Neural population encoding and decoding of sound source location across sound level in the rabbit inferior colliculus

Mitchell L. Day\textsuperscript{1,2} and Bertrand Delgutte\textsuperscript{1,2}

\textsuperscript{1} Eaton-Peabody Laboratories, Massachusetts Eye and Ear, Boston, MA
\textsuperscript{2} Department of Otology and Laryngology, Harvard Medical School, Boston, MA

Abbreviated title: Neural coding of sound location across level in IC

Address for Correspondence: Mitchell Day, Dept. of Biological Sciences, Ohio University, 107 Irvine Hall, Athens, OH 45701. E-mail: daym1@ohio.edu.

Author contributions: M.L.D. conception and design of research; M.L.D. performed experiments; M.L.D. analyzed data; M.L.D. and B.D. interpreted results of experiments; M.L.D. prepared figures; M.L.D. drafted manuscript; M.L.D. and B.D. edited and revised manuscript; M.L.D. and B.D. approved final version of manuscript.
Abstract

At lower levels of sensory processing, the representation of a stimulus feature in the response of a neural population can vary in complex ways across different stimulus intensities, potentially changing the amount of feature-relevant information in the response. How higher-level neural circuits could implement feature decoding computations that compensate for these intensity-dependent variations remains unclear. Here, we focused on neurons in the inferior colliculus (IC) of unanesthetized rabbits, whose firing rates are sensitive to both the azimuthal position of a sound source and its sound level. We found that the azimuth tuning curves of an IC neuron at different sound levels tend to be linear transformations of each other. These transformations could either increase or decrease the mutual information between source azimuth and spike count with increasing level for individual neurons, yet population azimuthal information remained constant across the absolute sound levels tested (35, 50 and 65 dB SPL), as inferred from the performance of a maximum-likelihood, neural population decoder. We harnessed evidence of level-dependent, linear transformations to reduce the number of free parameters in the creation of an accurate, cross-level population decoder of azimuth. Interestingly, this decoder predicts monotonic azimuth tuning curves, broadly sensitive to contralateral azimuths, in neurons at higher levels in the auditory pathway.

Keywords: sound localization; azimuth; rabbit; inferior colliculus; population code
Introduction

The ability of neural representations to code features of sensory stimuli in the face of changes in the intensity of the stimulus is a general issue across sensory modalities. For instance, organisms can visually recognize an object whether it is a sunny or overcast day, and can localize a sound source whether it is a loud car horn or the soft snap of a twig. Stimulus features, such as a visual object edge or sound source location, are encoded in the firing patterns of neural populations. At lower levels of the central nervous system these firing patterns also change with stimulus intensity. Does the amount of information from lower-level neurons regarding other stimulus features then also change with intensity? Moreover, how could higher-level neurons decode a stimulus feature from a lower-level population response that changes with intensity? We address these questions with respect to the specific feature of sound source azimuth (i.e., horizontal location) encoded in the firing rates of inferior colliculus (IC) neurons.

Perception of sound source azimuth relies on differences between the acoustical waveforms arriving at the two ears, specifically interaural time and level differences (ITD and ILD; (Middlebrooks and Green 1991)). Although these two binaural, acoustical cues are invariant to changes in sound level, sound level, as well as ITD and ILD, are implicitly or explicitly represented in the joint response of the left and right auditory nerves. Individual neurons whose firing rates are sensitive to ITD or ILD are first found in the auditory brainstem: the medial superior olive (MSO) mainly encodes the ITD of low-frequency sounds, while the lateral superior olive (LSO) mainly encodes the ILD of high-frequency sounds (Grothe et al. 2010). The predominant projections from MSO and LSO converge on the IC, as do additional projections from the cochlear nuclei carrying monaural information. The population of IC neurons therefore carries nearly all the information on both sound source azimuth and level fed forward to the cortex.
through the thalamus (but see Schofield et al. (2014) for evidence of small, direct olivothalamic projections).

Previous studies indicate that IC neurons tend to have broad “azimuth tuning curves” (i.e., mean firing rate as a function of azimuth), with both peaks that maintain the same preferred azimuth, and widths that can either increase or remain constant, with increases in sound level (Aitkin and Martin 1987; Delgutte et al. 1999; Irvine and Gago 1990; Kuwada et al. 2011; Kuwada and Yin 1983; Moore et al. 1984; Semple and Kitzes 1987; Sterbing et al. 2003; Yin et al. 1986). These studies all characterized the effects of sound level on coding of azimuth by classifying or quantifying changes in azimuth tuning curves across sound levels relative to each neuron’s threshold level, not absolute sound level, and did not take into account the intrinsic variability in neural responses over multiple presentations of a stimulus. In the present study, we instead quantified azimuthal information for individual neurons using mutual information (Pecka et al. 2010) and for the neural population using the performance of neural decoders. Both methods take trial-to-trial variability in firing rate (neural noise) into account, in addition to mean firing rate, thereby providing a more meaningful characterization of the effects of level on coding of azimuth appropriate for comparison to psychophysical performance in sound localization. Moreover, neural data were collected at the same absolute sound levels across neurons, allowing a direct comparison of population azimuthal information across level. Finally, all but the Kuwada et al (2011) study were performed on animals under anesthesia, which can alter ITD tuning curves (Kuwada et al. 1989). Our data were also collected from unanesthetized rabbits, thereby eliminating the confound of anesthesia.

We found that the azimuth tuning curves of an IC neuron at different sound levels tend to be linear transformations of each other. Some transformations increased azimuthal information,
while others decreased it, but population azimuthal information tended to remain the same across level. We further show that knowledge of level-dependent, linear transformations can reduce the complexity of decoding source azimuth from neural firing rates by reducing free parameters.

**Materials and Methods**

**Experimental methods**

Data were collected from two adult, female, Dutch-belted rabbits. All procedures were approved by the Institutional Animal Care and Use Committee of Massachusetts Eye and Ear. As described previously (Day et al. 2012; Devore and Delgutte 2010), the skull of each rabbit was surgically fitted with a metal bar for head-fixing during recording sessions, and a metal cylinder to environmentally isolate a small craniotomy overlying occipital cortex. Neural data were recorded from unanesthetized rabbits in head-fixed sessions, each up to 2.5 hours in duration, repeated at most once daily, and spread over 8 months. Rabbits were viewed over closed-circuit video and made frequent nose wiggles during recording sessions—a behavior which for rabbits indicates wakefulness.

In each session, a tungsten microelectrode (either A-M Systems or Microprobes; 1-kHz impedance = 5 MΩ) was lowered through occipital cortex into the IC. The neural signal was amplified, bandpass filtered from 1–3 kHz, digitally sampled at 100 kHz, and fed to a software spike detector. Extracellular spikes of single neurons were isolated based on voltage threshold crossings, consistency of the amplitude and shape of triggered waveforms, and online analysis of interspike intervals. Data were only included in the present study from neurons for which offline analysis confirmed that less than 1% of interspike intervals were less than 0.75 ms. Determination
that neural signals likely arose from the central nucleus of the IC was based on three criteria (Aitkin et al. 1975; Day et al. 2012; Nelson and Carney 2007): 1) robust, sound-evoked activity to noise stimuli; 2) non-habituating response across trials; and 3) a dorsoventral progression from low to high BFs. In one rabbit, an electrolytic lesion was made under anesthesia, subsequent to the last recording session in the general area of previous recordings. Histological analysis of the brain tissue revealed the lesion to be within the middle of the IC.

Recordings were made inside a double-walled, sound-attenuating chamber. Stimuli were created in MATLAB (Mathworks), digitally filtered to correct for the transfer function of the acoustic assembly, and converted to analog signals by a 24-bit digital-to-analog converter at a sampling rate of 50 kHz. The acoustic signal was produced by a pair of speakers (Beyer-Dynamic DT-48) attached to sound tubes running through custom-fitted ear molds. A probe-tube microphone (Etymotic ER-7C) measured acoustic pressure at the entrance to the ear canal at the end of the sound delivery tube. At the beginning of each recording session, we measured sound pressure in each ear in response to a broadband chirp stimulus and created inverse filters over the range of 0.1–25 kHz to correct for filtering by the acoustic assembly.

Sound stimuli were presented in virtual acoustic space via the sealed ear molds. To make a stimulus directional, we filtered the left and right waveforms with directional transfer functions (DTFs) corresponding to a particular azimuthal location. These DTFs were originally measured from a different rabbit, but processed to remove idiosyncratic spectral features, while keeping the appropriate ITD and ILD across both frequency and azimuth (Day et al. 2012).

IC neurons were searched for using a 60-dB SPL noise burst that alternated between 0 and 500 μs ITD. Measurements were collected from any neuron that responded to either stimulus.
Once a neuron was isolated, we measured the rate-level function: 200-ms bursts of broadband noise (0.1–17 kHz; 5-ms cos² on/off ramp) were presented at 0° azimuth every 500 ms at levels 0–70 dB SPL in 5-dB increments. Five trials were collected for each level, and levels were presented randomly. The rate-level function was computed as the average firing rate over stimulus duration at each level. Spontaneous rate was defined as the average firing rate of the last 100 ms of silence between stimuli. Sound level threshold was then defined as the first level for which the standard error of the firing rate did not overlap with the spontaneous rate.

Next, we measured azimuth tuning curves: 300-ms bursts of broadband noise were presented every 600 ms at each of 13 azimuthal locations in the front horizontal plane (15° resolution) and at sound levels of 35, 50 and 65 dB SPL. Between 4–10 trials were collected for each azimuth/level combination; 60% of neurons had 10 trials while only 13% had less than 6. Azimuth/level combinations were presented randomly. The azimuth tuning curve at a particular level was computed as the average firing rate over stimulus duration at each azimuth. For 38 of the 60 neurons, a noise burst was randomly generated for each trial, while for the rest the same burst of noise was used across trials (frozen noise).

Finally, we measured frequency tuning to tone pips in the contralateral ear. Tone pips were presented in an automatic, threshold-tracking procedure (Kiang and Moxon 1974), where characteristic frequency was defined as the frequency with the lowest threshold. In cases where the tracking procedure failed due to suppressive responses, tone pips were instead presented across frequency at a single low sound level near threshold. In this case best frequency was defined as the frequency with the greatest firing rate. In some cases frequency tuning was measured from background neural activity if isolation was lost prematurely, since neighboring IC neurons tend to have similar characteristic frequencies (Chen et al. 2012; Seshagiri and Delgutte 2007). For
simplicity, we referred to both characteristic and best frequencies throughout the paper as “BF.”

Information theoretic analysis

The mutual information (MI; (Cover and Thomas 2006)) between azimuth $X$ and spike count $Y$ of an individual neuron was computed as:

$$I(X; Y) = \sum_x \sum_y p(x) \ p(y|x) \ \log_2 \left( \frac{p(y|x)}{p(y)} \right)$$

[Equation 1]

in units of bits. We assumed a uniform stimulus distribution over the $M = 13$ azimuthal locations, $p(x) = 1/M$. The maximum possible MI is equal to the entropy of the stimulus ensemble, in this case $\log_2 M = \log_2(13) = 3.70$ bits. We assumed a parametric form of the conditional spike count distribution, $p(y|x)$, similar to Pecka et al (2010), but while they assumed a Laplace distribution, we assumed a gamma distribution whose shape and scale parameters were estimated by maximum likelihood from the neural data using the MATLAB function, gamfit. The gamma distribution was useful to model spike count distributions because, like the Poisson distribution, it is nonnegative, can take an approximately exponential shape at low spike counts and a Gaussian shape at high spike counts, but unlike Poisson, both mean and variance can be set independently to fit the data. $p(y|x)$ was discretized by setting it equal to the difference in the cumulative gamma distribution of neighboring points midway between each spike count. The marginal spike count distribution, $p(y)$, was then computed as the joint distribution summed over $x$.

Estimating MI from data with a finite number of trials yields an upwardly biased estimate (Treves and Panzeri 1995). To debias MI, we used a bootstrap method (Chase and Young 2005):
sample size bias was estimated as the difference between the mean of 500 bootstrapped estimates of MI and the original MI estimate. A bootstrapped data set for each stimulus azimuth was created by randomly selecting $m$ spike counts, with replacement, from the $m$ trials measured in response to that stimulus azimuth. The median bias across our sample was 0.14 bits. All reported MI values are debiased: the original MI estimate minus the sample size bias.

Neurophysiological studies often compute MI via nonparametric estimation of spike count distributions, necessitating a high number of repetitions for reliable estimation. We instead assumed gamma spike count distributions, whose two free parameters can be reliably estimated from fewer repetitions. We checked the accuracy and precision of our method of estimating MI in two ways: 1) using simulated data for which MI was exactly known, and 2) using experimental data measured with a higher number of trials. In the first case, we selected the parameters of the simulated data by setting neural variability to be Poisson and mean spike counts across azimuth to a typical, monotonically increasing tuning curve selected from one of the neurons in our sample (Fig. 1A, green curve). In this way, MI can be exactly computed. We then created 500 simulated data sets for each number of trials between 4 and 10 by randomly generating Poisson spike counts. Figure 1B shows the mean MI values computed from the 500 simulated data sets for each sample size, along with the exact MI. The mean MI was at most 0.02 bits greater than the exact MI for all sample sizes, indicating a very small amount of residual bias in our method of computing debiased MI. The SD of the MI estimates decreased with increasing number of trials from 0.16 bits for 4 trials to 0.08 bits for 10 trials. Next, we tested our method of computing MI on real data collected from a neuron using 30 trials (Fig. 1A, purple curve). We created 500 subsampled data sets for each number of trials by sampling spike counts, without replacement, from the original data at each azimuth. Figure 1B again shows MI values computed from the 500 subsampled sets for each
sample size, along with the MI computed from all 30 trials. The SD of the MI estimates again
decreased with increasing sample size as for the simulated data, but in this case the residual bias
between mean MI and the high-repetition MI also decreased, from 0.05 bits for 4 trials to 0.01 bits
for 10 trials. Overall, the results from testing our method of computing MI with limited trial
repetitions suggest high accuracy and reasonable precision.

Population decoding analysis

We decoded azimuthal location from the population IC response using a linear decoder
based on maximum-likelihood estimation (Day and Delgutte 2013; Jazayeri and Movshon 2006).
The population response is the set of spike counts from all neurons in the population from a single
trial of source azimuth: \( \{n_i\}_{i=1}^N \), where \( n_i \) is the spike count of the \( i \)th neuron and \( N \) is the number
of neurons. If we assume that spike counts are Poisson distributed and conditionally independent
between neurons, then the logarithm of the likelihood that the population response from a given
trial is associated with an azimuth may be expressed as:

\[
\log \mathcal{L} (\theta | \{n_i\}) = \sum_{i=1}^{N} n_i \log f_i (\theta) - \sum_{i=1}^{N} f_i (\theta) - \sum_{i=1}^{N} \log (n_i !)
\]

[Eqn. 2]

where \( \theta \) is azimuth and \( f_i (\theta) \) is the azimuth tuning curve of the \( i \)th neuron expressed as mean spike
count, not mean firing rate. Comparing the equation to the schematic in Figure 4A, the set of
possible \( \theta \) are the output-layer neurons, the log likelihood of each azimuth represent the responses
of the output-layer neurons, the set of \( n_i \) represents the spike counts of the first-layer neurons, the
set of \( \log f_i (\theta) \) represent the synaptic weights, and the second term represents the biases on the
output-layer neurons. The third term may be ignored since it is not dependent on $\theta$ and therefore
does not affect which $\theta$ maximizes the likelihood. To test the decoder on experimental data, a spike
count from one trial was randomly selected from each neuron in response to a given azimuth, then
the estimated azimuth was chosen to maximize the log likelihood. Test spike counts were removed
from the data set before computing the set of $f_i(\theta)$ (i.e., “training” the decoder) to avoid overfitting.
This procedure was then iterated 500 times for each azimuth. Instances when $f_i(\theta) = 0$ were set to
$1/(m + 1)$, where $m$ is number of trials, in order to prevent taking the logarithm of zero. Decoder
performance was summarized by the RMS error across all test azimuths.

We used the same test procedure to implement a cross-level decoder of the combination of
azimuth and sound level (“combination” decoder) by making both the likelihood and tuning curves
also dependent on sound level: $L(\theta, \ell)$ and $f_i(\theta, \ell)$, respectively. In this case, a spike count was
randomly selected from each neuron in response to a given combination of azimuth and level, then
the estimated azimuth and level were chosen to maximize the log likelihood.

We also implemented an alternative cross-level decoder (“transformation” decoder) by
assuming that azimuth tuning curves at different sound levels were linearly related: $f_i(\theta, \ell) = \alpha_i(\ell) + \beta_i(\ell)$. The level-independent azimuth tuning shape, $g_i(\theta)$, was chosen for each neuron as the
azimuth tuning curve at the sound level with greatest MI. The level-dependent factors, $\alpha_i(\ell)$ and
$\beta_i(\ell)$, were 1 and 0, respectively, at the level from which $g_i(\theta)$ was selected; at each other level they
were computed by linear regression between the azimuth tuning curve at that level and $g_i(\theta)$.

Substituting for $f_i(\theta)$ in Equation 2 yields:

$$
\log \mathcal{L}((\theta|([n_i], \ell)) = \sum_{i=1}^{N} n_i \log(\alpha_i(\ell)g_i(\theta) + \beta_i(\ell)) - \sum_{i=1}^{N} (\alpha_i(\ell)g_i(\theta) + \beta_i(\ell)) - \sum_{i=1}^{N} \log(n_i!)
$$
By expanding and rearranging terms, this equation may be rewritten as:

$$\log \mathcal{L}(\theta|\{n_i\}, \ell))$$

$$= \sum_{i=1}^{N} n_i \log(g_i(\theta) + \gamma_i(\ell)) - \sum_{i=1}^{N} \alpha_i(\ell) g_i(\theta) + \sum_{i=1}^{N} n_i \log \alpha_i(\ell) - \sum_{i=1}^{N} \beta_i(\ell)$$

$$- \sum_{i=1}^{N} \log(n_i!)$$

where $\gamma_i(\ell) = \beta_i(\ell) / \alpha_i(\ell)$. The last three terms are independent of $\theta$ and may be ignored. The first two terms indicate that the decoder is still linear, but with both weights and biases dependent on sound level. The weights are represented by $\log(g_i(\theta) + \gamma_i(\ell))$ and the biases by the second term. In implementing the transformation decoder, it was assumed sound level, and therefore the set of $\alpha_i(\ell)$ and $\beta_i(\ell)$ were known (see Discussion). Again, test spike counts were removed from the data set before computing $g_i(\theta)$, $\alpha_i(\ell)$ and $\beta_i(\ell)$ to avoid overfitting.

The decoders were also tested on simulated spike count data. The set of azimuth tuning curves, $f_i(\theta)$, or alternatively $g_i(\theta)$, $\alpha_i(\ell)$ and $\beta_i(\ell)$, were chosen to match those of measured tuning curves, depending on which decoder was being tested. Test spike counts were randomly drawn from Poisson distributions with means and numbers of trials that matched the measured tuning curves.

Midway through experimental data collection, we noticed that decoder estimation errors were larger for experimental data than for simulated Poisson data. We considered the possibility that the uncontrolled stimulus variability caused by using a random burst of noise in each trial may
have increased spike count variability above that expected from a Poisson distribution. We therefore collected subsequent data using the same burst of noise in each trial. Decoder errors using this later data were still greater than those using simulated data, indicating that deviations from Poisson distributions in our earlier data set were not simply due to uncontrolled stimulus variability. We therefore combined both sets of data in our decoder analyses.

Lastly, we implemented an alternative maximum-likelihood decoder based on the assumptions that spike count distributions were both gamma distributed and conditionally independent between neurons. We assumed the conditional spike count distribution of each neuron, $p(n_i|\theta)$, was a gamma distribution whose shape and scale parameters were estimated by the method of moments from the neural data following removal of test data. Again, $p(n_i|\theta)$ was discretized by setting it equal to the difference in the cumulative gamma distribution of neighboring points midway between each spike count. Unlike the Poisson decoder, the likelihood of the gamma decoder was nonlinear on the spike counts.

Statistical tests

We created a paired bootstrap test to determine the statistical significance of the effect of sound level on decoder performance. The null hypothesis was that the difference in decoder RMS error between any two sound levels was zero. First, $N$ sampled neurons were selected, with replacement, from the population of $N$ neurons. Then the decoder was tested on spike counts from the resampled population separately at 35, 50 and 65 dB SPL. The difference in RMS errors was then computed for three groups: $\varepsilon_{35} - \varepsilon_{35}$, $\varepsilon_{65} - \varepsilon_{50}$, and $\varepsilon_{65} - \varepsilon_{35}$. Finally, the whole procedure was iterated 2,000 times. A two-sided $p$-value for each group of decoder error differences was
computed directly from the distribution of bootstrap-estimated error differences by doubling the most significant one-sided p-value. Finally, the Benjamini-Hochberg method was used to correct for multiple comparisons (Wasserman 2004).

In a separate paired bootstrap test, the null hypothesis was that the difference in RMS error between the combination and transformation decoders was zero. Both the combination and transformation decoders were tested on spike counts from the same resampled population, separately at each sound level. The difference of RMS error was then computed for three level conditions: $\varepsilon_{T,35} - \varepsilon_{C,35}$, $\varepsilon_{T,50} - \varepsilon_{C,50}$, and $\varepsilon_{T,65} - \varepsilon_{C,65}$ (T: transformation; C: combination). Iteration and p-value computation were the same as for the other decoder bootstrap test.

Associations between measured or computed values were assessed with the nonparametric Kendall’s rank correlation. Significance of the effect of sound level on median MI was assessed with a Friedman’s test, which is a nonparametric test that adjusts for the effects of cross-level measurements being made in the same group of neurons. Finally, in plotting the log variance of spike counts in Figure 5A, simply taking the logarithm of the sample variance leads to a biased estimate of log variance. We computed an unbiased estimate using a Taylor series expansion, adding $\log_{10}e/(m-1)$ to the $\log_{10}$ sample variance, where $m$ is the number of trials (Gershon et al. 1998).

Results

Azimuthal information in IC firing rates changes with absolute sound level for individual neurons, but not over the population
To determine how azimuthal information encoded in the firing rates of IC neurons changes with sound level, we measured azimuth tuning curves of IC neurons at three sound levels. Data presented are from 60 single neurons in the left and right IC of two unanesthetized, female rabbits. Based on physiological criteria detailed in Methods, single neurons were likely located in the central nucleus of the IC, which comprises the main feed-forward projection to the auditory thalamus on the way to the auditory cortex (Winer and Schreiner 2005). Moreover, neurons in the sample had BFs from 0.38 to 25 kHz (25th, 50th and 75th percentiles of 1.1, 3.6 and 8.8 kHz, respectively), covering most of the audible range of rabbits (Heffner and Masterton 1980). For each neuron, average firing rates were measured in response to broadband noise bursts presented in virtual acoustic space at 13 frontal locations between ±90° in the horizontal plane (15° resolution), and at sound levels of 35, 50 and 65 dB SPL. To quantify the azimuthal information from each neuron at each sound level, we computed the MI between spike count and azimuth. MI quantifies what can be learned about sound source azimuth by observing the spike count of a neuron without making any assumption about the nature of the neural code.

The MI of individual neurons could change greatly with sound level, and could either increase or decrease as sound level increased (Fig. 2). Points away from both the x- and y-axes in Figure 2 indicate neurons whose firing rates were informative about azimuth at both levels. For example, the firing rates of the neuron in Figure 3A increased as level increased from 35 to 50 dB SPL, yet MI changed little from 1.85 to 1.95 bits. The broad selectivity for azimuthal locations contralateral to the recording site (positive) in Figure 3A is characteristic of most IC neurons (Aitkin et al. 1984; Day and Delgutte 2013; Delgutte et al. 1999). Azimuth tuning curves could be monotonic, as in Figure 3A, or nonmonotonic, as in Figure 3B. A small minority of neurons exhibited broad selectivity for ipsilateral azimuths or more complex tuning shapes.
Points near the y-axes in Figure 2 indicate neurons whose firing rates became informative about azimuth as level increased, while those near the x-axes indicate neurons whose firing rates became uninformative. Neurons whose firing rates became informative had lower-level tuning curves that were approximately flat, often near the spontaneous rate (Fig. 3B) because the lower sound level was near or below the neuron’s threshold of response to sound. However, some neurons had lower-level tuning curves that were flat and well above spontaneous rate, such as in Figure 3C, indicating the neuron responded to the noise stimulus at the lower level yet was not yet sensitive to source azimuth.

The opposite case was seen for neurons whose firing rates became uninformative as sound level increased in that the higher-level tuning curves became flatter. Most often the flattening was consistent with saturation of firing rate, such as in Figure 3D. There, the firing rate began to saturate to a maximal value of 85 sp/s at 65 dB SPL, so that sensitivity to azimuth weakened and the MI became small. In some neurons the higher-level tuning curves flattened to a non-maximal firing rate (Fig. 3E), inconsistent with rate saturation and instead indicative of a nonmonotonic relationship between firing rate and sound level as might be created by inhibition.

While the MI of individual neurons could change greatly with sound level, there was no significant effect of level on median MI across the neural sample (Fig. 2; $p = 0.89$, $\chi^2(2) = 0.23$, Friedman’s test; median MI: 0.57, 0.66 and 0.67 bits for 35, 50 and 65 dB SPL, respectively). Neurons that became more informative with an increase in sound level tended to be balanced by other neurons that became less informative.

The lack of change in median MI with sound level suggests that azimuthal information encoded in the joint response of the IC population (i.e., population MI) may remain unchanged.
over the 30-dB range of levels investigated. However population MI is not the same as the median MI over individual neurons because population MI depends on the joint distribution of spike counts across azimuth for the entire population of neurons. Computation of population MI is impractical because of the enormous number of possible combinations of spike count values over which the joint distribution must be estimated for each azimuth and because, strictly speaking, such computation would require simultaneous recordings from all the neurons in the population. For this reason, we adopted an indirect approach to inferring azimuthal information in the population response by evaluating the performance of a neural population decoder (Quian Quiroga and Panzeri 2009).

We used a linear decoder where the output of a putative output layer of neurons, each of which corresponds to a particular azimuth, is a weighted and biased sum of spike counts from the IC sample (Fig. 4A; (Day and Delgutte 2013)). Under the assumptions that IC spike counts are both conditionally independent and distributed in a Poisson manner, the weights and biases were computed such that the activities of each output-layer neuron are monotonically related to the likelihood of the population response being associated with a given azimuth (Jazayeri and Movshon 2006). Estimated azimuth was therefore chosen to maximize likelihood. Three different types of decoders were implemented and tested, only one of which is discussed in this section. This first decoder was trained and tested on data from each sound level separately, hence it is called the “single-level decoder”; figure 4B shows decoder performance at each level. Consistent with the stability of median MI over level, there was no significant effect of sound level on the performance of the decoder, as measured by RMS error across all azimuths (Fig. 4C; p > 0.05, paired bootstrap tests with correction for multiple comparisons). This result held when tested on data from each rabbit separately (N = 32 and 28 neurons; not shown).
Decoder estimates of azimuth will deviate from true maximum-likelihood estimates when the assumptions of conditional independence and Poisson statistics are not met. To verify that potential cross-level differences in performance were not being masked by incorrect assumptions, we evaluated the validity of each assumption with respect to our data set. Since neurons were recorded serially instead of simultaneously, we must assume conditional independence. This assumption is supported by recent data where simultaneous recordings in gerbil IC showed extremely weak pairwise noise correlations (Belliveau et al. 2014; Garcia-Lazaro et al. 2013). Other inaccuracies may occur when spike counts deviate from a Poisson distribution, or by training the decoder on a data set with limited stimulus repetitions. The average behavior of spike count distributions in our data set was consistent with Poisson statistics in that the mean spike count was equal to the variance of the spike count (Fig. 5A). However, the spike count distributions of some neurons at some azimuthal locations could have variance either smaller or larger than that expected for a Poisson distribution (Fig. 5B,C). We tested the decoder on simulated data in which spike counts were drawn randomly from Poisson distributions with repetitions and means that matched those of the measured azimuth tuning curves. Estimation errors were reduced by a factor of 2 for the simulated data (Fig. 5D), indicating that non-Poisson distributions in the real data did increase decoder error. However, there was still no significant effect of sound level on localization error with simulated data (p > 0.05, paired bootstrap tests with correction for multiple comparisons). We further considered the possibility that decoder performance may be improved by assuming gamma spike count distributions, in which mean and variance may be set independently to fit the data. We therefore estimated azimuth from the real data with a maximum-likelihood decoder assuming conditionally independent, gamma statistics. Decoder errors were, however, nearly identical to those for the Poisson decoder, and similar across
sound level ($\epsilon = 13.7, 11.0$ and $12.9^\circ$ for 35, 50 and 65 dB SPL, respectively; data not shown).

Altogether, our additional analyses do not provide evidence that cross-level differences in decoder performance were masked by unrealistic assumptions.

Some previous investigations of neural coding of azimuth highlighted neurons with tuning properties that were relatively invariant to changes in sound level (Aitkin and Martin 1987; Kuwada et al. 2011); the idea being that a subpopulation of level-invariant neurons may facilitate cross-level encoding of azimuth. We instead looked for evidence of neurons with high MI at all three sound levels. Figure 6 shows the cumulative fraction of neurons for which MI was greater than some minimum value at all three sound levels, as this minimum was swept down to zero. A distinct subpopulation of highly informative neurons across all levels should lead to an initial plateau in the cumulative function, followed by an increase as neurons that are not informative at every level accumulate. Instead, the cumulative function of Figure 6 gradually increases throughout, giving no indication of a distinct, level-robust subpopulation. This suggests that at each sound level, azimuthal information is distributed over partially overlapping subpopulations of IC neurons.

**MI of individual IC neurons is consistent across the tonotopic range**

Azimuthal sensitivity of firing rates in IC is dominated by different binaural cues in different frequency ranges: neurons with low or high BFs (2-kHz approximate boundary) have azimuth tuning curves largely dominated by ITD or ILD, respectively (Day and Delgutte 2013; Day et al. 2012). We therefore looked for differences in the MI values of individual neurons across BF. There was no correlation between MI and BF at 50 or 65 dB SPL, and only a weak correlation
at 35 dB SPL due to a decrease in MI below 1 kHz (Fig. 7A; 35 dB: \( r = 0.22, p = 0.01 \); 50 dB: \( r = -0.05, p = 0.57 \); 65 dB: \( r = -0.06, p = 0.49 \); Kendall’s rank correlation). Figure 7A shows MI divided by the maximum possible value it could attain given a uniform distribution of the 13 azimuthal locations (3.70 bits). This normalized MI is therefore the fraction of total azimuthal information at 15° resolution captured by IC spike counts. It is evident from Figure 7A that individual neurons in almost every frequency band at every level can provide a substantial fraction of maximum azimuthal information, consistently reaching around 40% for the best neurons.

For most neurons, we measured rate-level functions at 0° azimuth in order to estimate the threshold of response to noise. Sound level threshold was negatively correlated with BF (Fig. 7B; \( r = -0.36, p = 0.001 \), Kendall’s rank correlation), consistent with data from rabbit auditory nerve and behavioral audiogram over the same frequency range (Borg and Engstrom 1983; Borg et al. 1988; Heffner and Masterton 1980). In particular, most neurons with BF below 1 kHz had sound level thresholds near 35 dB SPL or higher. MI decreased to small values at sound levels near and below level threshold, as expected (Fig. 7C). The decrease in MI below 1 kHz at 35 dB SPL (Fig. 7A) therefore likely occurs because most neurons in these frequency bands are stimulated at or below their level thresholds.

Azimuth tuning curves in the IC undergo approximately pairwise linear transformations with sound level

Although the comparison of MI across sound levels is blind to the particular shapes of azimuth tuning curves, the extraction of azimuthal information from IC neurons by higher-level neural circuits may be facilitated by consistency in the shapes of tuning curves across sound level.
To quantify how azimuth tuning curves change with sound level, in each neuron we computed the correlations between tuning curves measured for every pair of levels. Figure 8 shows the Pearson product-moment correlation coefficient of every within-neuron, cross-level pair versus the minimum MI of the neuron between the two levels. Points with minimum MI near zero indicate neurons with little azimuthal information at either one or both levels. The relatively flat shapes of weakly informative tuning curves are dominated by measurement noise, not sensitivity to azimuth, therefore the correlation coefficients of pairs with small minimum MI were spread across all values. Further, those pairs with insignificant correlation (p ≥ 0.05, t-test; N = 41; closed circles) tended to be clustered near small minimum MI. Overall, there was a clear association between minimum MI and correlation coefficient (r = 0.5, p < 10^{-22}, Kendall’s rank correlation), such that cross-level tuning curve pairs that were more informative about azimuth tended to be highly correlated. Specifically, pairs in the top half of minimum MI values (> 0.4 bits) had a median correlation coefficient of 0.93. These high Pearson correlation values imply that the shapes of azimuth tuning curves are strongly similar at both sound levels and that changes in azimuth tuning curves across level can be reasonably approximated as pair-wise linear transformations.

Linear transformations can be additive (a vertical shift of the tuning curve), multiplicative (a scaling of the tuning curve) or both. For instance, the azimuth tuning curves of the neuron in Figure 9A at 50 and 65 dB SPL appear to be shifted versions of each other, indicating an additive transformation. A plot of the firing rate at 50 versus 65 dB SPL for each azimuth (Fig. 9B) shows an approximately linear relation, with a correlation coefficient of 0.96. Linear regression of these data yielded a bias (additive factor) of 33 sp/s, which was significantly different from 0 sp/s (p < 0.001, t-test), and a gain (multiplicative factor) of 0.94, which was not statistically different from 1 (p = 0.52, t-test). The transformation of tuning between these two levels was therefore additive.
For each of the 133 cross-level pairs that had a significant, positive correlation coefficient \((p < 0.05, t\text{-test})\), we performed linear regression to determine whether the additive factor significantly differed from 0 and whether the multiplicative factor significantly differed from 1 \((p < 0.05, t\text{-test})\). It was common for a neuron to have different types of linear transformations (e.g., additive vs. multiplicative) between different pairs of sound levels (32 of 41 neurons), so that strictly speaking the transformations can only be described as pairwise linear. In order to relate changes in MI to the type of linear transformation, we plotted histograms of the change in MI from lower to higher level across pairs, grouped by type of transformation (Figure 9C). Most multiplicative transformations had gains greater than 1 and tended to increase MI. For these transformations, the range of mean spike count across azimuth increased with the increase in sound level (Fig. 9D). Although this should reduce the overlap in spike count distributions for different azimuths, neural noise also increases at higher firing rates. Across our neural sample the variance of the spike count was, on average, equal to the mean spike count (Fig. 5A). Neural noise therefore tended to increase with mean count, but only as approximately the square root of the mean count. For the neuron in Figure 9D and others with multiplicative transformations, the increase in the range of mean spike count across azimuth was greater than the increase in neural noise, leading to an increase in MI.

All but one of the additive transformations had positive biases (Fig. 9C), and these most often decreased MI. For these transformations, the tuning curve shifted up with increasing level (Fig. 9A), leaving the range of mean spike count across azimuth the same while the neural noise increased, thereby leading to a greater overlap of spike count distributions and decrease in MI. All but two mixed transformations (i.e., both additive and multiplicative) had positive biases, but their gains could be either greater or lesser than 1. The combination of positive bias and gain lesser than
1 decreased MI, as expected (Fig. 9C). These transformations occurred as tuning curves began to saturate (Fig. 3D) or flatten to a non-maximal rate. Mixed transformations with positive bias and gain greater than 1 had additive and multiplicative factors that pushed the MI in opposite directions, yet these tended to increase MI (Fig. 9C), indicating the dominance of the multiplicative factor, which determines the range of firing rates.

In summary, the azimuth tuning curves of an IC neuron at different sound levels tend to be linear transformations of each other. These transformations tended to increase firing rates with increasing level, which could lead to either an increase or a decrease in MI. Increases in MI occurred when the range of firing rates increased. Decreases in MI occurred when tuning curves simply shifted upwards in firing rate, or when tuning curves flattened.

Knowledge of linear transformations reduces the complexity of cross-level decoding of source azimuth without sacrificing accuracy

The single-level linear decoder of Figure 4A has weights and biases that are completely determined by the mean spike counts of each neuron at each azimuth, i.e., the collection of azimuth tuning curves for a particular sound level (see Methods). However, since azimuth tuning curves can change substantially with sound level, and in different ways across neurons, the decoder weights and biases can be dramatically different across sound levels. How can azimuth be decoded from the IC population response at all sound levels when azimuth tuning changes in complex ways?

One obvious solution is to decode the combination of azimuth and level instead of azimuth alone. We augmented the linear decoder of Figure 4A so that each of the output-layer neurons
corresponds to a particular combination of azimuth and sound level (Fig. 10A). In a similar manner as before, the weights and biases were chosen such that the activities of the output-layer neurons were related to the likelihood of the population response being associated with each combination of azimuth and level. Figure 10B shows the performance of this “combination” decoder, with the azimuthal locations at each sound level concatenated on both the x and y axes. It is immediately apparent that azimuthal estimation errors only occurred within level; i.e., the decoder perfectly estimated sound level, albeit with a relatively coarse level resolution of 15 dB. Since there were no level estimation errors, the performance of the combination decoder in estimating azimuth was essentially the same as that of the single-level decoders in Figure 4B.

The combination decoder provides a theoretically optimal, maximum-likelihood estimate of azimuth and level, but is also complex, as measured by the number of free parameters. The weights and biases of this decoder are determined by the mean spike counts at each combination of azimuth and level. Therefore if there are \(N\) neurons, \(M\) azimuthal locations and \(L\) sound levels, then there are \(N \cdot M \cdot L\) free parameters. This becomes an enormous number if azimuth and level are sampled with resolution consistent with psychophysical acuity.

We therefore considered how our finding that azimuth tuning curves linearly transform with level could be harnessed to devise a less complex cross-level decoding strategy with fewer parameters. We assumed that the azimuth tuning curve of a neuron at a given level, \(f(\theta, l)\) where \(\theta\) and \(l\) are azimuth and level, respectively, can be expressed as \(\alpha(l) \cdot g(\theta) + \beta(l)\), where \(g(\theta)\) is the underlying azimuth tuning shape and \(\alpha(l)\) and \(\beta(l)\) are level-dependent multiplicative and additive factors, respectively. Substituting this expression into the likelihood equation (assuming conditionally independent, Poisson neurons) yields a linear decoder with output-layer neurons that correspond only to azimuth as in Figure 4A, but with level-dependent weights and biases (Eqn. 4).
The free parameters for this “transformation” decoder consist of the azimuth tuning shapes of every neuron, yielding $N \cdot M$ parameters, and the multiplicative and additive factors of every neuron at every level, yielding $2 \cdot N \cdot L$. In general, the number of free parameters of the combination decoder, $N \cdot M \cdot L$, is greater than that of the transformation decoder, $N \cdot M + 2 \cdot N \cdot L$, whenever $M > 4$ and $L > 1$. For instance, with the 13 azimuthal locations and 3 sound levels used in our measurements, the transformation decoder reduced the number of free parameters to about half that of the combination decoder. Even greater benefits would be obtained for denser azimuth and level sampling more consistent with psychophysical acuity.

We implemented the transformation decoder by setting $g(\theta)$ for each neuron to the azimuth tuning curve at the sound level with greatest MI. $\alpha(l)$ and $\beta(l)$ were then computed by linear regression with respect to $g(\theta)$. The localization performance of the transformation decoder was accurate and very similar to that of the combination decoder (Fig. 11A). In fact, there was no significant difference in RMS error between the combination and transformation decoders (Fig. 11B; $p > 0.05$, paired bootstrap tests with correction for multiple comparisons).

**Decoder predicts output layer neurons with monotonic azimuth tuning**

The linear decoders used in the present study can be implemented in plausible, feed-forward neural circuits (Fig. 4A and 10A), so it is reasonable to ask what azimuth tuning curves of the hypothetical output-layer neurons looks like. Figure 12A shows the azimuth tuning curves of the output-layer neurons corresponding to $-90$, $0$ and $90^\circ$ azimuth for the single-level decoder trained and tested on data at 65 dB SPL (Fig. 4B, bottom panel). The output responses were computed by the log likelihood equation (Eqn. 2), with $\theta$ set to either $-90$, $0$ or $90^\circ$, and spike
counts selected randomly from the data in response to each source azimuth. Interestingly, the azimuth tuning curve of each output layer neuron increases monotonically from ipsilateral to contralateral sources with a sigmoidal shape. At a source azimuth of 90°, the response of the output layer neuron corresponding to 90° is greater than that of the other two output layer neurons; similarly, the responses of the 0° and −90° output layer neurons are greater than the other two for sources at 0° and −90°, respectively. Altogether, the responses of output layer neurons generally increase from ipsilateral to contralateral sources, but are organized with respect to each other to ensure the appropriate maximum response at each source azimuth. While the tuning curves of output-layer neurons depicted in Figure 12A are from a decoder trained and tested on data at 65 dB SPL, similar monotonic tuning curves with source-specific ordering were obtained for decoders trained and tested at the other sound levels, as well as for the combination and transformation decoders when tested at each level.

The maximum response across output layer neurons is not perfectly aligned to the source azimuth from trial to trial, hence the decoder errors in Figure 4B. For instance, Figure 12B shows the responses of all output layer neurons to a source at 0° (the shaded area in Fig. 12A), on two different trials. These response profiles show the likelihood that the pattern of spike counts across IC neurons in each trial are associated with a sound source at a particular azimuth. The maximum response, and therefore the maximum likelihood, occurs for the 0° output-layer neuron in the first trial, while it occurs for the −15° neuron in the second trial.

Discussion

Using single-neuron recordings in unanesthetized rabbit, we showed that while MI
between source azimuth and spike counts of individual IC neurons could change greatly with sound level, the median MI across the neural sample was largely invariant with level and with location along the tonotopic axis. A maximum-likelihood decoder operating on the pattern of spike counts across IC neurons at each level estimated sound source azimuth equally accurately at all three sound levels. The transformations of azimuth tuning curves across sound levels could reasonably be characterized as pairwise linear, and this finding was used to reduce the complexity of a cross-level decoder of sound source azimuth that performed accurately for all three levels.

The collicular encoding of azimuth across level is summarized in Figure 13. A typical IC neuron first becomes informative about azimuth as sound level exceeds neural threshold, so that the azimuth tuning curve goes from a flat, unresponsive function to one in which firing rate varies appreciably (Fig. 13A). The tuning curve then undergoes linear transformations with further increases in level. Azimuthal information continues to increase so long as the range of firing rates increases with level. Azimuthal information begins to decrease as the tuning curve only shifts, but not scales, upward in firing rate, then decreases further as the tuning curve flattens either completely (solid curve) or partially (dashed curve).

The information-level curve depicted in Figure 13A is not the same across neurons. Neurons within each frequency band have information-level curves with different level thresholds (Fig. 7B), and possibly different maximum information values, peak levels, and widths (Fig. 13B). Ultimately, the heterogeneity of information-level curves leads to population azimuthal information that rises to a plateau at a relatively low sound level above the detection threshold (Fig. 13C). An outstanding question is why, between any two sound levels in the plateau, increases and decreases in MI tend to be balanced, leading to constant median MI and constant decoder performance over a wide range of levels.
Psychophysical studies in both human (Inoue 2001; Macpherson and Middlebrooks 2000; Miller and Recanzone 2009; Sabin et al. 2005; Vliegen and Van Opstal 2004) and cat (Gai et al. 2013) show that azimuthal localization performance remains constant at levels well above the detection threshold, but deteriorates near threshold. Further, performance asymptotes at an imperfect value, indicating the plateau in performance is not simply due to a ceiling effect. Our finding that collicular azimuthal information remains constant between 35–65 dB SPL is therefore consistent with psychophysical behavior. Whether performance degrades at higher sound levels is unknown: localization performance has only been measured up to 73 dB SPL in humans (Vliegen and Van Opstal 2004) and 78 dB SPL in cats (Gai et al. 2013). Our use of an unanesthetized, head-fixed animal preparation limited neural data collection to sound levels within a comfortable range.

Our finding that both median MI and decoder localization error are constant over a range of absolute sound levels is novel. Previous studies in MSO (Pecka et al. 2010) and cortex (Middlebrooks et al. 1998; Stecker et al. 2005) have also investigated changes in azimuthal information with sound level using either MI or population decoders. However in these studies, sound level was referenced to each neuron’s threshold, not absolute sound level. Thresholds vary substantially across neurons (Fig. 7B), therefore population trends over relative sound levels do not necessarily carry over to absolute sound levels. Miller and Recanzone (2009) tested the same spike count-based, maximum-likelihood decoder used in the present study on populations of auditory cortical neurons at several absolute sound levels. Decoder performance appeared to be similar across level in cortical areas CL and R, and less so in area A1. However, a cortical, spike count-based decoder is unlikely to capture all location-specific information, since cortical decoders that incorporate spike timing information significantly outperform those based on counts.
alone (Middlebrooks et al. 1994; Middlebrooks et al. 1998). In the IC, on the other hand, spike
counts contain the majority of information between spike trains and binaural cues (Belliveau et al.
2014; Chase and Young 2008).

We found that the transformation of azimuth tuning curves between sound levels was
reasonably described as pairwise linear. To the best of our knowledge, this has not been explicitly
stated before. However, linear transformations are consistent with numerous previous
observations. For instance, the preferred azimuth or ITD of IC and MSO neurons tends to remain
constant across level (Aitkin and Martin 1987; Kuwada et al. 2011; Kuwada and Yin 1983; Pecka
et al. 2008; Pecka et al. 2010; Sterbing et al. 2003; Yin et al. 1986). Further, the halfwidth of ITD
tuning in IC and MSO neurons also tends to remain constant across level (Pecka et al. 2010; Yin et
al. 1986). Studies in anesthetized animals have shown that the rising portion of an ILD tuning
curve in response to pure tones often shifts with sound level (Irvine and Gago 1990; Semple and
Kitzes 1987), which would indicate a nonlinear transformation. However, the dependence of firing
rate on sound level takes different shapes for tone and noise stimuli (Aitkin 1991). In other studies
using noise stimuli, most LSO and IC neurons were found to have azimuth tuning curves that
exhibited only small changes in half-maximal azimuth with level (Delgutte et al. 1999; Tollin and
Yin 2002). Most similar to our study is that by Kuwada et al (2011), who measured azimuth tuning
curves in IC of unanesthetized rabbits in response to noise stimuli over the full 360° of azimuth.
They found that 85% of their sample of combined single and multiunits had both vector angle and
vector strength (similar to preferred azimuth and halfwidth, respectively) that remained constant
between 30 and 50 dB re: level threshold, consistent with a linear transformation. Tuning curves
became more dissimilar at 10 dB re: threshold, at which point responses were largely monaural.

The linear decoders in the present study map the pattern of firing rates across the
population of IC neurons onto a set of likelihoods represented in the responses of a hypothetical output layer, with each output-layer neuron representing the likelihood of a particular source azimuth (Fig. 4A). One might naïvely assume that, since a given output-layer neuron is associated with a particular azimuth, its azimuth tuning would be highly selective to that azimuth. Contrary to this assumption, we found that all output-layer neurons had similar, monotonic azimuth tuning, broadly sensitive to contralateral azimuths (Fig. 12A). This means that even though each output-layer neuron is “labeled” with a particular azimuth, the selectivity of each neuron is not apparent from its tuning to azimuth, but rather from the relationships among tuning curves of different output layer neurons. Auditory neurophysiologists have long sought neurons in the mammalian auditory pathway that would be sharply tuned to source location ever since the original proposal of a place code of sound location (Jeffress 1948). However, azimuthal selectivity tends to be broadly contralateral in IC neurons (Fig. 3; (Day and Delgutte 2013; Kuwada et al. 2011), remains relatively broad in primary auditory cortex, and only narrows to a small extent in belt auditory cortex (Mickey and Middlebrooks 2003; Woods et al. 2006; Zhou and Wang 2012). Our decoder results raise the interesting possibility that the firing rates of broadly-tuned neurons at some higher level in the auditory pathway may represent the likelihoods of specific sound source locations. There is a growing body of evidence that humans and animals make use of uncertainty in stimulus representations when making decisions (Ma and Jazayeri 2014), and would therefore need a neural representation of probabilities. For instance, imagine that the response profile of the output layer shown in Figure 12B was elicited by a threatening sound in the night coming from directly in front of a rabbit. In one trial the rabbit would estimate the sound at 0°, while in the other at −15°. In either trial, the likelihoods are very similar between −15, 0 and 15°, so the rabbit may make use of this uncertainty by avoiding all three locations altogether.
We considered two methods by which azimuth could potentially be decoded from IC activity patterns in the face of changes in sound level: a “combination” decoder where the likelihoods of specific combinations of source azimuth and level are explicitly represented, and a “transformation” decoder where the likelihood of source azimuth is computed based on knowledge (or parallel estimation) of the sound level. One drawback of the combination decoder is the necessity of a large number of hypothetical output-layer neurons coding for every possible azimuth-level combination. The transformation decoder requires many fewer neurons in the output layer, as well as fewer free parameters overall, but its drawback is the necessity of separate sound level information to appropriately adjust decoder weights and biases. This could possibly be achieved through synaptic adaptation and regulation of hyperpolarizing inhibitory currents, respectively, although it is not straightforward to conceive how these would be specifically modulated by sound level. Recent studies in IC have shown that the rising portion of rate-level functions (the ”dynamic range”) adapts within around 160 ms to better-encode sound levels presented more often (Dean et al. 2005; Dean et al. 2008; Rabinowitz et al. 2013). One interesting possibility is that adaptation to preceding sound levels may quickly normalize azimuth tuning curves (Carandini and Heeger 2012) such that level-dependent adjustments of decoder weights and biases would be unnecessary or at least less computationally demanding. While the present study looked specifically at level-dependent changes in azimuthal coding, spectral and temporal stimulus features undoubtedly also change the population response to azimuth (Goodman et al. 2013), and it will be interesting to characterize azimuthal coding in IC in the face of changes in these other features.
Acknowledgements

We thank Ken Hancock for software support, and Dan Goodman for a critical reading of an earlier version of the manuscript. Current address of Mitchell Day is Dept. of Biological Sciences, Ohio University, Athens, OH.

Grants

This work was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health (NIH) under award numbers R03 DC013388 (MLD), R01 DC002258 (BD) and P30 DC005209. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References


Belliveau LA, Lyamzin DR, and Lesica NA. The neural representation of interaural time


Jazayeri M, and Movshon JA. Optimal representation of sensory information by neural


Middlebrooks JC, Xu L, Eddins AC, and Green DM. Codes for sound-source location in


Semple MN, and Kitzes LM. Binaural processing of sound pressure level in the inferior


Zhou Y, and Wang X. Level dependence of spatial processing in the primate auditory cortex. *J
Figure 1. Accuracy and precision of method of estimating mutual information (MI).

(A) Tuning curves (mean ± SD) either used to create simulated Poisson data (thin black line) or from an experimental data set with 30 trial repetitions (thick gray line). Positive azimuth indicates location contralateral to recording site. (B) MI (mean ± SD) computed from each of 500 simulated (thin black line) or subsampled (thick gray line) data sets using the specified number of trials. Exact MI for simulated data (solid gray line). MI computed from all 30 trials of experimental data (dashed gray line).

Figure 2. Cross-level comparison of the MI between azimuth and spike count for individual neurons.

Each circle marks the MI values of one neuron at the sound levels indicated on the x- and y-axes (N = 60). Diagonal lines indicate equality lines. Gray and black triangles on the right-hand side indicate median MI of the lower and higher levels, respectively. No significant effect of sound level on median MI (p = 0.89, Friedman’s test).

Figure 3. Changes in azimuth tuning curves with changes in sound level.

Azimuth tuning curves at 35, 50 and 65 dB SPL for five different neurons. SR indicates spontaneous firing rate. Best frequency (BF) listed in upper left, and MI at each level listed to the right of each panel. (A) Large MI at all levels. (B) Becomes informative above 35 dB SPL. (C) MI
increases greatly above 50 dB SPL, but responsive to sound at lower levels. (D) MI decreases above 35 dB SPL due to rate saturation. (E) MI decreases above 35 dB SPL due to flattening to a nonmaximal rate.

Figure 4. Cross-level comparison of the performance of a neural population decoder of azimuth. (A) Decoder schematic. Each output-layer neuron points to a particular azimuth and receives a weighted and biased sum of spike counts over IC neurons. The response of the output layer is related to the likelihood of the IC spike count pattern being associated with each azimuth. Azimuth is then estimated to maximize likelihood. (B) Localization performance of decoder, tested separately on data measured at each sound level (N = 60 neurons). Bubble diameter indicates fraction of estimates made at each azimuth. Each column sums to one. RMS error across locations listed in lower right of each panel. (C) RMS error over bootstrap-resampled neural populations, compared across level (median and 95% confidence interval). No significant effect of sound level on RMS error (p > 0.05, paired bootstrap tests with correction for multiple comparisons).

Figure 5. Comparing neural data to simulated Poisson data. (A) Mean vs variance of spike count for every neuron, level and azimuth (N = 2,340). Black line indicates equality line. Colored lines mark the average variance in several geometrically spaced bins of mean spike count for data measured in response to either frozen (blue) or random (red) noise bursts across trials (see Methods). (B) Histogram of the log variance of spike counts for those distributions with mean spike counts falling within a bin centered at 18 (dashed line in panel A). Black curve marks the histogram generated from simulated data in which spike counts were
drawn from Poisson distributions with repetitions and means that matched those of the measured azimuth tuning curves. Colored curves mark histograms generated from measured data using either frozen (blue) or random (red) noise bursts. All histograms were approximately Gaussian.

(C) Plot of the SD of the distribution of log variance across neurons and stimuli vs log mean. Dotted line marks the SDs of the histograms plotted in panel B. The SD of log variance of the measured data is always greater than that of the simulated data. (D) RMS decoder error over resampled neural populations, same as in Figure 4C except tested on simulated Poisson data (median and 95% confidence interval; paired bootstrap tests with correction for multiple comparisons; n.s. indicates p > 0.05). Median decoder error from experimental data (Fig. 4C) included in figure (green squares).

Figure 6. Fraction of neurons with MI at every sound level greater than a minimum MI value.

Figure 7. Dependence of MI on best frequency and sound level re: threshold.

(A) Each circle marks the normalized MI of a single neuron (N = 60), which is the ratio of MI to maximum possible MI given 15° azimuthal resolution. Gray curve is the average within each of 6 octave-wide bins of BF. Correlation assessed with Kendall’s rank correlation. (B) Sound level threshold plotted vs BF for all neurons with available data (N = 43). (C) MI values at 35, 50 and 65 dB SPL, plotted at the sound level with respect to each neuron’s level threshold. Gray curve is 10-dB-wide roving average.

Figure 8. Cross-level similarity of azimuth tuning curves.
Each circle marks the Pearson product-moment correlation between tuning curves of one neuron at two sound levels vs the minimum MI of that neuron between the two levels. Data is shown for every neuron and for all three cross-level pairs (N = 180). Y-axis is Fisher-transformed for clarity at high correlation values. Black curve is 0.25-bit-wide roving average. Open circles indicate statistically significant correlation (p < 0.05). Association between correlation coefficient and minimum MI assessed with Kendall’s rank correlation.

Figure 9. Relationship between linear transformation type and change in information.

(A) Example of an additive transformation between azimuth tuning curves at 50 and 65 dB SPL (mean ± SD). SR indicates spontaneous rate. (B) Same data as in (A). Each circle marks the firing rates at one azimuth at the two sound levels. Fit line, correlation, gain and bias found by linear regression. (C) Stacked histograms of the change in MI from lower to higher sound level for all cross-level tuning curve pairs with significant, positive correlation (N = 133), grouped by type of linear transformation. Transformation type determined by significance of gain and bias (p < 0.05, t-test): gain only (multiplicative), bias only (additive), gain and bias (mixed), or neither (no change). (D) Example of a multiplicative transformation of tuning between 35 and 65 dB SPL.

Figure 10. Neural population decoding of the combination of azimuth and sound level.

(A) Schematic for “combination” decoder. Same as in Fig. 4A except each output-layer neuron points to a particular combination of azimuth and sound level. Azimuth and level combination chosen to maximize likelihood. (B) Performance of decoder (N = 60 neurons), same as in Fig. 4B except azimuthal locations at each level are concatenated on the x- and y-axes. RMS error across
Figure 11. Cross-level neural population decoding of azimuth using knowledge of level-dependent transformations.

“Transformation” decoder schematic is same as in Fig. 4A, except with level-dependent weights and biases (see text for details). (A) Localization performance of transformation decoder, trained on data across all sound levels and tested separately at each level (N = 60 neurons). RMS error across locations listed in lower right of each panel. (B) RMS error across resampled populations, compared between combination and transformation decoders separately at each level (median and 95% confidence interval). Significant differences between combination and transformation decoder errors determined by paired bootstrap tests with correction for multiple comparisons (n.s. indicates p > 0.05).

Figure 12. Response of decoder output layer.

(A) Responses of three output layer neurons corresponding to three different azimuths (Fig. 4A) as source azimuth is varied (mean ± SD over 1,000 decoder repetitions at each source azimuth). Decoder trained and tested on data at 65 dB SPL. Output response in arbitrary units (a.u.). (B) Responses across output layer neurons to two different trials of a stimulus presented at a source azimuth of 0° (shaded area in panel A). Arrows point to the output layer azimuth with greatest response (i.e., maximum likelihood), with solid and dotted lines matched to each trial.
Figure 13. Schematic summary of cross-level encoding of azimuth in the IC.

(A) Black curve represents the MI of one neuron as sound level is varied ("information-level" curve). Azimuth tuning becomes significant at some level above level threshold. Tuning undergoes linear transformations with increases in level. Scaling (or scaling and shifting) upwards increases MI. Shifting upwards, or scaling downwards and shifting upwards, decreases MI. Some tuning functions ultimately flatten to a maximal or nonmaximal firing rate, but others may not (dashed curve). (B) Information-level curves of many different neurons, distinguished by color. Curves have different thresholds, widths, maxima, etc. (C) Population azimuthal information grows to a constant above detection threshold. Whether this plateau remains at higher levels is unknown.
Figure 1. Accuracy and precision of method of estimating mutual information (MI).

(A) Tuning curves (mean ± SD) either used to create simulated Poisson data (thin black line) or from an experimental data set with 30 trial repetitions (thick gray line). Positive azimuth indicates location contralateral to recording site. (B) MI (mean ± SD) computed from each of 500 simulated (thin black line) or subsampled (thick gray line) data sets using the specified number of trials. Exact MI for simulated data (solid gray line). MI computed from all 30 trials of experimental data (dashed gray line).
Figure 2. Cross-level comparison of the MI between azimuth and spike count for individual neurons. Each circle marks the MI values of one neuron at the sound levels indicated on the x- and y-axes (N = 60). Diagonal lines indicate equality lines. Gray and black triangles on the right-hand side indicate median MI of the lower and higher levels, respectively. No significant effect of sound level on median MI ($p = 0.89$, Friedman’s test).
Figure 3. Changes in azimuth tuning curves with changes in sound level.
Azimuth tuning curves at 35, 50 and 65 dB SPL for five different neurons. SR indicates spontaneous firing rate. Best frequency (BF) listed in upper left, and MI at each level listed to the right of each panel. (A) Large MI at all levels. (B) Becomes informative above 35 dB SPL. (C) MI increases greatly above 50 dB SPL, but responsive to sound at lower levels. (D) MI decreases above 35 dB SPL due to rate saturation. (E) MI decreases above 35 dB SPL due to flattening to a nonmaximal rate.
Figure 4. Cross-level comparison of the performance of a neural population decoder of azimuth.

(A) Decoder schematic. Each output-layer neuron points to a particular azimuth and receives a weighted and biased sum of spike counts over IC neurons. The response of the output layer is related to the likelihood of the IC spike count pattern being associated with each azimuth. Azimuth is then estimated to maximize likelihood.

(B) Localization performance of decoder, tested separately on data measured at each sound level (N = 60 neurons). Bubble diameter indicates fraction of estimates made at each azimuth. Each column sums to one. RMS error across locations listed in lower right of each panel. (C) RMS error over bootstrap-resampled neural populations, compared across level (median and 95% confidence interval). No significant effect of sound level on RMS error (p > 0.05, paired bootstrap tests with correction for multiple comparisons).
Figure 5. Comparing neural data to simulated Poisson data.

(A) Mean vs variance of spike count for every neuron, level and azimuth (N = 2,340). Black line indicates equality line. Colored lines mark the average variance in several geometrically spaced bins of mean spike count for data measured in response to either frozen (blue) or random (red) noise bursts across trials (see Methods).

(B) Histogram of the log variance of spike counts for those distributions with mean spike counts falling within a bin centered at 18 (dotted line in panel A). Black curve marks the histogram generated from simulated data in which spike counts were drawn from Poisson distributions with repetitions and means that matched those of the measured azimuth tuning curves. Colored curves mark histograms generated from measured data using either frozen (blue) or random (red) noise bursts. All histograms were approximately Gaussian.

(C) Plot of the SD of the distribution of log variance across neurons and stimuli vs log mean. Dotted line marks the SDs of the histograms plotted in panel B. The SD of log variance of the measured data is always greater than that of the simulated data.

(D) RMS decoder error over resampled neural populations, same as in Figure 4C except tested on simulated Poisson data (median and 95% confidence interval; paired bootstrap tests with correction for multiple comparisons; n.s. indicates p > 0.05). Median decoder error from experimental data (Fig. 4C) included in figure (green squares).
Figure 6. Fraction of neurons with MI at every sound level greater than a minimum MI value.
Figure 7. Dependence of MI on best frequency and sound level re: threshold.
(A) Each circle marks the normalized MI of a single neuron (N = 60), which is the ratio of MI to maximum possible MI given 15° azimuthal resolution. Gray curve is the average within each of 6 octave-wide bins of BF. Correlation assessed with Kendall’s rank correlation. (B) Sound level threshold plotted vs BF for all neurons with available data (N = 43). (C) MI values at 35, 50 and 65 dB SPL, plotted at the sound level with respect to each neuron’s level threshold. Gray curve is 10-dB-wide roving average.
Figure 8. Cross-level similarity of azimuth tuning curves.
Each circle marks the Pearson product-moment correlation between tuning curves of one neuron at two sound levels vs the minimum MI of that neuron between the two levels. Data is shown for every neuron and for all three cross-level pairs (N = 180). Y-axis is Fisher-transformed for clarity at high correlation values. Black curve is 0.25-bit-wide roving average. Open circles indicate statistically significant correlation (p < 0.05). Association between correlation coefficient and minimum MI assessed with Kendall’s rank correlation.
Figure 9. Relationship between linear transformation type and change in information.
(A) Example of an additive transformation between azimuth tuning curves at 50 and 65 dB SPL (mean ± SD). SR indicates spontaneous rate. (B) Same data as in (A). Each circle marks the firing rates at one azimuth at the two sound levels. Fit line, correlation, gain and bias found by linear regression. (C) Stacked histograms of the change in MI from lower to higher sound level for all cross-level tuning curve pairs with significant, positive correlation (N = 133), grouped by type of linear transformation. Transformation type determined by significance of gain and bias (p < 0.05, t-test): gain only (multiplicative), bias only (additive), gain and bias (mixed), or neither (no change). (D) Example of a multiplicative transformation of tuning between 35 and 65 dB SPL.
Figure 10. Neural population decoding of the combination of azimuth and sound level.
(A) Schematic for “combination” decoder. Same as in Fig. 4A except each output-layer neuron points to a particular combination of azimuth and sound level. Azimuth and level combination chosen to maximize likelihood. (B) Performance of decoder (N = 60 neurons), same as in Fig. 4B except azimuthal locations at each level are concatenated on the x- and y-axes. RMS error across within-level locations listed for each level.
Figure 11. Cross-level neural population decoding of azimuth using knowledge of level-dependent transformations.

"Transformation" decoder schematic is same as in Fig. 4A, except with level-dependent weights and biases (see text for details). (A) Localization performance of transformation decoder, trained on data across all sound levels and tested separately at each level (N = 60 neurons). RMS error across locations listed in lower right of each panel. (B) RMS error across resampled populations, compared between combination and transformation decoders separately at each level (median and 95% confidence interval). Significant differences between combination and transformation decoder errors determined by paired bootstrap tests with correction for multiple comparisons (n.s. indicates p > 0.05).
Figure 12. Response of decoder output layer.
(A) Responses of three output layer neurons corresponding to three different azimuths (Fig. 4A) as source azimuth is varied (mean ± SD over 1,000 decoder repetitions at each source azimuth). Decoder trained and tested on data at 65 dB SPL. Output response in arbitrary units (a.u.). (B) Responses across output layer neurons to two different trials of a stimulus presented at a source azimuth of 0° (shaded area in panel A). Arrows point to the output layer azimuth with greatest response (i.e., maximum likelihood), with solid and dotted lines matched to each trial.
Figure 13. Schematic summary of cross-level encoding of azimuth in the IC.
(A) Black curve represents the MI of one neuron as sound level is varied ("information-level" curve). Azimuth tuning becomes significant at some level above level threshold. Tuning undergoes linear transformations with increases in level. Scaling (or scaling and shifting) upwards increases MI. Shifting upwards, or scaling downwards and shifting upwards, decreases MI. Some tuning functions ultimately flatten to a maximal or nonmaximal firing rate, but others may not (dashed curve). (B) Information-level curves of many different neurons, distinguished by color. Curves have different thresholds, widths, maxima, etc. (C) Population azimuthal information grows to a constant above detection threshold. Whether this plateau remains at higher levels is unknown.