus onset (see Fig. 5D).

## Results

We recorded the responses of  $n = 70$  neurons from primary auditory cortex during the presentation of natural sounds (con-specific vocalizations, vocalizations or noises of other animals) presented at unpredictable times. Within this population some neurons responded with short and highly reproducible latency to each sound (Fig. 1B, top examples), while other neurons were more selective, responded with variable latencies only to some stimuli and not in all trials (bottom examples).

We characterized the responsiveness and single-trial latencies of these neurons by means of several quantitative metrics (see Materials and Methods): the mean latency across trials and stimuli; the fraction of trials for which a well defined latency could be detected; the latency variability, defined as the SD of the latency on all responsive trials (and averaged across stimuli); and an index of stimulus selectivity. These response metrics varied markedly from neuron to neuron and revealed a subset of neurons responding with short and little-varying latencies on almost all repeats of the stimuli. For the present analysis we selected neurons with such highly stereotyped responses by partitioning the population into two groups using a single criterion. We applied a threshold to the latency variability (threshold = 19.5 ms, dashed gray line in lower right, Fig. 1C) and termed one resulting

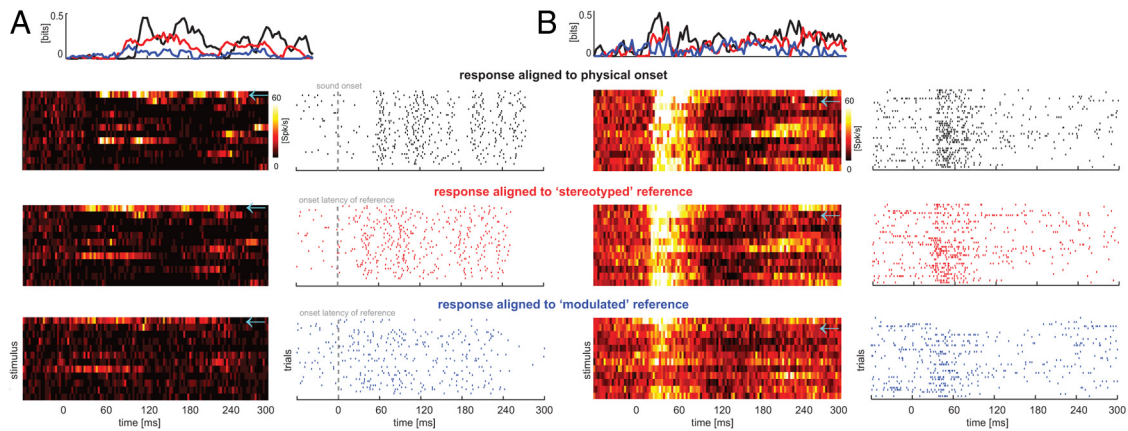
subpopulation stereotyped latency (briefly “stereotyped”) neurons and one subpopulation modulated latency (“modulated”) neurons. These terms reflect the observation that neurons within the first group had similar onset latencies across stimuli and responded to all tested stimuli (Fig. 1C, red), while latencies of neurons in the second group were modulated by stimulus identity and varied across repeats of the same stimulus as well as across stimuli (Fig. 1C, blue).

A quantitative comparison of these response metrics revealed several differences between neurons in the two subpopulations, despite them being partitioned based only upon a single criterion (single-trial latency variability). Stereotyped neurons (17 of 70, 24%) responded with short mean latencies ( $21.7 \pm 0.8$  ms), had a well identifiable response latency on almost all trials (>96%, mean 99%), and responded significantly to almost all stimuli ( $11.5 \pm 0.2$  of 12, mean  $\pm$  SEM). By definition, these neurons had low trial-by-trial latency variability ( $11.9 \pm 1.4$  ms). Modulated neurons (53 of 70, 76%), in contrast, responded with longer mean latencies ( $72.0 \pm 4.6$  ms; two-sample  $t$  test,  $p < 10^{-7}$ ), only on a subset of trials ( $78.6 \pm 3.4\%$  trials,  $p < 10^{-3}$ ), responded only to some stimuli ( $9.8 \pm 0.4$ ,  $p < 0.05$ ), and by definition responded with variable onset latencies ( $59.5 \pm 3.0$  ms;  $p < 10^{-6}$ ). In addition, stereotyped neurons also responded with significantly higher firing rates, creating a strong population response ( $28.2 \pm 2.6$  vs  $7.6 \pm 1.0$  spikes during the first 300 ms, two-sample  $t$  test  $p \approx 0$ ).

The stereotyped neurons clearly stand out from the entire population because of their rapid, reliable and comparatively strong responses (see Fig. 5A). Their rapid and unspecific responses make these stereotyped neurons natural candidates to form an intrinsic temporal reference frame, relative to which the latencies and the time-varying responses of other stimulus modulated neurons could be interpreted. In the following we test this hypothesis quantitatively using methods of information theory and stimulus decoding.

#### Response patterns relative to an intrinsic reference

We calculated the information about the identity of the presented stimulus carried by modulated neurons when their responses were referenced (aligned) either to the single-trial response onset of a simultaneously recorded stereotyped neuron or to the single-trial response of another modulated neuron. We compared these values to the information available from the same responses when referenced to the actual stimulus onset to evaluate the efficiency of the two internal references in preserving stimulus information. Figure 2 displays the responses of one example neuron when aligned to each of these three references. Examination of the trial-averaged responses aligned to stimulus onset (Fig. 2A, left) shows that these are stimulus specific and time dependent. In particular, stimulus-specific episodes of high mean firing rates are localized in time periods of few tens to hundreds of milliseconds poststimulus and correspond to epochs during which spike-rasters reveal a reliable response across stimulus repeats (black rasters, Fig. 2A, right). These temporal response patterns can only be correctly interpreted by downstream decoders that have some knowledge of the poststimulus time at which the response was emitted. For example, a downstream decoder could determine whether a low firing rate was elicited by a noneffective stimulus or by an effective stimulus at a non-optimal poststimulus time only if it had some information about the time of the current response with respect to the stimulus onset. Similar stimulus-specific response patterns are visible in the responses of two additional example neurons shown in Figure 3.



**Figure 3.** Additional example data showing responses aligned using different reference frames. *A, B*, Responses of one example neuron each. The uppermost panels display the time course of the stimulus information for this example neuron calculated from the full spike train using each reference condition. Subsequent panels display the trial averaged responses for each stimulus (left) and the single-trial spike rasters for one stimulus indicated by a blue arrow (right) for each type of reference frame. Note that for the sound onset reference time  $t = 0$  corresponds to the actual stimulus onset time, while for the intrinsic references  $t = 0$  corresponds to the reference time point, i.e., the onset latency of the reference neuron on each trial. Conventions otherwise as in Figure 2.

To gain some intuition about how the nervous system may form an intrinsic time frame for decoding time-varying responses, it is useful to visualize the spike rasters when referenced to an internal reference frame provided by the firing of another neuron. For the example case the single-trial responses and the detected onset latencies for a stereotyped neuron are shown for one stimulus in Figure 2*B*. When the single-trial spike trains of the investigated neuron are aligned relative to the onset latencies of this reference neuron the resulting responses (Fig. 2*C*; trial-averaged responses, left, and single-trial rasters in red, right) are still stimulus-selective, temporally modulated and resemble the responses as aligned to stimulus onset. Thus, the use of a stereotyped neuron as reference qualitatively preserves the stimulus selectivity and temporal profile of the response (compare black and red rasters in Fig. 2*A, C*). However, when the same responses are aligned relative to the response onset latencies derived from a simultaneously recorded modulated neuron, the structured response patterns disappear (blue rasters, Fig. 2*D*). Together with the additional examples (Fig. 3) this suggests that aligning spike trains of a considered neuron relative to the response of another stereotyped neuron largely preserves the stimulus related response modulation, an important prerequisite for the internal readout of the stimulus-specific information carried by a neurons response.

### Information in spike trains on a short time scale

We used information theoretic analyses to quantify how much the stimulus information carried by these neurons is affected by the choice of reference frame. Figure 2*E* (see also Fig. 3) displays the stimulus information obtained from the spike trains of the example neurons using a sliding window analysis. Noteworthy, the information values are considerably higher when using a stereotyped neuron compared with a modulated neuron as reference, and they fall only little short of the values attained when responses are referenced relative to the actual sound onset time.

Analysis of the entire population confirmed that information values were larger when using a stereotyped rather than a modulated reference. For each neuron we calculated the average information in spike patterns (averaged over all sliding window positions along the first 300 ms of the response) and we then compared these values when using different types of reference neurons. Across the population ( $n = 48$  modulated neurons that

could be paired with a simultaneously recorded stereotyped and a modulated reference neuron) information relative to the stereotyped reference was significantly higher than relative to the modulated reference (median values: 0.07 vs 0.05 bits, sign-rank test  $p < 10^{-6}$ , Fig. 4*A*). Of these 48 neurons, 44 (92%) carried more information relative to the stereotyped rather than to the modulated reference. For comparison, information in the same neurons when referenced to the actual stimulus onset time was 0.08 bits, significantly higher than relative to the stereotyped neurons (Fig. 4*A*, inset;  $p < 10^{-6}$ ). Noteworthy, the prevalence of higher information obtained using the stereotyped reference extended over the entire response time, as shown by the population averaged information time course (Fig. 4*B*). To evaluate how effective intrinsic reference frames are with respect to preserving knowledge about the precise stimulus onset, we expressed the information values obtained using the intrinsic references relative to the information carried by same responses when aligned to stimulus onset (Fig. 4*C*). We found that the stereotyped reference preserved 86% (median value) while the modulated reference preserved only 66% of the information available relative to the actual stimulus onset.

We verified that these results do not depend on the specific parameters used to characterize the responses as binned  $n$ -spike patterns. We repeated the analysis using bins of 4, 6 and 8 ms duration (resulting in time windows  $T$  of 20, 30 and 40 ms respectively) and we obtained very comparable results (Fig. 4*C*). Importantly, the fraction of information preserved by the stereotyped reference was largely independent on the specific parameter choice and was larger than when using the modulated reference for all choices of bins (sign-rank tests,  $p < 10^{-6}$ ).

The above demonstrates that the responses of auditory cortex neurons remain highly stimulus informative, even when analyzed relative to a temporal reference frame provided by the subset of stereotyped neurons. However, it remains unclear what aspects of neural response actually carry this information. In particular, when considering binary  $n$ -spike patterns information can be provided by variations in the overall spike count within the time window  $T$ , by temporal variations in the firing rate (spike density) within this window, or by higher-order correlations of spikes within the  $n$ -spike pattern. Previous studies have demonstrated that auditory cortex neurons encode information about naturalistic sounds by patterns of activity on the time scale of few

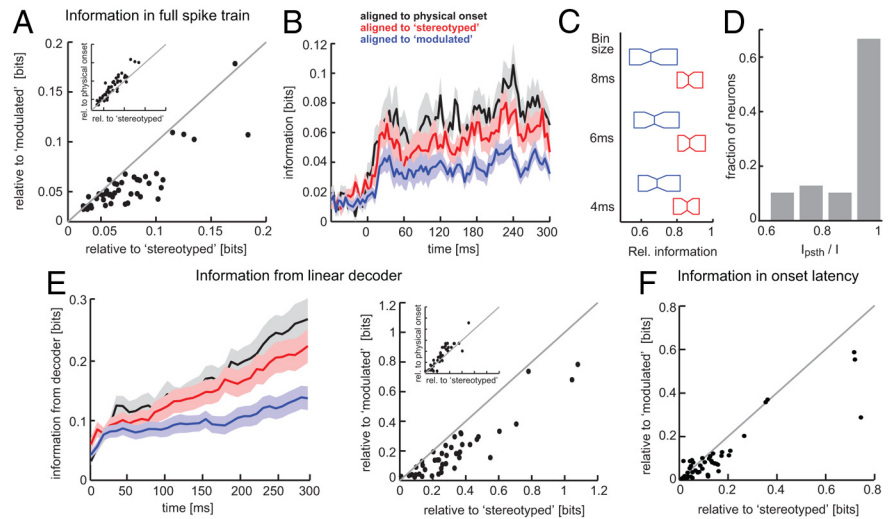


(or few tens of) milliseconds (Engineer et al., 2008) and that most of this information is carried by short-time variations of firing rate (spike density) rather than higher-order correlations of spikes (Kayser et al., 2010). We confirmed this dominance of firing rate variations in carrying information for the present dataset. Specifically, we compared the information carried by the actual spike train ( $I(S;R)$ , Eq. 1, Materials and Methods) to the information provided by the spike train of a hypothetical Poisson neuron with the same time-dependent rate as the one under analysis ( $I_{PSTH}(S;R)$ , Eq. 2, Materials and Methods). Note that for this analysis we analyzed only responses of modulated neurons when aligned to stimulus onset. Across neurons  $I_{PSTH}$  accounted for the largest fraction of the total information for most of the neurons. For example, when using 4 ms bins the median ratio was 0.94 (Fig. 4D), and similar values were obtained when using longer time bins (e.g., median ratio of 0.92 for 8 ms bins). These findings are in good agreement with previous work (Kayser et al., 2010). We conclude that stereotyped neurons provide an internal temporal reference frame that is sufficiently precise to allow recovering the temporal variations of modulated neurons' firing rates that carry sensory information.

### Cumulative information in spike trains

We performed additional analysis to compare the performance of different reference frames when stimulus information is characterized on longer time scales. While the above analysis focused on spike trains on the short time scale (up to 40 ms) and used direct estimates of stimulus information, in an additional analysis we used a stimulus decoding framework to estimate the stimulus information from responses in longer time windows. Specifically, we estimated the cumulative information provided by spike trains between stimulus onset and a specific time point later during the response.

This confirmed the above result that stereotyped (but not modulated) neurons provide a suitable intrinsic reference frame. The overall stimulus information increased with increasing length of the considered response epoch, reflecting the accumulation of stimulus information over time (Fig. 4E). In addition, the benefit of using stereotyped (over modulated) neurons as a reference became larger when considering progressively longer time windows. For example, when using the full 300 ms window stimulus information was significantly higher when using stereotyped rather than modulated neurons as reference (median 0.25 vs 0.13 bits; sign-rank test  $p < 10^{-7}$ ; Fig. 4E). For comparison, information relative to the actual onset was 0.3 bits (median,  $p < 10^{-3}$ ). When calculated relative to the information available in the same responses when aligned to actual stimulus onset the stereotyped reference preserved 84% (median) while the modulated reference preserved only 52% of the available information (sign-rank test,  $p < 10^{-6}$ ).



**Figure 4.** Population analysis of stimulus information using different reference frames. **A**, Average information in the full spike train for each modulated neuron, when referenced to either a stereotyped another modulated neuron (large panel), and when referenced to the physical sound onset (small inset). Each dot corresponds to one neuron, information values were computed using the direct method, using 4 ms bins, and were averaged over all sliding window ( $T = 20$  ms duration) positions. **B**, Time course of the stimulus information following stimulus (or reference neuron) onset (each aligned to  $t = 0$ ) when using each reference frame. Lines denote the mean across all neurons ( $n = 48$ ), shaded areas the standard error (SEM). Information was computed using the direct method, using 4 ms bins, and were averaged over all sliding window ( $T = 20$  ms duration) positions. **C**, Fraction of stimulus information preserved by each type of reference (red, stereotyped; blue, modulated neurons) when stimulus information was computed using different bin sizes. The relative information was computed as the ratio of the information available using both types of reference neuron to the information available relative to the physical stimulus onset. Box plots display the median (central bar), 25th and 75th percentiles (box). **D**, Distribution (across the population of all modulated neurons when aligned to stimulus onset) of the ratio between the information provided by the spike train of a hypothetical Poisson neuron with the same time-dependent rate as the one under analysis ( $I_{PSTH}(S;R)$ , Eq. 2) and the information carried by the actual spike train ( $I(S;R)$ , Eq. 1).  $I_{PSTH}$  accounted for a large fraction of the total information for most of the neurons. Information was calculated using 4 ms bins ( $T = 20$  ms windows). **E**, Stimulus information obtained from a linear decoder applied to response epochs of increasing duration. Left, Information obtained when using the response in progressively longer time windows for decoding (starting at  $t = 0$  and ending at each indicated time point). Lines denote the mean across neurons, shaded areas the standard error (SEM). Right, Information obtained when using the full 300 ms of the response for each modulated latency neuron, when referenced to either a stereotyped or another modulated neuron (large panel), and when referenced to the physical sound onset (small inset). Each dot corresponds to one neuron. **F**, Stimulus information in the response onset latency for each neuron, when referenced to either a stereotyped or a modulated reference neuron. Information was calculated using the direct approach and by considering only the onset latency as response variable.

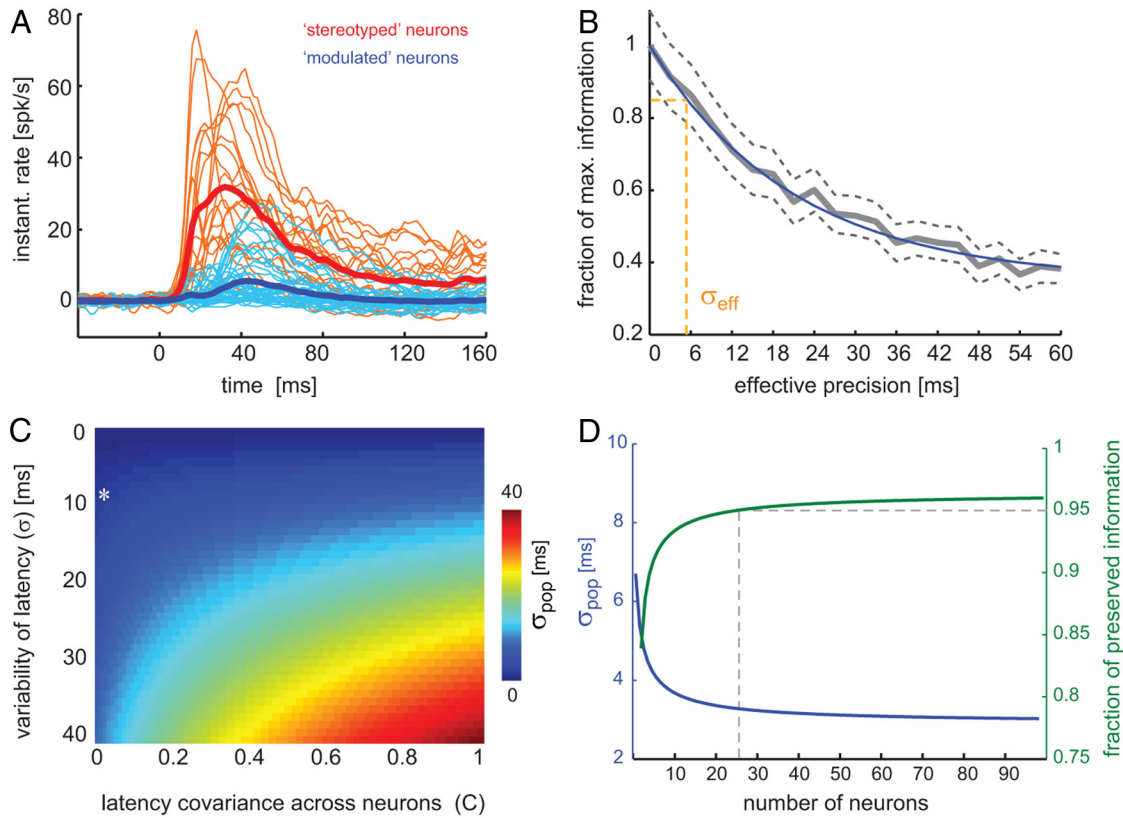
### Information in response onset latencies

In addition to the information provided by the full spike train, we also calculated the information provided by the response onset latency of each neuron. Onset latencies of auditory cortical neurons are known to carry information about sounds features such as spatial location or pitch and are considered as a rapid and potentially valuable code for auditory processing (Nelken et al., 2005; Chechik et al., 2006; Bizley et al., 2010). We found that the information carried by single-trial onset latencies was higher when referenced to a stereotyped rather than a modulated reference (median values 0.10 vs 0.07 bits, sign-rank test,  $p < 10^{-4}$ ; Fig. 4F). When expressed to the information available relative to the actual stimulus onset, information in onset latencies relative to stereotyped neurons reached 91%, relative to modulated neurons only 63%. This demonstrates that stereotyped neurons can act as an effective internal reference frame for decoding in the context of several putative codes.

### Population responses as intrinsic reference

The above analysis focuses on the information available in neural responses when referenced to the onset time of a simultaneously recorded individual reference neuron. Although referencing to single stereotyped neurons was effective, it seems conceivable





**Figure 5.** Analysis of the effective precision and information provided by a population of reference neurons. **A**, Population response showing the average response time course for stereotyped (red) and modulated (blue) neurons around stimulus onset ( $t = 0$ ). Thick lines denote the mean across stereotyped and modulated neurons. **B**, Fraction of stimulus information preserved when reducing the precision about the knowledge of the actual stimulus onset time. For this analysis, a Gaussian jitter of increasing variance was added to the single-trial stimulus onset times and the graph displays the ratio of the information at each jitter value (“effective precision”) relative to the information obtained with zero jitter (dashed lines: SEM from a jackknife procedure). The decrease of information with increasing jitter is well fit by an exponential (blue line). Using individual stereotyped neurons as reference allowed recovering 86% of the information available relative to stimulus onset, which corresponds to a latency variability ( $\sigma_{\text{eff}}$ ) of 5.1 ms (orange line). **C, D**, Estimate of the minimal size of a population of stereotyped neurons required to achieve a given effective precision in the estimate of the population response onset. The exponential fit in **B** was used to relate stimulus information to the latency variability and a population model was used to derive the variability of the summed population response ( $\sigma_{\text{pop}}$ ) for a given single-neuron latency variability  $\sigma$ , latency covariance between pairs of neurons  $c$ , and population size  $N$  (Eq. 4). **C** displays the variability of the summed population response as a function of  $\sigma$  and  $c$ , with the experimentally observed values marked by \*. **D** displays the decrease of  $\sigma_{\text{pop}}$  with increasing population size  $N$  and the corresponding increasingly preserved fraction of stimulus information (green curve). The latter reveals that a population size of  $N = 25$  neurons is sufficient to recover at least 95% of the information available when using the physical sound as reference (dashed line).

that the nervous system may use an aggregate population signal as reference, such as for example a pooled population response (Chase and Young, 2007; Panzeri and Diamond, 2010). Such a population signal is not only relatively easy to evaluate and access by cortical microcircuits, but in addition may provide a timing signal more precise and reliable than that provided by the latencies of individual reference neurons. Indeed, considering that the stereotyped neurons are the first to respond from the entire population, pooling the responses across stereotyped neurons directly corresponds to the aggregate response of the entire sampled auditory cortex population during the first tens of milliseconds following the occurrence of a stimulus (Fig. 5A). We hence performed a population-based analysis to investigate whether using a larger population of stereotyped neurons as reference would increase the amount of extractable information. To this end we considered a population of stereotyped neurons assuming their latencies were distributed according to a multivariate Gaussian distribution with trial-to-trial covariance matching that of the actually recorded neurons. From this we then obtained the latency variability of a modeled population of varying population size  $N$  following previous theoretical studies (Abbott and Dayan, 1999).

We first determined the relation between the amount of recovered stimulus information and the effective temporal preci-

sion (latency variability) of a presumed intrinsic reference. By adding Gaussian noise to the physical stimulus onset time we obtained the dependency of stimulus information provided by the full spike train of each neuron on the precise temporal alignment of each trial to stimulus onset (Fig. 5B). Using the above result that stereotyped neurons can recover 86% of the information available relative to stimulus onset this provided an equivalent temporal jitter of 5.1 ms. This equivalent jitter was used together with the measured trial-by-trial latency covariation between simultaneously recorded stereotyped neurons as parameters to obtain the effective jitter from a presumed population of size  $N$  (see Materials and Methods; Eq. 4).

The dependency of the latency variability of the modeled population ( $\sigma_{\text{pop}}$ ) on the values of the single-neuron variability ( $\sigma$ ) and the latency covariance is shown in Figure 5C. To convert the variability of the population latency estimates into estimates of information loss, we again used the measured dependency of stimulus information on effective reference precision: we used the exponential curve fit to the data in Figure 5B to relate effective precision ( $\sigma_{\text{pop}}$ ) to the fraction of preserved information. Performing this calculation using the experimentally observed values for  $\sigma$  and  $c$  for a range of population sizes (Fig. 5D) we found that a population of 25 stereotyped neurons provided an estimate

of stimulus onset that is sufficiently precise to recover at least 95% of the stimulus information available when using the physical stimulus onset as reference (dashed line; Fig. 5D). This illustrates the efficiency of small sets of stereotyped neurons as intrinsic reference frame of high temporal precision that allows the formation of highly informative relative coding schemes.

## Discussion

We recognize natural sounds such as animal noises in a forest or the call of our name despite their unpredictable occurrence. Natural sounds vary on multiple time scales, especially fast ones, and neurons in auditory areas represent them by finely timed responses (Liu et al., 2006; Engineer et al., 2008; Wang et al., 2008; Kayser et al., 2010). Our ability to detect and recognize these sounds suggests that the auditory system features mechanisms to extract information from precisely timed responses, such as those found in auditory cortex (Yang et al., 2008; Sharpee et al., 2011). What these mechanisms are and how they are implemented in neural populations remains a matter of debate.

### The problem of reference frames

Typical experimental analyses of time-dependent responses align single-trial spike trains to stimulus onset, and thereby exploit a priori knowledge about experimental design and timing. The auditory system, however, has to extract information from time-varying responses without access to precise and independent knowledge of stimulus timing. Instead, the auditory system must rely on some internal temporal reference frame and previous work suggested the relative timing of responses to a population-defined event as one potential reference (Reich et al., 2000; Gollisch and Meister, 2008). Such relative timing does not rely on external evidence about stimulus occurrence, and studies on the representation of spatial acoustic cues in midbrain and cortex provided evidence for the feasibility of such relative coding schemes, at least in the context of simplistic stimuli (tone or noise bursts) and the anesthetized state (Furukawa et al., 2000; Chase and Young, 2007; Zohar et al., 2011). Here, we made significant progress by demonstrating that relative coding schemes can operate efficiently in the alert animal for decoding latencies and sustained time-varying responses to natural sounds. Our work enhances previous insights about the use of relative temporal references by demonstrating an explicit neural signature of stimulus time in the auditory cortex of alert animals, and by proving its effectiveness in the decoding of behaviorally relevant sounds such as communication signals.

Our findings highlight a role for neurons with rapid and stereotyped responses to various natural stimuli, which may easily go unnoticed in typical experiments due to lack of sought-after specificity. Such rapid and apparently unselective neurons may provide a saliency signal that constitutes a reliable reference for decoding information from the responses of other more selective neurons. In our data, the stereotyped neurons responded first and with stronger amplitude than selective neurons, hence contributing most to the initial population response. Our modeling analysis demonstrated that about two dozen reference neurons are sufficient to create a population signature with sufficiently high precision to recover nearly as much information from relative responses as could be obtained from the original responses aligned to stimulus onset. It will be worthwhile to investigate whether similarly stereotyped neurons that qualify as intrinsic reference signal exist in other sensory cortices or modalities.

The division we made between “reference” and “encoding” neurons is not meant to refer to a strict dichotomy. Rather, our

data suggest that auditory cortex neurons form a continuum, at one end of which are neurons responding selectively and with variable latencies and at the other end are neurons with stereotyped, rapid and highly reliable responses (cf. Fig. 1C). The latter evidently excel in the aggregate population response (compare Fig. 5A) and collectively provide an early saliency signal that may serve as intrinsic reference.

### Putative coding schemes in auditory cortex and stereotyped neurons as reference

Auditory cortex neurons encode naturalistic sounds using rapid variations of firing and to extract most information from their responses decoding mechanisms need to read these responses using a precision of a few to a few tens of milliseconds (Schnupp et al., 2006; Engineer et al., 2008; Kayser et al., 2010). Previous work has demonstrated that the exact time scale of neural coding precision depends on the precise stimulus context and that the vast majority of stimulus information carried by auditory cortical spike trains results from rapid variations of firing rate and not from higher-order patterns within the spike train (Kayser et al., 2010). Our analysis confirmed this dominance of rapid modulations of firing rates in carrying stimulus information and compared different reference frames for decoding such informative temporal response patterns. In the following we discuss the implications of our findings at the light of previous proposals for candidate decoding mechanisms of time-varying responses.

One possibility is that sensory systems rely on an explicit internal representation of stimulus timing. Information about temporal stimulus structure is crucial for sensation and behavior, and this is particularly true for the auditory system (Heil and Irvine, 1997; Schnupp et al., 2006). An internal explicit representation of stimulus timing by a short latency population response as reported here is appealing, as it is rapidly available following stimulus occurrence and precedes subsequent neural events that carry stimulus-specific information. Our results assign a representation of stimulus time to neurons with short and stereotyped latencies while at the same time imputing the representation of stimulus identity to neurons with longer latency and time-varying responses. Such separated instantiation of saliency signal and stimulus representation differs from previous proposals that attributed stimulus information to the latencies of all neurons while considering separate subsets of neurons for the encoding of distinct stimulus features. In the rat somatosensory cortex, for example, each potential object location is encoded by the latency of a specific population, and the stimulation time of each location can be estimated from the pooled activity of local populations (Panzeri et al., 2001; Foffani et al., 2008; Panzeri and Diamond, 2010). This scheme, however, has the disadvantage of requiring similarly stimulus tuned latencies in each subpopulation, hence introducing considerable redundancy in the neural representation. Exploiting separate neuron subsets for stimulus encoding and the saliency signal, in contrast, puts little constraints on what and how the remaining majority of neurons may encode, thereby offering the benefits of a high dimensional coding space.

### Efference motor copies as reference signals

A different possibility for sensory systems to receive temporal information about stimulus context is by direct feedback related to motor commands. Especially for systems that critically rely on the active exploration of the environment, such as the rodent somatosensory system, olfaction or vision, the problem of decoding information from precise spike times could be solved if sensory inputs were generated in response to an active motor

command. Indeed, feedback or motor-efference signals related to saccadic eye movements (Gawne et al., 1996; Gollisch, 2009), to sniffing during olfactory exploration (Shusterman et al., 2011) or to the rhythmic movement of facial whiskers (Diamond et al., 2008b; Hill et al., 2011) have been reported in the respective sensory areas. These motor efference signals could provide sensory cortices with some estimate of stimulus timing, or at least narrow down the processing into a window of “expectation” that may be used to constrain the decoding of time-dependent responses. Rats, for example, sweep their vibrissae toward objects of interest and may be able to register incoming spike trains with respect to their whisker protraction with a resolution of some tens of milliseconds (Kleinfeld et al., 2006; Diamond et al., 2008a). Similarly, neurons within the olfactory bulb can elicit responses that are precisely timed relative to the sniff cycle (Shusterman et al., 2011).

In the auditory system, however, there are no obvious motor efference copies. The problem of establishing an intrinsic reference frame is therefore even more compelling for auditory processing than for other modalities, especially in absence of external predictive clues of stimulus dynamics. While it may well be that attention or cross-modal inputs to the auditory system provide feedback related to active exploration (Lakatos et al., 2009; Schroeder et al., 2010), general sound processing mechanisms likely rely on intrinsic reference signals driven by sensory inputs (Chase and Young, 2007).

### Encoding based on purely internally defined reference frames

Another proposed mechanism for an intrinsic temporal reference is the phase of firing, which does not rely on a direct relationship to stimulus onset. The phase of firing codes information by the relative time of a spike with respect to an ongoing intrinsic rhythm (Hopfield, 1995; Lisman, 2005; Tiesinga et al., 2008; Panzeri et al., 2010), and thereby facilitates decoding and the organizing of information over time and across populations (Hopfield, 1995; Lisman, 2005; Fries, 2009; Panzeri et al., 2010). However, using oscillations as intrinsic temporal reference naturally constrains the speed of computations by the cycle length of the respective oscillation. For slow oscillations such as the auditory theta rhythm (Luo and Poeppel, 2007; Kayser et al., 2009; Chandrasekaran et al., 2010) this would result in a relatively slow encoding process, which seems at odds with the fast speed at which sensory systems can detect or recognize natural stimuli (Thorpe et al., 1996; VanRullen et al., 2005; Murray et al., 2006). The high speed of perception seems better accommodated by intrinsic references that are immediately available following stimulus occurrence and which do not necessitate integration over longer time windows. One such example is the rapid onset of stereotyped neurons described here.

### References

- Abbott LF, Dayan P (1999) The effect of correlated variability on the accuracy of a population code. *Neural Comput* 11:91–101.
- Bair W, Koch C (1996) Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Comput* 8:1185–1202.
- Bizley JK, Walker KM, King AJ, Schnupp JW (2010) Neural ensemble codes for stimulus periodicity in auditory cortex. *J Neurosci* 30:5078–5091.
- Chandrasekaran C, Tureson HK, Brown CH, Ghazanfar AA (2010) The influence of natural scene dynamics on auditory cortical activity. *J Neurosci* 30:13919–13931.
- Chase SM, Young ED (2007) First-spike latency information in single neurons increases when referenced to population onset. *Proc Natl Acad Sci U S A* 104:5175–5180.
- Chechik G, Anderson MJ, Bar-Yosef O, Young ED, Tishby N, Nelken I (2006) Reduction of information redundancy in the ascending auditory pathway. *Neuron* 51:359–368.
- deCharms RC, Zador A (2000) Neural representation and the cortical code. *Annu Rev Neurosci* 23:613–647.
- Diamond ME, von Heimendahl M, Arabzadeh E (2008a) Whisker-mediated texture discrimination. *PLoS Biol* 6:e220.
- Diamond ME, von Heimendahl M, Knutsen PM, Kleinfeld D, Ahissar E (2008b) ‘Where’ and ‘what’ in the whisker sensorimotor system. *Nat Rev Neurosci* 9:601–612.
- Engineer CT, Perez CA, Chen YH, Carraway RS, Reed AC, Shetake JA, Jakamsetti V, Chang KQ, Kilgard MP (2008) Cortical activity patterns predict speech discrimination ability. *Nat Neurosci* 11:603–608.
- Foffani G, Chapin JK, Moxon KA (2008) Computational role of large receptive fields in the primary somatosensory cortex. *J Neurophysiol* 100:268–280.
- Fries P (2009) Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu Rev Neurosci* 32:209–224.
- Furukawa S, Xu L, Middlebrooks JC (2000) Coding of sound-source location by ensembles of cortical neurons. *J Neurosci* 20:1216–1228.
- Gawne TJ, Kjaer TW, Richmond BJ (1996) Latency: another potential code for feature binding in striate cortex. *J Neurophysiol* 76:1356–1360.
- Gollisch T (2009) Throwing a glance at the neural code: rapid information transmission in the visual system. *HFSP J* 3:36–46.
- Gollisch T, Meister M (2008) Rapid neural coding in the retina with relative spike latencies. *Science* 319:1108–1111.
- Heil P, Irvine DR (1997) First-spike timing of auditory-nerve fibers and comparison with auditory cortex. *J Neurophysiol* 78:2438–2454.
- Hill DN, Curtis JC, Moore JD, Kleinfeld D (2011) Primary motor cortex reports efferent control of vibrissa motion on multiple timescales. *Neuron* 72:344–356.
- Hopfield JJ (1995) Pattern recognition computation using action potential timing for stimulus representation. *Nature* 376:33–36.
- Jaramillo S, Zador AM (2011) The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat Neurosci* 14:246–251.
- Kayser C, Montemurro MA, Logothetis NK, Panzeri S (2009) Spike-phase coding boosts and stabilizes the information carried by spatial and temporal spike patterns. *Neuron* 61:597–608.
- Kayser C, Logothetis NK, Panzeri S (2010) The millisecond encoding precision of auditory cortex neurons. *Proc Natl Acad Sci U S A* 107:16976–16981.
- Kleinfeld D, Ahissar E, Diamond ME (2006) Active sensation: insights from the rodent vibrissa sensorimotor system. *Curr Opin Neurobiol* 16:435–444.
- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320:110–113.
- Lakatos P, O’Connell MN, Barczak A, Mills A, Javitt DC, Schroeder CE (2009) The leading sense: supramodal control of neurophysiological context by attention. *Neuron* 64:419–430.
- Lisman J (2005) The theta/gamma discrete phase code occurring during the hippocampal phase precession may be a more general brain coding scheme. *Hippocampus* 15:913–922.
- Liu LF, Palmer AR, Wallace MN (2006) Phase-locked responses to pure tones in the inferior colliculus. *J Neurophysiol* 95:1926–1935.
- Luo H, Poeppel D (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54:1001–1010.
- Magri C, Whittingstall K, Singh V, Logothetis NK, Panzeri S (2009) A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. *BMC Neurosci* 10:81.
- Montemurro MA, Senatore R, Panzeri S (2007) Tight data-robust bounds to mutual information combining shuffling and model selection techniques. *Neural Comput* 19:2913–2957.
- Murray MM, Camen C, Gonzalez Andino SL, Bovet P, Clarke S (2006) Rapid brain discrimination of sounds of objects. *J Neurosci* 26:1293–1302.
- Nelken I, Chechik G (2007) Information theory in auditory research. *Hear Res* 229:94–105.
- Nelken I, Chechik G, Mscic-Flogel TD, King AJ, Schnupp JW (2005) Encoding stimulus information by spike numbers and mean response time in primary auditory cortex. *J Comput Neurosci* 19:199–221.
- Panzeri S, Diamond ME (2010) Information carried by population spike

- times in the whisker sensory cortex can be decoded without knowledge of stimulus time. *Front Synaptic Neurosci* 2:17.
- Panzeri S, Petersen RS, Schultz SR, Lebedev M, Diamond ME (2001) The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron* 29:769–777.
- Panzeri S, Senatore R, Montemurro MA, Petersen RS (2007) Correcting for the sampling bias problem in spike train information measures. *J Neurophysiol* 98:1064–1072.
- Panzeri S, Brunel N, Logothetis N, Kayser C (2010) Sensory neural codes using multiplexed temporal scales. *Trends Neurosci* 33:111–120.
- Perkel DH, Bullock TH (1968) Neural coding. *Neurosci Res Prog Bull* 6:221–348.
- Quiñero Quiroga R, Panzeri S (2009) Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci* 10:173–185.
- Reich DS, Mechler F, Purpura KP, Victor JD (2000) Interspike intervals, receptive fields, and information encoding in primary visual cortex. *J Neurosci* 20:1964–1974.
- Remedios R, Logothetis NK, Kayser C (2009) An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. *J Neurosci* 29:1034–1045.
- Rieke F, Warland D, de Ruyter van Steveninck RR, Bialek W (1999) *Spikes—exploring the neural code*. Cambridge, MA: MIT.
- Russ BE, Ackelson AL, Baker AE, Cohen YE (2008) Coding of auditory-stimulus identity in the auditory non-spatial processing stream. *J Neurophysiol* 99:87–95.
- Schnupp JW, Hall TM, Kokelaar RF, Ahmed B (2006) Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J Neurosci* 26:4785–4795.
- Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P (2010) Dynamics of active sensing and perceptual selection. *Curr Opin Neurobiol* 20:172–176.
- Sharpee TO, Atencio CA, Schreiner CE (2011) Hierarchical representations in the auditory cortex. *Curr Opin Neurobiol* 21:761–767.
- Shusterman R, Smear MC, Koulakov AA, Rinberg D (2011) Precise olfactory responses tile the sniff cycle. *Nat Neurosci* 14:1039–1044.
- Sokal RR, Rohlf FJ (1995) *Biometry*. New York: W.H. Freeman and Company.
- Soteropoulos DS, Baker SN (2009) Quantifying neural coding of event timing. *J Neurophysiol* 101:402–417.
- Strong SP, Koberle R, de Ruyter van Steveninck RR, Bialek W (1998) Entropy and information in neural spike trains. *Phys Rev Lett* 80:197–200.
- Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381:520–522.
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292:290–293.
- Tiesinga P, Fellous JM, Sejnowski TJ (2008) Regulation of spike timing in visual cortical circuits. *Nat Rev Neurosci* 9:97–107.
- VanRullen R, Guyonneau R, Thorpe SJ (2005) Spike times make sense. *Trends Neurosci* 28:1–4.
- Victor JD (2000) How the brain uses time to represent and process visual information. *Brain Res* 886:33–46.
- Victor JD, Purpura KP (1996) Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J Neurophysiol* 76:1310–1326.
- Wang X, Lu T, Bendor D, Bartlett E (2008) Neural coding of temporal information in auditory thalamus and cortex. *Neuroscience* 157:484–494.
- Yang Y, DeWeese MR, Otazu GH, Zador AM (2008) Millisecond-scale differences in neural activity in auditory cortex can drive decisions. *Nat Neurosci* 11:1262–1263.
- Zohar O, Shackleton TM, Nelken I, Palmer AR, Shamir M (2011) First spike latency code for interaural phase difference discrimination in the guinea pig inferior colliculus. *J Neurosci* 31:9192–9204.