Effective sound localization coding by noisy populations of mouse inferior colliculus neurons revealed by fast volumetric imaging

Juan C. Boffi[1,3], Brice Bathellier[2], Hiroki Asari[3] & Robert Prevedel[1,3-6]

Affiliations

[1] Cell Biology and Biophysics Unit, European Molecular Biology Laboratory, Heidelberg, Germany.

[2] Institut Pasteur, Université Paris-Cité, INSERM, Institut de l'Audition, 63 rue de Charenton, F-75012 Paris, France.

[3] Epigenetics and Neurobiology Unit, European Molecular Biology Laboratory, Monterotondo, Italy.

[4] Developmental Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany.

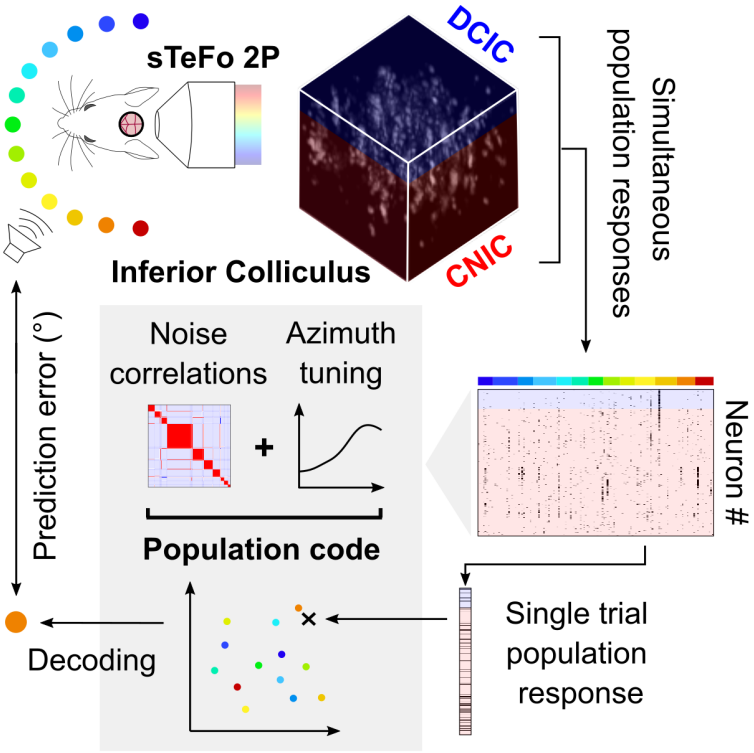
[5] Molecular Medicine Partnership Unit, European Molecular Biology Laboratory, Heidelberg, Germany.

[6] Interdisciplinary Center for Neurosciences, Heidelberg University, Heidelberg, Germany.

Correspondence should be addressed to J.C.B. (juan.boffi@embl.de) and R.P. (prevedel@embl.de).

Abstract

The mammalian inferior colliculus (IC) is thought to encode sound location. However, large-scale simultaneous recordings are lacking to characterise the underlying neuronal population code. Using volumetric two-photon calcium imaging in awake mice, here we simultaneously recorded the activity of ~1,000 IC neurons in response to sounds delivered from 13 different frontal horizontal locations (azimuths). The population responses showed marked trial-to-trial variability and revealed the recruitment of distinct groups of noise-correlated IC neurons as a function of stimulus azimuth. We tested decoding model performance on our data and found that template matching outperformed previously proposed models, where noise correlations helped improve the decoding accuracy. Finally, our volumetric imaging approach revealed that both the central nucleus and dorsal cortex of IC contribute to sound location coding. These findings indicate that noise correlations, together with azimuth tuning, play a key role in shaping the population code for sound location in the mammalian IC.



Introduction

Locating a sound source facilitates essential behaviours such as foraging, mating and predator avoidance, thanks to the omni-directionality and long reach of the auditory system (King et al., 2001). In vertebrates, sound localization relies on both binaural cues such as interaural level and time differences (ILD, ITD) (Knudsen and Konishi, 1979) and monaural cues including spectral notches (SN) (Kulkarni and Colburn, 1998). These sound localization cues are processed independently at brainstem nuclei in the ascending auditory pathway and integrated altogether for the first time at the inferior colliculus (IC) (Adams, 1979; Brunso-Bechtold et al., 1981; Felix et al., 2018; Grothe et al., 2010). This makes the IC a crucial early relay of the ascending auditory pathway to study how a primary neural representation of auditory space is formed (Felix et al., 2018; Grothe et al., 2010).

Previous studies involving extracellular recordings from the mammalian IC revealed that average responses from IC neurons were tuned to multiple sound location cues, evidencing the integration of such cues at IC (Chase and Young, 2005, 2008). However, the amount of sound location information carried by such individual neurons throughout the auditory pathway is limited and quite variable (Clarey et al., 1995; Day and Delgutte, 2016; Panniello et al., 2018), which argues in favour of a population code for auditory space. Fundamental work by Day and Delgutte (2013) supports that sound location in the horizontal plane (azimuth) is represented at the mammalian IC by the activity patterns of populations of neurons that display average response tuning to azimuth, while ruling out other possibilities such as a topological code, an opposing two channel model or a population vector code (Georgopoulos et al., 1982, 1986). Due to the limitations of traditional *in vivo* microelectrode techniques, this knowledge stems from prospective IC

population activity patterns that were approximated by pooling non-simultaneous single unit recordings, under the assumption of independent firing between cells. However, it remains an open question whether the actual activity patterns occurring in the IC of awake animals can encode sound location given emerging properties such as correlated neuronal noise (Averbeck et al., 2006; Kohn et al., 2016). Noise correlation and trial-to-trial response variability can have a profound impact on the representation of sound location by IC population activity patterns, but it can only be determined by simultaneous recordings that capture them (Averbeck et al., 2006; Kohn et al., 2016). Finally, traditional microelectrode recordings were mostly performed at relatively deep anatomical subdivisions of IC, due to the technical difficulty to target superficial regions (Chase and Young, 2008; Day and Delgutte, 2013, 2016; Lesica et al., 2010; Schnupp and King, 1997). To overcome these limitations and explore the importance of trial-to-trial response variability, noise correlation and IC anatomical subdivisions, we relied on volumetric Ca^{2+} imaging (Prevedel et al., 2016) and monitored IC population response patterns that encode sound location in the single trial basis.

Recent technical developments, including multichannel silicon probes (Jun et al., 2017) and fast volumetric Ca^{2+} imaging (Prevedel et al., 2016; Stringer et al., 2019; Weisenburger et al., 2019) enable routine simultaneous recordings from large numbers of neurons *in vivo*. Because neurons sensitive to azimuth are distributed throughout the central nucleus of IC without any topographical arrangement or specific distribution (Day and Delgutte, 2013), fast volumetric imaging is the most promising approach to sample large numbers of them across an IC volume, due to its denser sampling ability. In this work, we implemented scanned temporal focusing two-photon microscopy (sTeFo 2P) (Prevedel et al., 2016), a cutting-edge fast volumetric Ca^{2+} imaging modality, to simultaneously record the activity of unprecedentedly large samples of IC neurons. This approach revealed that 1) IC populations recruit distinct ensembles of noise-correlated neurons as a function of stimulus azimuth; and 2) together with azimuth response tuning, this noise correlation shapes the population response patterns, across IC subdivisions to generate a representation of sound azimuth. These findings open new perspectives about unexpected features of population responses that can contribute to sensory information processing.

Results

Simultaneous *in vivo* recordings of IC population activity in response to sound azimuth

In this work, we implemented sTeFo 2P for fast volumetric Ca^{2+} imaging, which enabled simultaneous sampling of the activity from unprecedentedly large numbers of IC neurons. Moreover, TeFo based methods have been shown to be more resilient to highly scattering (optically opaque) tissues such as the IC (Dana and Shoham, 2012). Adeno-associated viral vector (AAV) induced IC neurons expressing jRGECO1a were imaged through a cranial window in awake head-fixed mice, passively listening to 500 ms-long broad-band noise stimuli (20-40 kHz white noise) delivered by a motorised speaker every 5 s located at one of 13 different frontal azimuth angles in a random order covering the frontal hemifield in 15º steps (Fig. 1A). We performed volumetric Ca^{2+} imaging (4 Hz volume rate) simultaneously sampling the activity of up to 1520 (on

average 753 ± 478, mean ± S.D., n = 11 mice) IC neurons within a 470x470x500 µm volume from the right IC unilaterally (Fig. 1B, C). To compensate for the relatively slow decay time of the Ca^{2+} indicator signal and non-linearities in the summation of Ca^{2+} events, we evaluated the spike probabilities associated to the neuronal Ca^{2+} signals (Deneux et al., 2016) (Fig. 1D). The spike probability estimation does not involve predicting spike timings (Deneux et al., 2016; Huang et al., 2021; Pachitariu et al., 2018; Rupprecht et al., 2021; Vanwalleghem et al., 2021). Hence, we do not make any observation or interpretation in relation to the fine temporal structure of neuronal spiking. Instead, we treated the spike probabilities as a measure proportional to the spike count evaluated within the time interval set by our volumetric imaging rate (4 Hz). Taking into account that the main carrier of sound location information is IC neuron firing rate (Chase and Young, 2008), we defined the evoked IC population responses as the mean spike probabilities of each recorded cell during the sound stimulation epoch (500 ms, Fig. 1D, E). In this way, we systematically collected the simultaneous response patterns of IC populations to every azimuth trial (Fig. 1E).

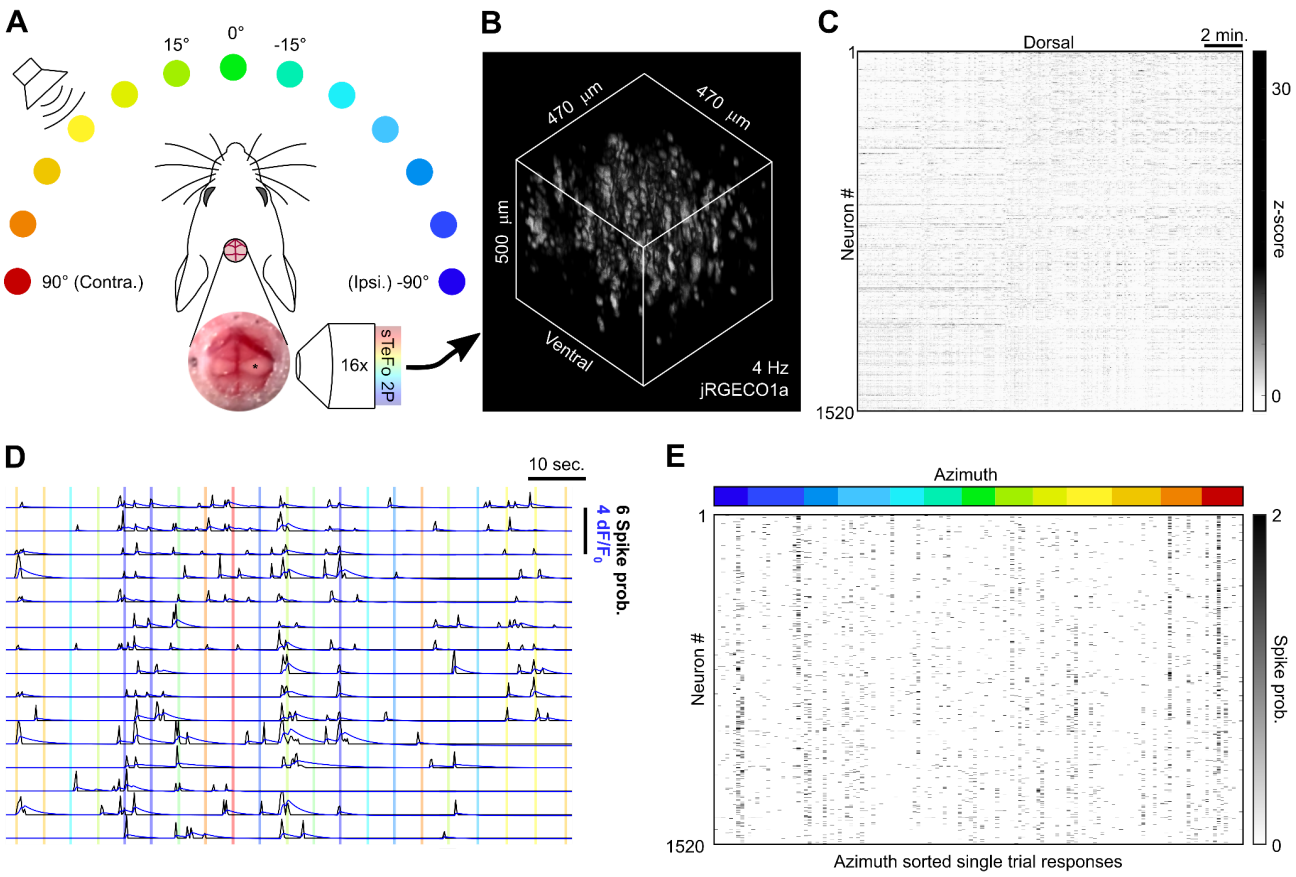


Figure 1: *In vivo* fast volumetric Ca^{2+} imaging of IC population activity. A) Schematic representation of the experimental design. B) Rendering of a sampled IC volume covering the jRGECO1a signal. C) Neuropil-corrected and denoised jRGECO1a signals extracted from a representative full dataset. Extracted signals are arranged from dorsal (top) to ventral (bottom) ROI position across the IC volume imaged. D) Representative neuropil corrected and denoised jRGECO1a traces (blue) with their corresponding spike probability traces (black) and stimulation epochs (colour-coded based on stimulus azimuth angle according to (A)) super-imposed. E) Population response vectors (in columns), sorted by their corresponding stimulus azimuth (colour-coded at the top) from the dataset in C).

Trial-to-trial variability, azimuth tuning and noise correlation.

The simultaneously recorded IC population activity from passively listening awake mice displayed sound-evoked response patterns that varied markedly across trials, as did the overall, average spike probability of the population (Fig. 2A). Moreover, when considering the same-azimuth trials, the simultaneously recorded responses also showed marked trial-to-trial variability both at the level of the population (Fig. 1E) and individual cells (Fig. 2B). To explore how trial-to-trial variability affects sound location representation by the IC populations, we evaluated azimuth tuning. Using one-way ANOVA (Kruskal-Wallis test), we identified IC neurons whose responses significantly depended on stimulus azimuth, which we defined as azimuth tuned neurons (Fig. 2B). Nevertheless, we qualitatively observed that neurons that were not significantly tuned to azimuth (N.S. tuned) could still display some response selectivity to certain azimuths (Fig. 2C). To compare the response selectivities of tuned and N.S. tuned neurons, we evaluated the ratio between the mean (signal) and the standard deviation (noise) of the responses to the best azimuth for each neuron (i.e. the azimuth evoking its maximal responses), which we refer to as a neuronal signal-to-noise ratio (nS/N ratio). Both tuned and N.S. tuned neurons showed generally low nS/N ratios, pointing out that trial-to-trial variability has a prominent influence on IC neuron responses in awake, passively listening mice (Fig. 2D). The only quantitative difference observed between the azimuth selectivities of the tuned and N.S. tuned neurons was a slight but statistically significant shift in the nS/N ratio distribution of tuned neurons towards larger values with respect to the distribution of N.S. tuned neurons (Fig. 2D). Neurons that displayed significant azimuth tuning represented 9 ± 3% (mean ± S.D., n = 8 mice, broad-band noise stimulus) of the total number of neurons imaged simultaneously in one session, which is close to, but significantly above, the false positive detection rate of our ANOVA tests (α = 0.05, Suppl. Fig. 1A). Altogether, these results suggest that the occurrence of statistically significant azimuth tuning and the distinction between tuned and N.S. tuned neurons are not biologically meaningful due to the observed trial-to-trial variability in the responses. Instead, the influence of trial-to-trial variability on single-cell responses would argue in favour of a population representation of sound azimuth by both tuned and N.S. tuned neurons altogether.

An important factor affecting population coding under a large trial-to-trial response variation is neuronal noise correlation (Averbeck et al., 2006; Kohn et al., 2016). Here we defined neuronal noise as the difference between a single-trial response and the average response to the corresponding azimuth. Due to the large trial-to-trial response variability, the average response amplitudes were generally small (Fig. 2B, C), resulting in mostly positive or close to baseline values of neuronal noise when subtracted from single trial responses (see methods). We evaluated Pearson's correlation coefficient between the noise activity of all possible pairs of simultaneously recorded IC neurons. We found that the noise correlations were not homogeneous across the population, with some groups of cells showing relatively high noise correlations between each other and low with the rest (Fig. 2E, F). Importantly, tuned and N.S. tuned neurons did not form distinct noise-correlated groups but were intermingled as is evidenced with a hierarchical clustering of their noise activity patterns (Fig. 2E). This again suggests that the distinction between the tuned and N.S. tuned neurons is likely not biologically meaningful. Interestingly, the clusters of noise correlated IC neurons

did not show any particular trend in their azimuth sensitivity (Fig. 2F). The pairwise noise correlation coefficients of simultaneously imaged IC neurons yielded a bimodal distribution across mice (Fig. 2G). To determine if the observed noise correlation patterns were linked to the processing of the sound stimuli, we studied the correlation structure of the on-going IC population activity during the inter-trial periods without sound stimulation. We observed that the on-going IC population activity did not show clear-cut clusters of correlated neurons as observed in the noise activity during sound stimulations (Suppl Fig 1B, C), yielding a unimodal distribution across mice (Fig 2G). This supports that the observed noise correlation patterns are part of the evoked responses and likely linked to the processing of the sound stimulus. The segmented footprints of highly noise-correlated cell clusters scattered across large distances along the field of view and imaging depths, arguing against the possibility that the observed high noise correlations arise from over-segmentation of single cells (Suppl. Fig. 1D, E). We did not observe any particular spatial segregation between the detected clusters of noise correlated neurons across the imaged IC volumes (Suppl. Fig. 1D, E). These observations support the presence of correlated noise in IC population activity, which violates the common assumption of independent firing made in previous studies (Day and Delgutte, 2013, 2016). Moreover, our result indicates that the occurrence of noise correlations might be more complex than expected (Day and Delgutte, 2013; Lesica et al., 2010), as the observed bimodal distribution suggests that the average pairwise correlation coefficient is not a good estimator for characterising the noise correlation structure.

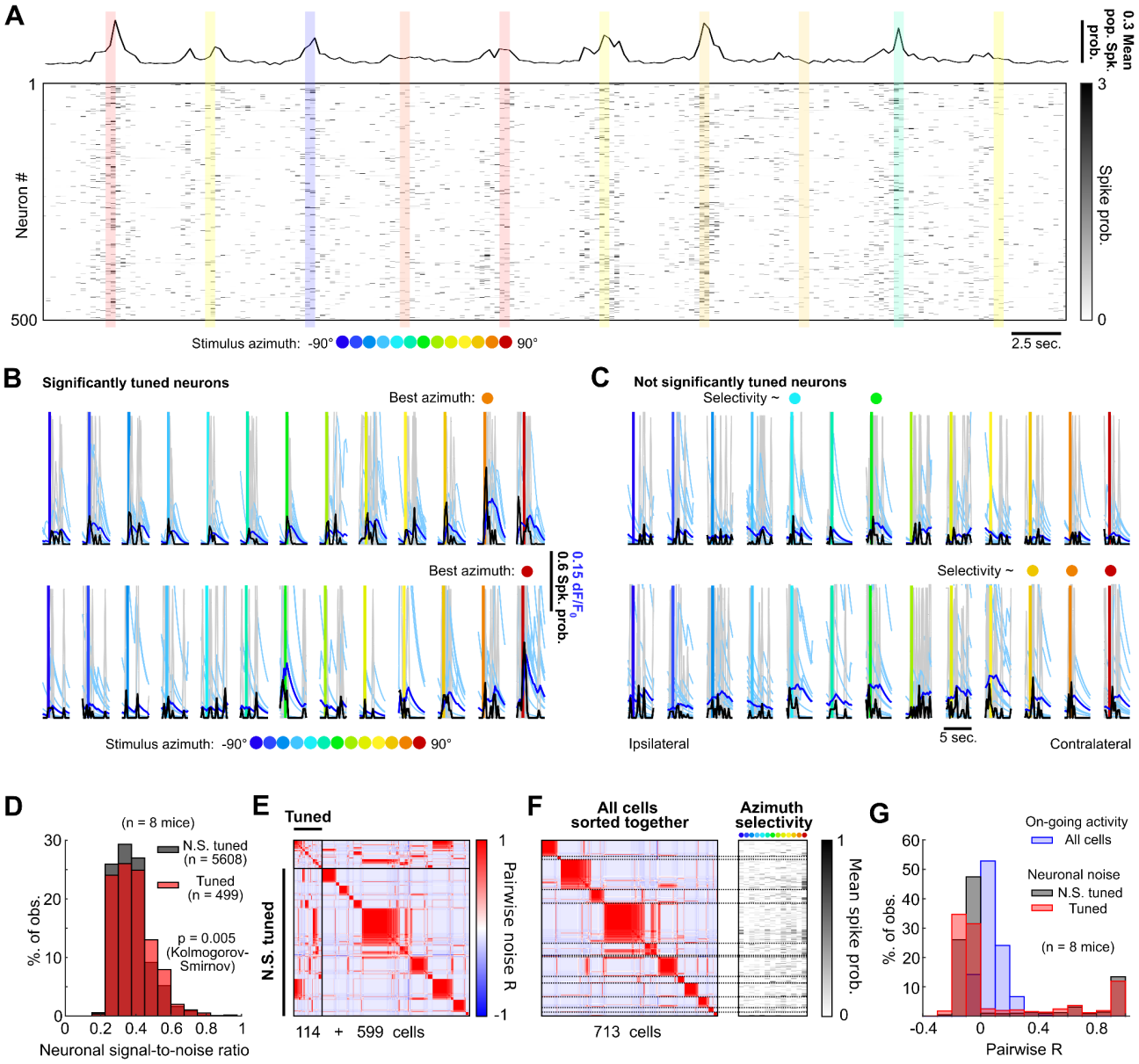
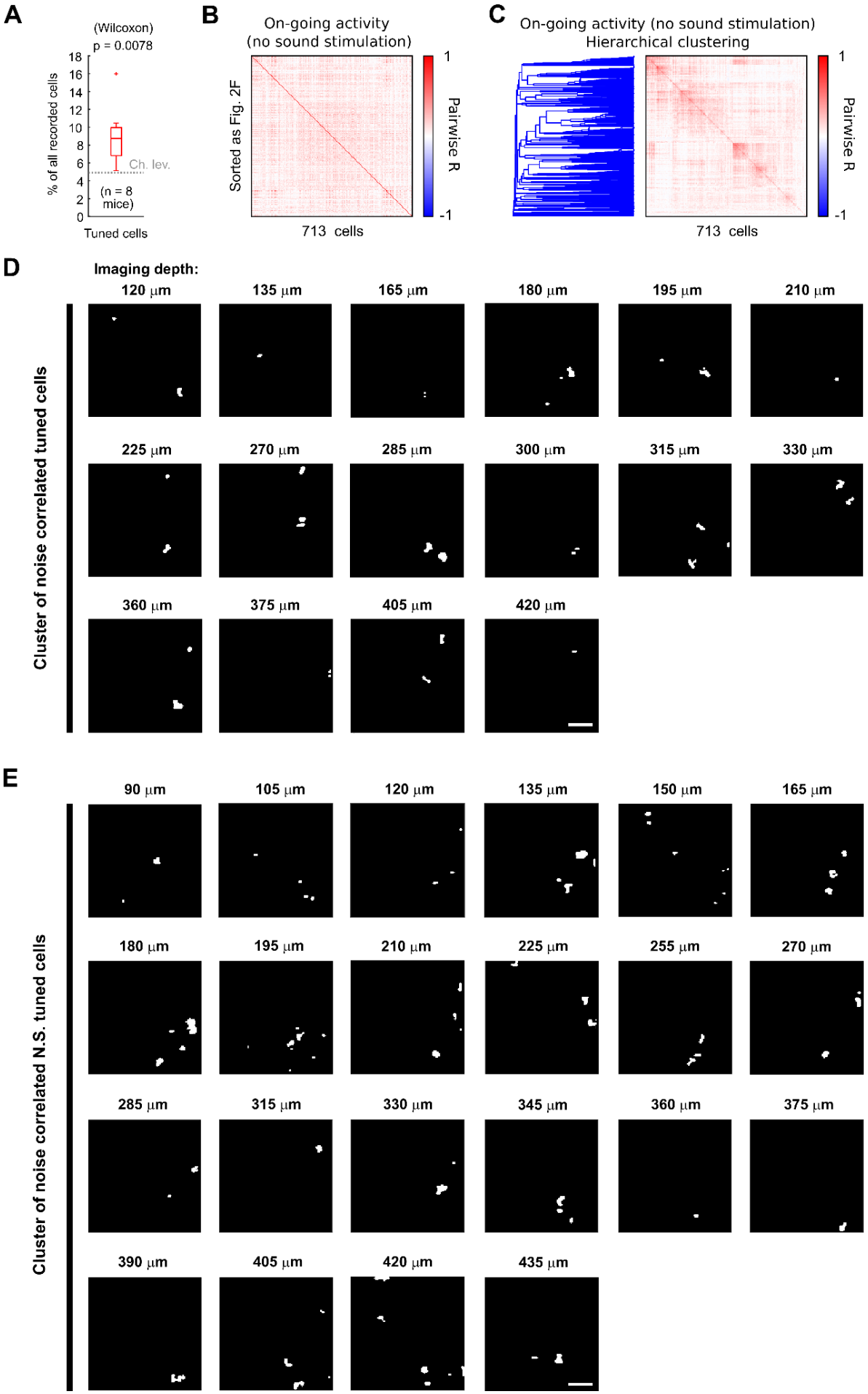


Figure 2: Trial-to-trial variability and azimuth tuning. A) Representative sound-evoked IC population response patterns. Top trace is the population average response. Sound stimulation epochs are colour-coded based on azimuth. B, C) Representative traces of detected IC neurons with significantly tuned responses as a function of stimulus azimuth (ANOVA, azimuth tuned neurons) and not significantly (N.S.) tuned neurons showing response selectivity. The Ca^{2+} signal responses in individual trials are shown in light blue and the average across the same-azimuth trials in dark blue. Grey traces show predicted spike probability traces in single trials. Black traces show the average spike probability across the same-azimuth trials. D) Histogram of the nS/N ratios from the best-azimuth responses of tuned and N.S. tuned neurons, estimated as the inverse of their coefficients of variation. E-F) Representative correlation matrices of pairwise noise correlations detected in simultaneous recordings where tuned and N.S. tuned neurons are considered separately (E) or all cells are sorted together (F; their azimuth selectivity is shown on the right). G) Histograms of pairwise correlations in on-going activity (without sound stimulation, blue) and neuronal noise from tuned (red) and N.S. tuned (grey) neurons simultaneously sampled across mice.


A
(Wilcoxon)
p = 0.0078
% of all recorded cells
Ch. lev.
(n = 8 mice)
Tuned cells
B
On-going activity
(no sound stimulation)
Sorted as Fig. 2F
Pairwise R
713 cells
C
On-going activity (no sound stimulation)
Hierarchical clustering
Pairwise R
713 cells
D
Imaging depth:
Cluster of noise correlated tuned cells
120 µm
135 µm
165 µm
180 µm
195 µm
210 µm
225 µm
270 µm
285 µm
300 µm
315 µm
330 µm
360 µm
375 µm
405 µm
420 µm
E
Cluster of noise correlated N.S. tuned cells
90 µm
105 µm
120 µm
135 µm
150 µm
165 µm
180 µm
195 µm
210 µm
225 µm
255 µm
270 µm
285 µm
315 µm
330 µm
345 µm
360 µm
375 µm
390 µm
405 µm
420 µm
435 µm

Supplemental figure 1: A) Percentage of detected tuned neurons across mice. B) Representative correlation matrix from the on-going IC population activity, in the absence of sound stimuli, recorded from the inter-trial epochs from the dataset shown in Fig. 2E, F. The cells are sorted as in Fig 2F. C) Representative hierarchical clustering dendrogram (left) depicting the clustered sorting order of the on-going activity correlation matrix (right, same data as in B). D, E) Scattering of the spatial footprints of segmented cell somas from a representative noise correlated cluster of simultaneously recorded tuned (D) and N.S. tuned (E) neurons across the imaged volume. The depicted footprints correspond to the largest clusters in Fig. 2E respectively. Scale bars: 100 µm.

How are these noise correlation patterns linked to stimulus azimuth processing? To explore this question, we compared the clustering of noise correlated neurons across the azimuths tested, and found that different groups of neurons were recruited for each stimulus azimuth (Fig. 3A). To have a quantitative estimate of how much do the groups of noise correlated neurons change in relation to azimuth, we evaluated the cosine distance between the sorted noise correlation matrices as a measure inversely proportional to their similarity (Fig. 3A, B). We found that the mean cosine distance between the sorted noise correlation matrices across mice was quite large for all the stimulus azimuths but still not as extreme as in decorrelated neuronal noise samples obtained by scrambling the noise from each cell across same-azimuth trials (Fig. 3C, D). These observations suggest that noise correlations in IC population responses are linked to stimulus azimuth through the recruitment of markedly distinct groups or ensembles of noise correlated neurons.

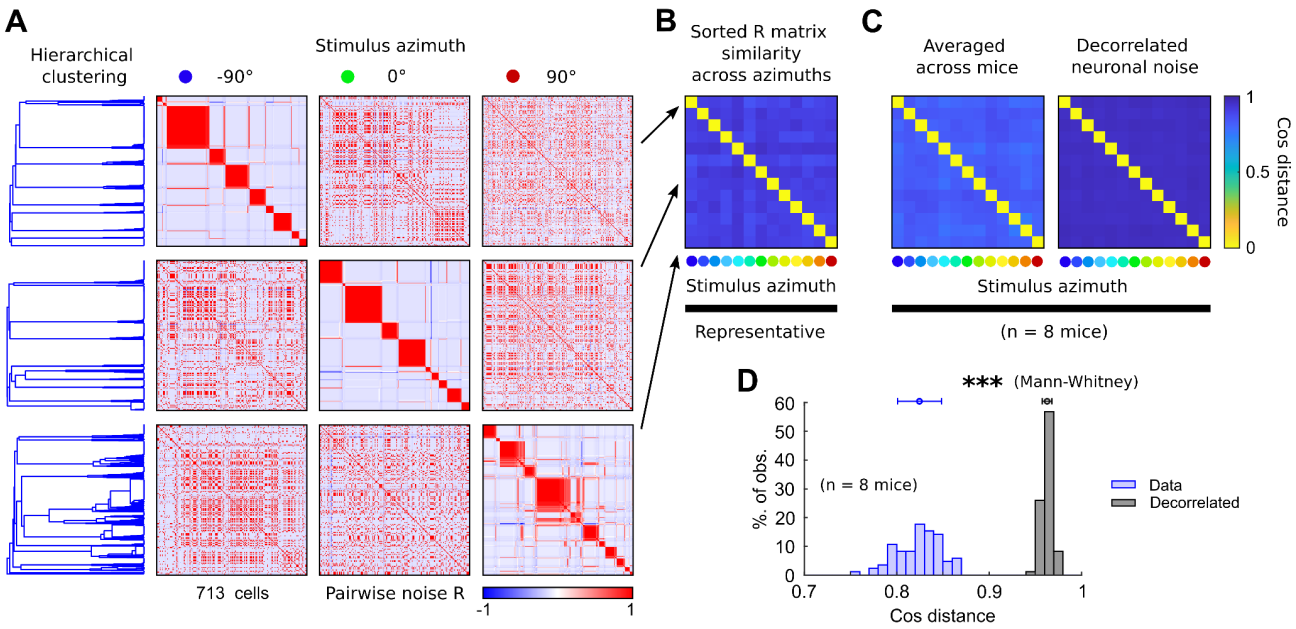


Figure 3: Noise correlations in relationship to stimulus azimuth. A) Representative hierarchical clustering dendrograms (left) showing the sorting order of the representative noise correlation matrices (right) as a function of stimulus azimuth (rows of matrices), from the same dataset shown in Fig. 2E, F. B) Pairwise cosine distances between the sorted noise correlation matrices following the clusterings for each azimuth. Each row/column in the chart corresponds to comparisons across a row of matrices in panel (A) (arrows). C) Mean cosine distance between the sorted noise correlation matrices following the clusterings for each azimuth across mice (left) and for the same datasets with scrambling of each cell's neuronal noise across same-azimuth trials to break noise correlations (decorrelated, right). D) Histogram of the mean pairwise cosine distances plotted in (C). *** $p < 0.0005$, Mann-Whitney test.

Sound location encoding and decoding by IC populations

The high trial-to-trial variability in our simultaneous recordings raised the question of to what extent the responses of the recorded neurons carry information about stimulus azimuth. To determine this, we evaluated how the collected single-trial IC population response patterns (Fig. 1E) could be associated with stimulus azimuth. First, we performed dimensionality reduction through linear discriminant analysis (LDA). This supervised method of dimensionality reduction allowed us to define dimensions along which the population responses should segregate maximally, in contrast to unsupervised methods such as the covariance-based principal component analysis (Fig. 4A). The scatter of single-trial population responses in the low dimensional LD space allowed us to visualise how well a few azimuth categories differentiate between each other (e.g. segregation of three azimuth categories in 2-D LD space), qualitatively showing that these few categories segregate from each other with some overlap (Fig. 4B). To observe the separation in single-trial population responses across all the 13 tested azimuth categories, we examined their scatter in a 13-D LD space. Specifically, we evaluated a 2-D uniform manifold approximation and projection (UMAP) (McInnes et al., 2018) of the complete 13-D LD space, and found that single-trial population responses to same-azimuth trials did cluster and segregate from the rest for some azimuths, but overlapped to different extents for the other azimuths (Fig. 4C). This qualitative observation suggests that single-trial population responses might not be fully separable by linear boundaries as a function of stimulus azimuth (linear separability).

To quantitatively assess how effectively these slightly overlapping population responses could represent stimulus azimuth, we performed cross-validated, non-linear, multi-class classification of single-trial responses (decoding, Fig. 4D). We compared different classification algorithms as decoding models, including the previously proposed maximum likelihood decoder (Day and Delgutte, 2013; Jazayeri and Movshon, 2006) (MLD), template matching algorithms (MCD, KNN) and sophisticated machine learning algorithms such as ensembles of classifiers (SVM ens., D. trees ens.) and a fully connected feed-forward artificial neural network (FF aNN, Fig. 4E-G). To determine if these algorithms could effectively estimate stimulus azimuth from the recorded single-trial population responses, we evaluated the cross-validated prediction errors as the absolute difference between the stimulus azimuth and the predicted azimuth (Fig. 4E). The cross-validated prediction error showed asymmetrical distributions across all decoding algorithms, skewed towards small prediction errors, as predicted by chance, with deviations from the chance level distribution (determined by considering all the possible prediction errors for the 13 tested azimuths) towards more frequent high or low prediction errors (Fig. 4F). The mode of these distributions was 15° across decoders, which suggests that the most frequent prediction error is below the minimum azimuth angle that trained mice can discriminate in a behavioural task (minimum audible angle, 31°) (Lauer et al., 2011). This supports that sound azimuth decoding accuracy from the IC population responses can be frequently finer than the behavioral discrimination ability, on the single trial basis. To find out how effectively each of the different classification algorithms decoded the single-trial population responses, we determined the medians

of the cross-validated prediction error distributions across mice, which typically took the value of a tested azimuth (Fig. 4G, grey points and their scatter). We evaluated if the median cross-validated prediction errors across mice were significantly different from the chance level, defined by averaging all the possible prediction errors for the 13 tested azimuths. Out of all the algorithms tested, two classifiers/decoders showed median cross-validated prediction errors across mice that were significantly lower than chance level: the previously proposed MLD (Day and Delgutte, 2013), and a maximum correlation decoder model (Franco et al., 2004) (MCD, template matching algorithm, Fig. 4G). This result supports that the simultaneously recorded IC population response patterns contain information about stimulus azimuth that is sufficient for different, independent classification models to effectively decode them with an error smaller than chance level. Consistently, the prediction errors from the trial-scrambled data were not significantly different from chance level (Suppl. Fig. 2A).

Next, we determined the influence of azimuth tuning on decoding performance by comparing the median cross-validated prediction errors across mice obtained with the MLD and MCD operating on the tuned subpopulation and subsamples of N.S. tuned neurons of equal size to the tuned subpopulation (see methods). Additionally, we evaluated the effect of breaking noise correlations on the decoding performance by scrambling the responses from each cell in the simultaneously imaged populations across the same-azimuth trials (decorrelated population responses). We observed that the median cross-validated prediction errors across mice were significantly smaller than chance level for both the MLD and MCD independently of tuning or the presence of noise correlations, except for the MCD operating on the decorrelated complete population responses (Fig. 4H, I). These results point out that the response patterns from both tuned and N.S. tuned neurons carry sufficient information to effectively decode the stimulus azimuth, further challenging the distinction between these cells. Interestingly, noise correlations seemed to affect only the performance of the MCD operating on the complete populations of the sampled cells. On the other hand, the lack of influence of noise correlations on the MLD model might be a consequence of the way this model is formulated under the assumption of independent firing, which is violated in our data. On the other hand, the observed influence of noise correlations on MCD decoding error might arise only when decoding a large enough number of cells, such as the complete sampled population, but not the subsets, such as tuned or N.S. tuned subpopulations (Fig. 4I). Importantly, we observed that MCD decoding error is not significantly correlated with the total number of cells in the complete samples of imaged neurons or N.S. tuned subsets (Pearson's R, all cells: 0.32, $p = 0.44$; N.S. tuned: 0.61, $p = 0.11$; $n = 8$ mice), supporting sufficient sampling thanks to our volumetric imaging approach (Suppl. Fig. 3A). Nevertheless, decoding errors did significantly correlate to the number of tuned neurons decoded (Pearson's R, tuned: 0.52, $p = 0.018$, $n = 8$ mice, Suppl. Fig. 3A). This correlation was positive, opposite to the expected negative relationship between number of cells decoded and prediction error (Stringer et al., 2021). This unexpected positive correlation might be the consequence of potential biases due to under-sampling, further arguing against pre-selection criteria leading to small subsets of neurons for decoding. Altogether, these results point out that the MCD model operating on the responses from the complete sampled population of neurons is likely the most informative model from the ones tested in this study.

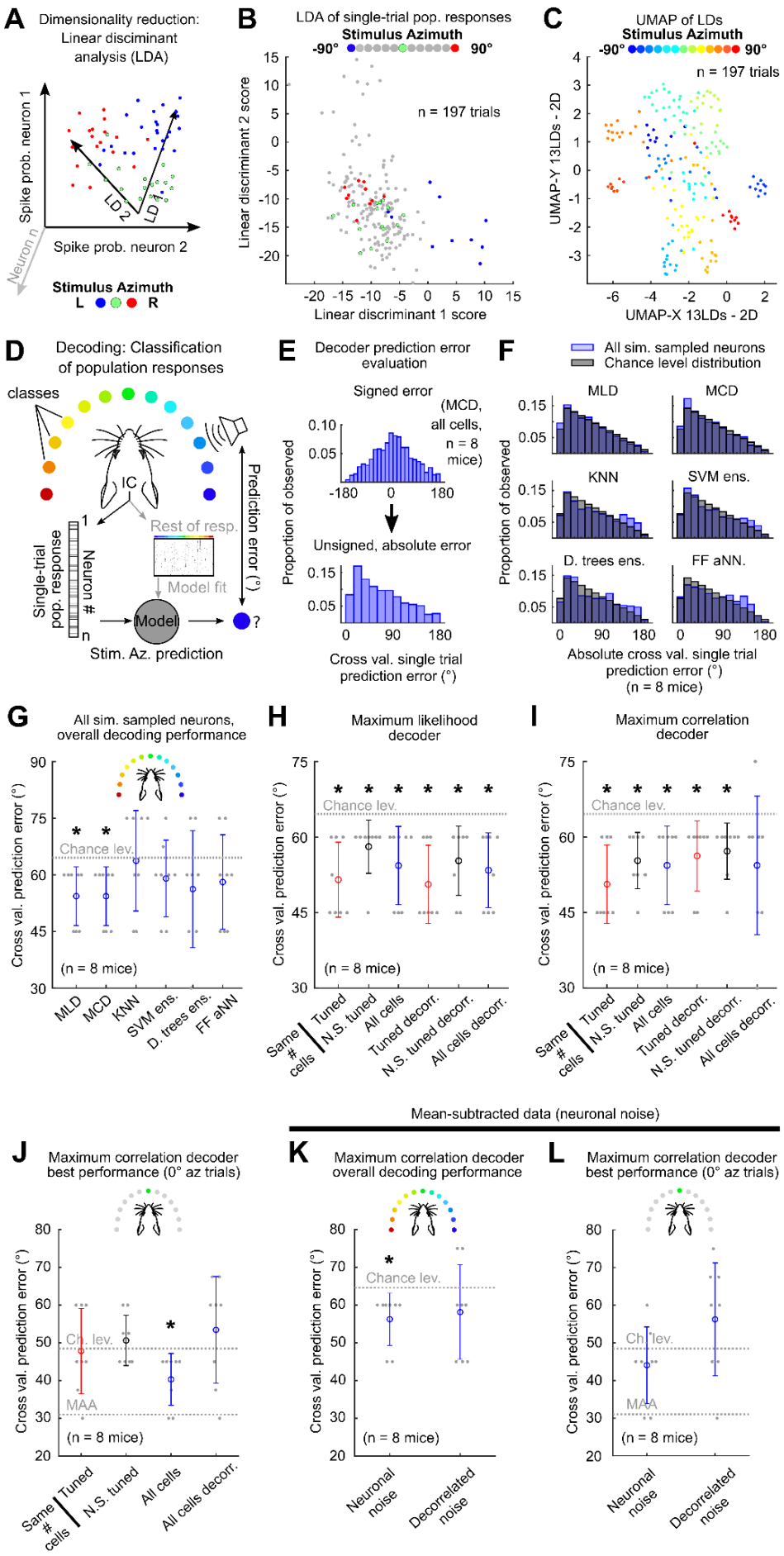
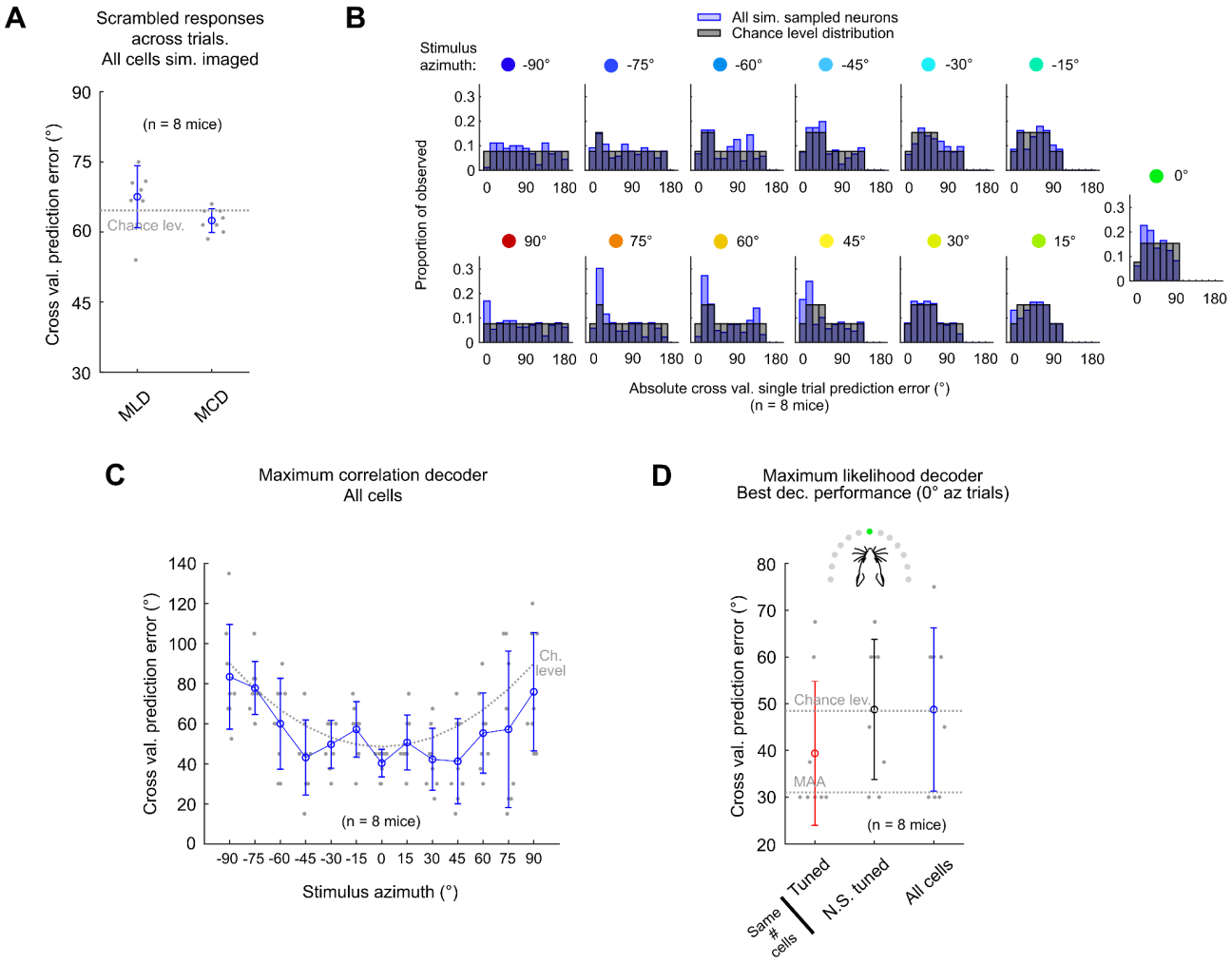

A
Dimensionality reduction: Linear discriminant analysis (LDA)
Spike prob. neuron 1
Spike prob. neuron 2
Neuron n
LD 1
LD 2
Stimulus Azimuth
L R
B
LDA of single-trial pop. responses
Stimulus Azimuth
-90° 90°
n = 197 trials
Linear discriminant 2 score
Linear discriminant 1 score
C
UMAP of LDs
Stimulus Azimuth
-90° 90°
n = 197 trials
UMAP-Y 13LDs - 2D
UMAP-X 13LDs - 2D
D
Decoding: Classification of population responses
classes
IC
Rest of resp.
Model fit
Model
Single-trial pop. response
Neuron #
Prediction error (°)
Stim. Az. prediction
E
Decoder prediction error evaluation
Signed error
(MCD, all cells, n = 8 mice)
Unsigned, absolute error
Proportion of observed
Cross val. single trial prediction error (°)
F
All sim. sampled neurons
Chance level distribution
MLD
MCD
KNN
SVM ens.
D. trees ens.
FF aNN.
Proportion of observed
Absolute cross val. single trial prediction error (°)
(n = 8 mice)
G
All sim. sampled neurons, overall decoding performance
Cross val. prediction error (°)
Chance lev.
(n = 8 mice)
MLD
MCD
KNN
SVM ens.
D. trees ens.
FF aNN
H
Maximum likelihood decoder
Cross val. prediction error (°)
Chance lev.
(n = 8 mice)
Same # cells
Tuned
N.S. tuned
All cells
Tuned decorr.
N.S. tuned decorr.
All cells decorr.
I
Maximum correlation decoder
Cross val. prediction error (°)
Chance lev.
(n = 8 mice)
Same # cells
Tuned
N.S. tuned
All cells
Tuned decorr.
N.S. tuned decorr.
All cells decorr.
Mean-subtracted data (neuronal noise)
J
Maximum correlation decoder best performance (0° az trials)
Cross val. prediction error (°)
Ch. lev.
MAA
(n = 8 mice)
Same # cells
Tuned
N.S. tuned
All cells
All cells decorr.
K
Maximum correlation decoder overall decoding performance
Cross val. prediction error (°)
Chance lev.
(n = 8 mice)
Neuronal noise
Decorrelated noise
L
Maximum correlation decoder best performance (0° az trials)
Cross val. prediction error (°)
Ch. lev.
MAA
(n = 8 mice)
Neuronal noise
Decorrelated noise

Figure 4: A) Schematic representation of supervised dimensionality reduction through linear discriminant analysis (LDA). B) Representative scatter of population response vectors in 2D linear discriminant space. C) Representative scatter of population response vectors in 13D linear discriminant space, remapped onto 2D UMAP. D) Schematic representation of the decoding strategy using multi-class classification algorithms on the recorded simultaneous population responses. E) Histograms depicting decoder accuracy evaluation across classes. For comparisons, we focused on the absolute, unsigned prediction error distributions. F) Histograms of the prediction error distributions from the decoders obtained by pooling the recorded errors from all tested mice (blue) and chance level distribution obtained by considering all possible prediction errors for the 13 azimuths tested (grey). G) Median absolute prediction errors obtained with the different decoding models across mice. H, I) Median absolute prediction errors throughout all azimuths tested obtained with the MLD (H) and MCD (I) models operating on simultaneous responses from tuned, N.S. tuned or all sampled cells across mice, including randomised samples to break noise correlations (Decorr.). J) Median absolute prediction errors obtained across mice for 0° azimuth trials with the MCD model considering tuned, N.S. tuned or all sampled cells. K, L) Median absolute prediction errors overall (K) and for 0° azimuth trials (L), obtained across mice with the MCD model operating on the neuronal noise from all sampled cells, including randomised samples to break noise correlations (Decorr.). *: comparisons against chance level, $p < 0.05$, Wilcoxon Signed Rank test with Sidak's correction for multiple comparisons. MLD: Maximum Likelihood decoder; MCD: Maximum Correlation Decoder; KNN: K-nearest neighbours; SVM ens.: Ensemble of Support Vector Machines; D. trees ens.: Decision trees ensemble; FF aNN: Fully connected feed-forward artificial neural net. Grey points depict median absolute prediction errors for each mouse; empty circles and whiskers depict their mean and std. deviation across mice, respectively. The plotted N.S. tuned neuron data from subsamples of equal size to tuned neuron samples is the averaged outcome of decoding 10 independent resampling iterations per mouse.

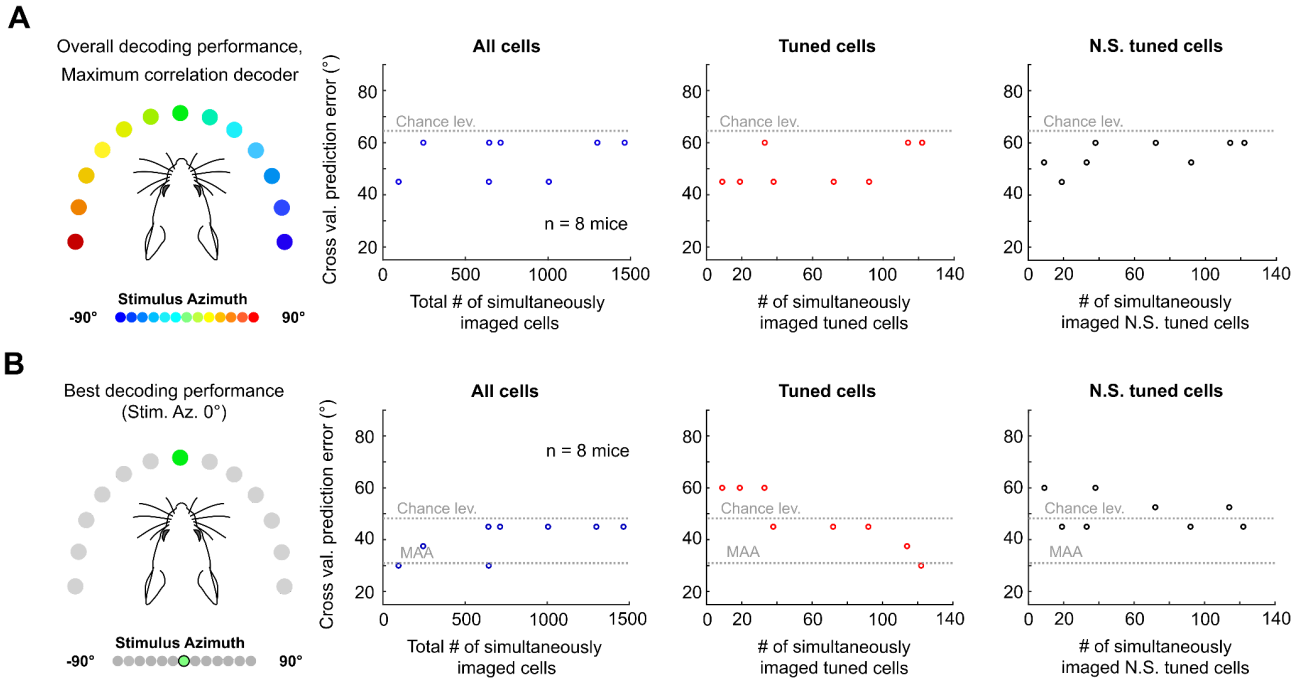


Supplemental figure 2. A) Median absolute prediction errors obtained after scrambling data. B) Histograms of the prediction error distributions from the MCD obtained by pooling the errors from same-azimuth trials recorded in all tested mice (blue) and chance level distribution obtained by considering all possible prediction errors for each azimuth tested (grey). C) Median prediction error trend across stimulus azimuth. D) Median absolute prediction errors obtained across mice for 0° azimuth trials with the MLD model for tuned, N.S. tuned or all sampled cells. Grey points depict median absolute prediction errors for each mouse; empty circles and whiskers depict their mean and std. deviation across mice, respectively. The plotted N.S. tuned neuron data from subsamples of equal size to tuned neuron samples is the averaged outcome of decoding 10 independent resampling iterations. The plotted randomised data with scrambled responses is the averaged outcome of decoding 100 independent randomisation iterations per mouse.

Decoding accuracy of sound location across azimuths and stimuli

The cross-validated prediction errors averaged across the tested 13 different azimuths were relatively large for both the MCD and MLD using the collected responses from passively listening mice (~55º, Fig. 4H,I). In contrast, mice trained to perform a binary discrimination task can discriminate 31º in azimuth with respect to 0º (minimum audible angle, MAA) (Lauer et al., 2011). Despite the fact that our experimental design involved passively listening mice with no discrimination training and multi-class representation of 13 different azimuths instead of binary discrimination, we explored the origins of this difference by studying if decoding performance varied as a function of stimulus azimuth to determine if there is a more precise azimuthal region. We decoded population responses to the same-azimuth trials by fitting models with complete datasets covering all azimuths, and compared the median cross-validated prediction errors across mice as a function of azimuth (Suppl. Fig. 2B, C). We observed that, under the experimental design used by us and others (Day and Delgutte, 2013, 2016; Panniello et al., 2018), the decoding showed different error distributions as a function of stimulus azimuth, as predicted by chance (Suppl. Fig. 2B), with a minimum median error for 0º azimuth trials (Suppl. Fig. 2C), which is also the reference azimuth used in the binary discrimination study (Lauer et al., 2011). Considering this observation, we compared the decoding performances of MLD and MCD models for 0º trials to determine the minimum decoding error achievable with these models, by fitting them to complete datasets covering all azimuths, respectively. The MLD produced median cross-validated prediction errors across mice that were not significantly different from the chance level, independently of the subset of neurons decoded (Suppl. Fig. 2D). In contrast, the MCD produced prediction errors that were significantly lower than chance level when decoding the complete population of sampled cells (Fig. 4J). This observation points out that the MCD outperforms the MLD when decoding population responses to 0º trials, reaching a median cross-validated prediction error across mice that does not significantly differ from the MAA (Fig. 4J, $p = 0.14$, Wilcoxon Signed Rank test with Sidak's correction for multiple comparisons). We observed that the MCD decoding error for 0º azimuth trials was not negatively correlated with the number of total cells decoded from the complete samples of imaged neurons (Pearson's R, all cells: 0.68, $p = 0.06$; n = 8 mice: Suppl. Fig. 3B), supporting sufficient sampling. In contrast, when we focused on the subset, such as the tuned or N.S. tuned cells, the decoding error showed a negative trend with respect to the number of cells decoded (Pearson's R, tuned: -0.92, $p = 0.001$; N.S. tuned: -0.35, $p = 0.40$; n = 8 mice: Suppl. Fig. 3B), indicating that smaller decoding errors would be possible if more neurons were sampled, especially for tuned cells. These observations point out that the decoding results for 0º azimuth trials obtained from the detected subsets of tuned neurons in our recordings must be taken

cautiously as the tuned neuron population is likely undersampled in our recordings. Keeping this in mind, we suggest that the responses from the complete population of sampled neurons is critical to achieve the best possible decoding performance (0º trials). Interestingly, the best decoding performance was only achieved by the MCD using the responses of the complete sampled population in the presence of noise correlations, as decorrelating the population responses produced median cross-validated prediction errors across mice that were not significantly different from chance level (Fig. 4J). This result supports that noise correlations contribute to the population code of sound azimuth and that they are key to achieve the best possible decoding performance (0º trials).



Supplemental figure 3: Median absolute prediction errors across all azimuths tested (A) or for 0º azimuth trials (B) as a function of the number of simultaneously sampled neurons decoded. Each point represents the median absolute prediction error recorded form one mouse using the MCD model operating on simultaneous responses from tuned, N.S. tuned or all sampled cells. The plotted N.S. tuned neuron data from subsamples of equal size to tuned neuron samples is the averaged outcome of decoding 10 independent resampling iterations per mouse.

Altogether, our decoding analysis points out that the population response patterns of both tuned and N.S. tuned neurons, including their correlated noise, are necessary for effective decoding of stimulus azimuth from single trial IC population responses. Next, to determine to what extent noise correlations carry azimuth information, we decoded the noise activity across azimuth trials. In this way, we minimised the influence of average response tuning for decoding. The overall median cross-validated prediction errors across mice obtained with the MCD operating on the noise activity alone during azimuth trials were significantly smaller than chance level, while decorrelating the noise produced chance level decoding errors (Fig. 4K). These results support that noise correlations indeed carry information about stimulus azimuth. However, decoding noise activity from 0º azimuth trials (best decoding performance conditions, Fig. 4J, Suppl. Fig. 2B) produced chance level errors (Fig 4L), arguing that the information carried by noise correlations is not enough but the

contribution of azimuth tuning is also necessary to achieve the smallest possible decoding errors (0º azimuth trials).

So far, we have focused on the population coding of azimuth using a sound stimulus rich in sound location information, carrying all the sound location cues (broad-band white noise, 20-40kHz, 96.2 dB). To test the MCD model with an independent stimulus, we repeated our experiments using 20kHz pure tones that carry only ILD cues for sound location, which is proposed to be the main sound location cue used by mice (Panniello et al., 2018). Mice are unlikely to rely on ITD due to the small size of their heads (Ono et al., 2020). Decoding population responses to pure tone stimuli using the MCD produced overall median cross-validated prediction errors across mice that were significantly lower than chance level, independently of the sound pressure level (SPL, Fig. 5A). Again, decorrelating the population responses to break noise correlations produced chance-level decoding performance (Fig. 5A). Altogether, these results support our observations using a broad-band stimulus and suggest that the population code of sound azimuth is robust across SPL, in consistency with previous reports (Day and Delgutte, 2016).

Trained mice are reported to have a poor azimuth discrimination ability for pure tone stimuli in a conditioned avoidance binary discrimination task (Lauer et al., 2011). Consistently, using the MCD to decode 0º azimuth trials from population responses to pure tone stimuli (best decoding performance conditions), we observed the chance level performance (Fig. 5B). Altogether, these results point out that sound azimuth can be decoded from IC population response patterns to pure tones (Fig. 5A), but broad-band stimuli are needed to reach the best decoding performance (Fig. 4J, 5B), as described in behavioural studies (Lauer et al., 2011).

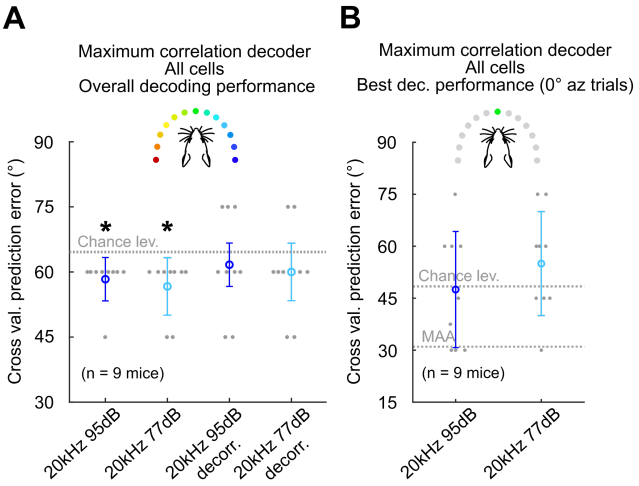


Figure 5: A) Median absolute prediction errors obtained with the MCD operating on the responses of all simultaneously recorded cells, across mice stimulated with pure tone stimuli of different intensities, including randomised samples to break noise correlations (Decorr.). B) Median absolute prediction errors obtained across mice stimulated with pure tone stimuli of different intensities for 0° azimuth trials with the MCD model for all sampled cells. *: comparisons against chance level, $p < 0.05$, Wilcoxon Signed Rank test with Sidak's correction for multiple comparisons. Grey points depict median absolute prediction errors for each mouse; empty circles and whiskers depict their mean and std. deviation across mice, respectively.

Azimuth coding across IC subdivisions

Our volumetric imaging approach allowed us to study the representation of sound azimuth across different anatomical subdivisions of IC. Previous *in vivo* imaging studies interrogating the functional architecture of the mouse IC suggested a subdivision based on sound frequency tuning consisting of a superficial 100 µm thick layer, referred to as the dorsal cortex of IC (DCIC), and a central nucleus (CNIC) (Barnstedt et al., 2015; Wong and Borst, 2019). So far, mammalian azimuth sensitive neurons were mainly sampled from the CNIC region, due to the difficulty of performing superficial recordings with traditional microelectrode techniques (Chase and Young, 2005, 2008; Day and Delgutte, 2013, 2016; Day et al., 2012), leaving the azimuth coding at DCIC unexplored. Our volumetric imaging approach enabled us to perform simultaneous recordings from DCIC populations within the superficial 100µm of IC (Fig. 6A, B), which were on average 92 neurons representing 11 ± 6 % (mean ± S.D., n = 8 mice, broad-band stimuli) of the complete imaged population.

In Fig. 1-4, we analysed the recordings from complete populations of sampled IC neurons across the imaged volumes without considering the anatomical subdivisions of IC. To determine if the anatomical subdivisions of IC sampled in our imaged volumes show any functional differences at the level of sound azimuth coding, we split the recorded populations into DCIC (within the superficial 100 µm of the imaged volume) and CNIC (the remaining deeper volume) neurons. Then, we compared the performance of the MCD when operating on the responses to broad-band stimuli from the DCIC subpopulation or subsamples of CNIC neurons of equal size to the DCIC subpopulation (see methods). The MCD produced overall median cross-validated prediction errors across mice that were significantly lower than chance level when operating on both DCIC and CNIC subpopulation responses (Fig. 6C), supporting the occurrence of sound azimuth coding at DCIC.

Interestingly, to reach the best decoding performance (decoding 0º azimuth trials) the complete population of sampled neurons is needed, including cells from DCIC and CNIC altogether (Fig. 6D). These results suggest that the azimuth information carried by DCIC and CNIC populations is likely complementary to generate an accurate representation of auditory space.

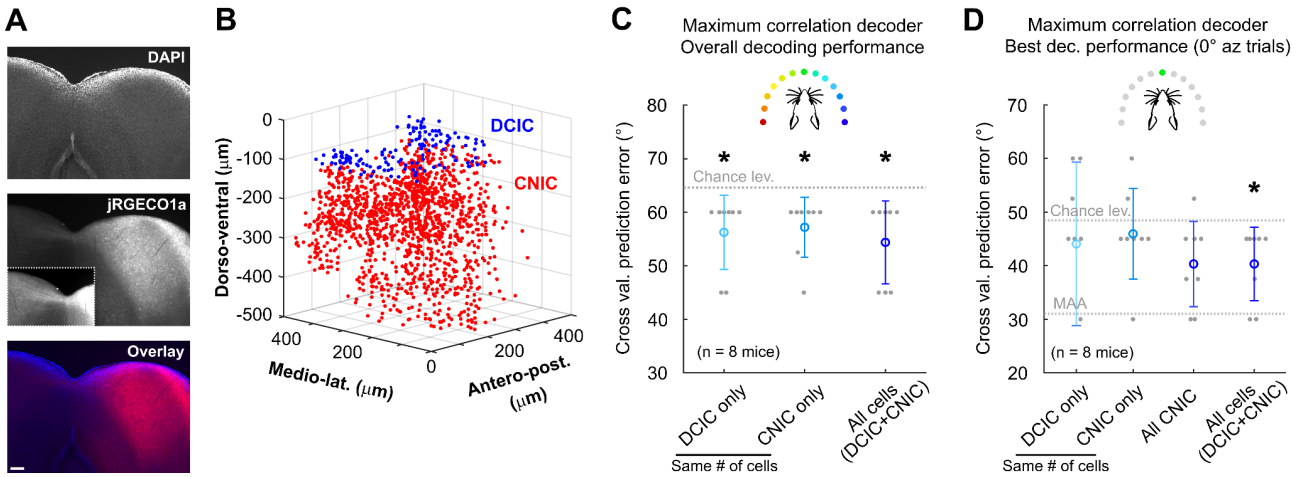


Figure 6: A) Representative coronal histological section of IC from an AAV injected and cranial window implanted mouse used in this study, showing synapsin driven jRGECO1a expression across the dorsal IC from the very surface (DCIC). Inset in middle panel displays a contrast enhanced section of the image, showing jRGECO1a labelled projections crossing over to the IC contralateral to the AAV injected one. Scalebar: 200 µm. B) Representative scatter plot of the position in the *in vivo* imaged volume of the centroids of the segmented cell somas. Cell positions corresponding to the dorsal cortex of IC (DCIC, superficial 100 µm) or central nucleus (CNIC), are coloured in blue and red, respectively. C) Comparison of the median absolute prediction errors obtained across mice with the MCD model operating on the simultaneously recorded responses from DCIC neurons, CNIC neurons or all sampled cells. D) Comparison of the median absolute prediction errors obtained across mice for 0° azimuth trials with the MCD model considering DCIC neurons, CNIC neurons or all sampled cells. *: comparisons against chance level, $p < 0.05$, Wilcoxon Signed Rank test with Sidak's correction for multiple comparisons. Grey points depict median absolute prediction errors for each mouse, and empty circles with whiskers depict their mean and std. deviation across mice, respectively. The plotted CNIC only data from subsamples of equal size to DCIC only neuron samples is the averaged outcome of decoding 10 independent resampling iterations.

Discussion

In this work, we tackled the technical challenge of recording simultaneously the response patterns from large numbers of IC neurons to sound azimuth by performing fast volumetric Ca^{2+} imaging through sTeFo 2P microscopy (Prevedel et al., 2016). We show that sTeFo 2P effectively achieves dense sampling across large imaging depths in highly light-scattering tissue such as the mouse IC, showcasing the capability of this method for interrogating neuronal population activity. Further improvements to this microscopy modality, including larger fields of view (Clough et al., 2021; Demas et al., 2021; Ota et al., 2021), multiplexing and 3P excitation (Weisenburger et al., 2017), will open new horizons in the study of neuronal population function and systems neuroscience.

The simultaneously recorded IC population responses showed prominent trial-to-trial variation that rendered response averages from single cells and their tuning little informative about single trial responses (Fig 2A-D). This supports that azimuth information is distributed in the coordinated response patterns of populations of IC neurons (Day and Delgutte, 2013). We show that, despite the trial-to-trial variability, the simultaneously acquired IC population responses to sound azimuth indeed effectively encode sound location on the single trial basis. Additionally, we characterise two key elements of these population activity patterns: azimuth sensitivity (tuning) and noise correlations. In particular, we found that noise correlations in IC population responses are linked to sound location through the recruitment of different noise correlated ensembles of neurons scattered across IC, as a function of azimuth. These findings point out that azimuth tuning is not the main or only feature of IC neurons carrying information about stimulus azimuth. This implication might extend to other brain regions relevant for sound location coding such as the auditory cortex. More generally, the clustered noise correlations observed across the recorded population responses outlook new perspectives on how noise correlations can occur, shape population activity and contribute to population coding of sensory information, opening exciting future research directions.

The decoding strategy proposed by Day and Delgutte (2013) was robust in effectively decoding stimulus azimuth from our imaging data, even though we used a different method and animal model to sample IC population responses. Our data therefore does not fully rule out maximum likelihood decoding

(MLD) as a decoding model for azimuth information in IC (Day and Delgutte, 2013). However, our data challenge some assumptions of the MLD model: for example, independent firing of IC neurons was not verified in our simultaneously recorded population responses. In order to go beyond these assumptions, we evaluated the more generic maximum correlation decoder (MCD) model which is based on matching the single-trial responses to the expected responses across different azimuths.

The decoding performance is not homogeneous across the frontal hemifield in the experimental design used by us and others (Day and Delgutte, 2013, 2016; Panniello et al., 2018), reaching the best possible decoding performance for 0º azimuth trials. This experimental design is aimed at studying the representation of sound azimuth across the frontal hemifield, rather than determining discrimination abilities. Nevertheless, the MCD model effectively decoded single-trial responses to 0º azimuth trials producing errors close to the discrimination ability of mice (MAA) and outperforming the MLD in this respect. Altogether, these observations support that the MCD is an informative model for generating hypotheses about the mechanisms of sound location processing between the IC and the ascending auditory pathway in mammals.

How many IC neurons are then needed to decode sound location? We show that at least both DCIC and CNIC neurons are needed, including both tuned and N.S. tuned ones (Fig. 4, 6). To give a clear answer to this question, however, we would need to find the most informative combination of neurons in the imaged sample, based on both azimuth tuning and noise correlations. Formulating an effective criterion to find this set of cells is challenging as it involves considering not only one simple feature of the population responses, like choosing the better tuned combination of cells, but also other complex features like noise correlated ensembles of neurons across azimuths, altogether. Future theoretical modelling studies aimed at finding an effective criterion to achieve this by understanding how these two features of IC population responses are integrated to encode sound location will be a key step to understand the minimum requirements for effective sound location coding at the mammalian IC.

It has been previously reported that the population responses of a single IC (unilateral recordings) can produce a complete representation of sound location throughout the frontal hemifield, and that this representation is robust across sound levels (Day and Delgutte, 2013, 2016). Our observations support both these notions with simultaneously acquired data, on the single trial basis (Fig. 4, 5). Discrepancies with respect to unilateral lesion studies (Day and Delgutte, 2013; Jenkins and Masterton, 1982), suggesting that one IC would carry information about the contralateral hemifield only, could be explained by the loss of commissural interactions between both ICs that contribute to sound location processing (Orton et al., 2016). In this respect, we observed that the jRGECO1a-labelled neurons we imaged could project contralaterally (Fig. 6A, inset), suggesting that the imaged IC populations would be involved in commissural interactions.

We show that sound location coding involves DCIC, following a similar mechanism to the more widely studied CNIC, and that DCIC and CNIC carry complementary information about sound location encoded in a similar format. One possibility is that DCIC information, which covers a broad range of frequencies (Barnstedt et al., 2015; Wong and Borst, 2019), compensates for information carried by the deeper part of CNIC that we could not access in our experimental setup but may cover azimuth information

present in higher sound frequencies. These findings open new and exciting research avenues, as DCIC is a higher order relay of the auditory pathway receiving descending corticofugal inputs (Bajo and King, 2012; Winer et al., 2002) where to study unexplored colliculo-cortical interactions for sound location information processing.

Despite the advantages of our imaging approach, our technique also has some limitations. In particular, large-scale volumetric imaging has a low temporal resolution, reaching sampling rates of 4 volumes per second, and the use of an indirect method of monitoring neuronal activity via Ca^{2+} sensors. The low temporal resolution effectively produces a low-pass filtering of neuronal activity, misrepresenting peak responses. Furthermore, indirectly inferring neuronal spiking responses from Ca^{2+} sensor signals can cause further information loss. Nevertheless, despite this information loss and misrepresentation of peak responses, two independent decoding models effectively classified the recorded population responses to the corresponding stimulus azimuth with errors smaller than chance level. This supports that our recordings still carried sufficient information about stimulus azimuth to study the neuronal population code of sound azimuth at the mouse IC. Moreover, the information loss associated to our approach suggests that the actual achievable decoding errors are likely smaller than the ones observed in this study, potentially reaching levels that could be more precise than the behavioural performance, as reported in the visual system (Stringer et al., 2021). Another aspect that could increase the prediction errors observed when decoding the recorded IC population responses includes neurons with activities associated to non-auditory processes (Gruters and Groh, 2012; Wong and Borst, 2019). Nevertheless, the detection and separation of these non-auditory cells are challenging due to the high trial-to-trial variability observed in our data, even for neurons with significant azimuth tuning. Another important limitation was that our custom microscope setup produced relatively high background SPL (79.5 dB SPL, see methods), despite our efforts to isolate the head-fixed mouse under the objective in a sound attenuating chamber as much as possible. Finally, despite the fact that our approach was focused on simultaneously sampling the responses from as large populations of IC neurons as technically possible, we did observe evidence of undersampling when studying the azimuth tuned neuronal subpopulation. Future studies based on characterising this or other small subpopulations of IC neurons will require further technical improvements to the 2-photon sTeFo imaging modality, such as multiplexing or larger fields of view (Clough et al., 2021; Demas et al., 2021; Ota et al., 2021; Weisenburger et al., 2019).

To conclude, this study points out a mechanistic role of IC in providing a general representation (encoding) of sound azimuth, based on azimuth tuning and noise correlations, to a multi-class template matching classification network for decoding sound azimuth. This prospective classification network potentially involves multiple ascending relays from the auditory pathway such as the medial geniculate nucleus or the auditory cortex. Future technical advancements enabling high throughput population recordings from different brain regions including IC and the auditory thalamus or cortex simultaneously will pave the way to further explore these possibilities.

Acknowledgements

We acknowledge Dr. Lina Streich and Ling Wang for their contribution to microscope assembly and setup, the mechanical and electronics workshops and the laboratory animal resource facility of EMBL Heidelberg for technical assistance. Dr. Evan Harrell for input on motorised speaker design and Dr. Jacques Bourg for input on IC cranial window surgeries. JCB acknowledges supporting fellowships from the EMBL Interdisciplinary Postdoc (EIPOD) Programme under Marie Skłodowska Curie Cofund Actions MSCA-COFUND-FP (664726). This work was funded by the European Molecular Biology Laboratory (EMBL) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project 458898724 awarded to RP.

Competing interests

The authors declare no competing financial interests.

Materials & correspondence

Datasets and code are available upon reasonable request.

References

Adams, J.C. (1979). Ascending projections to the inferior colliculus. J. Comp. Neurol. *183*, 519–538. https://doi.org/10.1002/cne.901830305.

Averbeck, B.B., Latham, P.E., and Pouget, A. (2006). Neural correlations, population coding and computation. Nat. Rev. Neurosci. *7*, 358–366. https://doi.org/10.1038/nrn1888.

Bajo, V.M., and King, A.J. (2012). Cortical modulation of auditory processing in the midbrain. Front. Neural Circuits *6*, 1–12. https://doi.org/10.3389/fncir.2012.00114.

Barnstedt, O., Keating, P., Weissenberger, Y., King, A.J., and Dahmen, J.C. (2015). Functional Microarchitecture of the Mouse Dorsal Inferior Colliculus Revealed through In Vivo Two-Photon Calcium Imaging. J. Neurosci. *35*, 10927–10939. https://doi.org/10.1523/JNEUROSCI.0103-15.2015.

Boffi, J.C., Knabbe, J., Kaiser, M., and Kuner, T. (2018). KCC2-dependent Steady-state Intracellular Chloride Concentration and pH in Cortical Layer 2/3 Neurons of Anesthetized and Awake Mice. Front. Cell. Neurosci. *12*, 1–14. https://doi.org/10.3389/fncel.2018.00007.

Brunso-Bechtold, J.K., Thompson, G.C., and Masterton, R.B. (1981). HRP study of the organization of auditory afferents ascending to central nucleus of inferior colliculus in cat. J. Comp. Neurol. *197*, 705–722. https://doi.org/10.1002/cne.901970410.

Chase, S.M., and Young, E.D. (2005). Limited segregation of different types of sound localization information among classes of units in the inferior colliculus. J. Neurosci. *25*, 7575–7585. https://doi.org/10.1523/JNEUROSCI.0915-05.2005.

Chase, S.M., and Young, E.D. (2008). Cues for sound localization are encoded in multiple aspects of spike trains in the inferior colliculus. J. Neurophysiol. *99*, 1672–1682. https://doi.org/10.1152/jn.00644.2007.

Clarey, J.C., Barone, P., Irons, W.A., Samson, F.K., and Imig, T.J. (1995). Comparison of Noise and Tone Azimuth Tuning of Neurons in Cat Primary Auditory-Cortex and Medial Geniculate-Body. J. Neurophysiol. *74*, 961–980. https://doi.org/10.1152/jn.1995.74.3.961.

Clough, M., Chen, I.A., Park, S.W., Ahrens, A.M., Stirman, J.N., Smith, S.L., and Chen, J.L. (2021). Flexible simultaneous mesoscale two-photon imaging of neural activity at high speeds. Nat. Commun. *12*, 1–7. https://doi.org/10.1038/s41467-021-26737-3.

Dana, H., and Shoham, S. (2012). Numerical evaluation of temporal focusing characteristics in transparent and scattering media: erratum. Opt. Express *20*, 28281. https://doi.org/10.1364/oe.20.028281.

Dana, H., Mohar, B., Sun, Y., Narayan, S., Gordus, A., Hasseman, J.P., Tsegaye, G., Holt, G.T., Hu, A., Walpita, D., et al. (2016). Sensitive red protein calcium indicators for imaging neural activity. Elife *5*, 1–24. https://doi.org/10.7554/eLife.12727.

Day, M.L., and Delgutte, B. (2013). Decoding Sound Source Location and Separation Using Neural Population Activity Patterns. J. Neurosci. *33*, 15837–15847. https://doi.org/10.1523/JNEUROSCI.2034-13.2013.

Day, M.L., and Delgutte, B. (2016). Neural population encoding and decoding of sound source location across sound level in the rabbit inferior colliculus. J. Neurophysiol. *115*, 193–207. https://doi.org/10.1152/jn.00643.2015.

Day, M.L., Koka, K., and Delgutte, B. (2012). Neural encoding of sound source location in the presence of a concurrent, spatially separated source. J. Neurophysiol. *108*, 2612–2628. https://doi.org/10.1152/jn.00303.2012.

Demas, J., Manley, J., Tejera, F., Barber, K., Kim, H., Traub, F.M., Chen, B., and Vaziri, A. (2021). High-speed, cortex-wide volumetric recording of neuroactivity at cellular resolution using light beads microscopy.

Deneux, T., Kaszas, A., Szalay, G., Katona, G., Lakner, T., Grinvald, A., Rozsa, B., and Vanzetta, I. (2016). Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations in vivo. Nat. Commun. *7*. https://doi.org/10.1038/ncomms12190.

Felix, R.A., Gourévitch, B., and Portfors, C. V. (2018). Subcortical pathways: Towards a better understanding of auditory disorders. Hear. Res. *362*, 48–60. https://doi.org/10.1016/j.heares.2018.01.008.

Franco, L., Rolls, E.T., Aggelopoulos, N.C., and Treves, A. (2004). The use of decoding to analyze the contribution to the information of the correlations between the firing of simultaneously recorded neurons. Exp. Brain Res. *155*, 370–384. https://doi.org/10.1007/s00221-003-1737-5.

Georgopoulos, A.P., Kalaska, J.F., Caminiti, R., and Massey, J.T. (1982). ON THE RELATIONS BETWEEN THE DIRECTION TWO-DIMENSIONAL ARM MOVEMENTS AND CELL DISCHARGE PRIMATE MOTOR CORTEX1 IN. J. Neurosci. *2*, 1527–1537.

Georgopoulos, A.P., Schwartz, A.B., and Kettner, R.E. (1986). Neuronal Population Coding of Movement Direction. Science. *233*, 1416–1419.

Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B.L., Koay, S.A., Taxidis, J., Najafi, F., Gauthier, J.L., Zhou, P., et al. (2019). CaImAn an open source tool for scalable calcium imaging data analysis. Elife *8*, 1–45. https://doi.org/10.7554/eLife.38173.

Grothe, B., Pecka, M., and McAlpine, D. (2010). Mechanisms of Sound Localization in Mammals. Physiol. Rev. *90*, 983–1012. https://doi.org/10.1152/physrev.00026.2009.

Gruters, K.G., and Groh, J.M. (2012). Sounds and beyond: Multisensory and other non-auditory signals in the inferior colliculus. Front. Neural Circuits *6*, 1–39. https://doi.org/10.3389/fncir.2012.00096.

Holtmaat, A., Bonhoeffer, T., Chow, D.K., Chuckowree, J., De Paola, V., Hofer, S.B., Hübener, M., Keck, T., Knott, G., Lee, W.C.A., et al. (2009). Long-term, high-resolution imaging in the mouse neocortex through a chronic cranial window. Nat. Protoc. *4*, 1128–1144. https://doi.org/10.1038/nprot.2009.89.

Huang, L., Ledochowitsch, P., Knoblich, U., Lecoq, J., Murphy, G.J., Reid, R.C., de Vries, S.E.J., Koch, C., Zeng, H., Buice, M.A., et al. (2021). Relationship between simultaneously recorded spiking activity and fluorescence signal in gcamp6 transgenic mice. Elife *10*, 1–19. https://doi.org/10.7554/eLife.51675.

Jazayeri, M., and Movshon, J.A. (2006). Optimal representation of sensory information by neural populations. Nat. Neurosci. *9*, 690–696. https://doi.org/10.1038/nn1691.

Jenkins, W.M., and Masterton, B. (1982). Sound Localization: Effects Of Unilateral Lesions in Central Auditory System. J. Neurophysiol. *47*, 978.

Jun, J.J., Steinmetz, N.A., Siegle, J.H., Denman, D.J., Bauza, M., Barbarits, B., Lee, A.K., Anastassiou, C.A., Andrei, A., Aydin, Ç., et al. (2017). Fully integrated silicon probes for high-density recording of neural activity. Nature *551*, 232–236. https://doi.org/10.1038/nature24636.

King, A.J., Schnupp, J.W.H., and Doubell, T.P. (2001). The shape of ears to come: Dynamic coding of auditory space. Trends Cogn. Sci. *5*, 261–270. https://doi.org/10.1016/S1364-6613(00)01660-0.

Knudsen, E.I., and Konishi, M. (1979). Mechanisms of sound localization in the barn owl (Tyto alba). J. Comp. Physiol. *133*, 13–21. https://doi.org/10.1007/BF00663106.

Kohn, A., Coen-Cagli, R., Kanitscheider, I., and Pouget, A. (2016). Correlations and Neuronal Population Information. Annu. Rev. Neurosci. *39*, 237–256. https://doi.org/10.1146/annurev-neuro-070815-013851.

Kulkarni, A., and Colburn, H.S. (1998). Role of spectral detail in sound-source localization. Nature *396*, 747–749. https://doi.org/10.1038/25526.

Lauer, A.M., Slee, S.J., and May, B.J. (2011). Acoustic basis of directional acuity in laboratory mice. JARO - J. Assoc. Res. Otolaryngol. *12*, 633–645. https://doi.org/10.1007/s10162-011-0279-y.

Lesica, N.A., Lingner, A., and Grothe, B. (2010). Population coding of interaural time differences in gerbils and barn owls. J. Neurosci. *30*, 11696–11702. https://doi.org/10.1523/JNEUROSCI.0846-10.2010.

McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. ArXiv *1802.03426*.

Ono, M., Bishop, D.C., and Oliver, D.L. (2020). Neuronal sensitivity to the interaural time difference of the sound envelope in the mouse inferior colliculus. Hear. Res. *385*, 107844. https://doi.org/10.1016/j.heares.2019.107844.

Orton, L.D., Papasavvas, C.A., and Rees, A. (2016). Commissural Gain Control Enhances the Midbrain Representation of Sound Location. J. Neurosci. *36*, 4470–4481. https://doi.org/10.1523/JNEUROSCI.3012-15.2016.

Ota, K., Oisi, Y., Suzuki, T., Ikeda, M., Ito, Y., Ito, T., Uwamori, H., Kobayashi, K., Kobayashi, M., Odagawa, M., et al. (2021). Fast, cell-resolution, contiguous-wide two-photon imaging to reveal functional network architectures across multi-modal cortical areas. Neuron *109*, 1810-1824.e9. https://doi.org/10.1016/j.neuron.2021.03.032.

Pachitariu, M., Stringer, C., and Harris, K.D. (2018). Robustness of spike deconvolution for neuronal calcium imaging. J. Neurosci. *38*, 7976–7985. https://doi.org/10.1523/JNEUROSCI.3339-17.2018.

Panniello, M., King, A.J., Dahmen, J.C., and Walker, K.M.M. (2018). Local and global spatial organization of interaural level difference and frequency preferences in auditory cortex. Cereb. Cortex *28*, 350–369. https://doi.org/10.1093/cercor/bhx295.

Pnevmatikakis, E.A., and Giovannucci, A. (2017). NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. J. Neurosci. Methods *291*, 83–94. https://doi.org/10.1016/j.jneumeth.2017.07.031.

Prevedel, R., Verhoef, A.J., Pernía-Andrade, A.J., Weisenburger, S., Huang, B.S., Nöbauer, T., Fernández, A., Delcour, J.E., Golshani, P., Baltuska, A., et al. (2016). Fast volumetric calcium imaging across multiple cortical layers using sculpted light. Nat. Methods *13*, 1021–1028. https://doi.org/10.1038/nmeth.4040.

Rupprecht, P., Carta, S., Hoffmann, A., Echizen, M., Blot, A., Kwan, A.C., Dan, Y., Hofer, S.B., Kitamura, K., Helmchen, F., et al. (2021). A database and deep learning toolbox for noise-optimized, generalized spike inference from calcium imaging. Nat. Neurosci. *24*, 1324–1337. https://doi.org/10.1038/s41593-021-00895-5.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: An open-source platform for biological-image analysis.Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

Schnupp, J.W., and King, a J. (1997). Coding for auditory space in the nucleus of the brachium of the inferior colliculus in the ferret. J. Neurophysiol. *78*, 2717–2731. https://doi.org/10.1152/jn.1997.78.5.2717.

Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. Science. *364*. https://doi.org/10.1126/science.aav7893.

Stringer, C., Michaelos, M., Tsyboulski, D., Lindo, S.E., and Pachitariu, M. (2021). High-precision coding in visual cortex. Cell *184*, 2767-2778.e15. https://doi.org/10.1016/j.cell.2021.03.042.

Vanwalleghem, G., Constantin, L., and Scott, E.K. (2021). Calcium Imaging and the Curse of Negativity. Front. Neural Circuits *14*, 1–10. https://doi.org/10.3389/fncir.2020.607391.

Weisenburger, S., Prevedel, R., and Vaziri, A. (2017). Quantitative evaluation of two-photon calcium imaging modalities for high-speed volumetric calcium imaging in scattering brain tissue. BioRxivru 115659. https://doi.org/10.1101/115659.

Weisenburger, S., Tejera, F., Demas, J., Chen, B., Manley, J., Sparks, F.T., Martínez Traub, F., Daigle, T., Zeng, H., Losonczy, A., et al. (2019). Volumetric Ca2+ Imaging in the Mouse Brain Using Hybrid Multiplexed Sculpted Light Microscopy. Cell *177*, 1050-1066.e14. https://doi.org/10.1016/j.cell.2019.03.011.

Winer, J.A., Chernock, M.L., Larue, D.T., and Cheung, S.W. (2002). Descending projections to the inferior colliculus from the posterior thalamus and the auditory cortex in rat, cat, and monkey. Hear. Res. *168*, 181–195. https://doi.org/10.1016/S0378-5955(02)00489-6.

Wong, A.B., and Borst, J.G.G. (2019). Tonotopic and non-auditory organization of the mouse dorsal inferior colliculus revealed by two-photon imaging. Elife *8*, 1–50. https://doi.org/10.7554/eLife.49091.

Methods

- Animals and ethics statement:

This work was performed in compliance to the European Communities Council Directive (2010/63/EU) to minimise animal pain and discomfort. EMBL's committee for animal welfare and institutional animal care and use (IACUC) approved all experimental procedures under protocol number 2019-04-15RP. Experiments were performed on 7-16-week-old CBA/j mice obtained from Charles River Laboratories and housed in groups of 1-5 in makrolon type 2L cages on ventilated racks at room temperature and 50% humidity with a 12 hr light cycle. Food and water were available ad libitum.

- Surgical procedures:

Cranial window surgeries were performed on 7-8-week-old mice of either sex following precedures published elsewere (Boffi et al., 2018). Briefly, anesthesia consisted of a mixture of 40 µl fentanyl (0.1 mg/ml; Janssen), 160 µl midazolam (5 mg/ml; Hameln) and 60 µl medetomidin (1 mg/ml; Pfizer), dosed in 5 µl/g body weight and injected i.p.. Hair removal cream was used to remove the fur over the scalp of anesthetized mice and eye ointment was applied (Bepanthen, Bayer). 1% xylocain (AstraZeneca) was injected as preincisional anesthesia under the scalp. Prepared mice were then placed in a stereotaxic apparatus (David Kopf Instruments, model 963) with their bodies on a heating pad (37°C). The scalp was surgically removed to expose the dorsal cranium. The periosteum was removed with fine forceps and scissors to prepare the

surface for cranial window implantation and stereotaxic delivery of viral vectors. Post-surgical pain relief was provided (Metacam, Boehringer Ingelheim) through s.c. injections (0.1 mg/ml, dosed 10µl/g body weight).

The Ca^{2+} indicator jRGECO1a (Dana et al., 2016) was expressed in IC neurons through induction with AAV vectors (Addgene #100854-AAV1), which were stereotaxically delivered as follows. A 4mm diameter circular craniectomy centred over IC (~1 mm posterior to lambda) was produced using a dental drill (Microtorque, Harvard Apparatus). Stereotaxic injections were performed at the centre of the right IC with pulled glass pipettes lowered to depths of 300, 400 and 500 µm, at a rate of ~4ul/hr using a syringe. ~300nl were injected per site. After injection, the craniectomy was sealed with a round 4mm coverslip (~170µm thick, disinfected with 70% ethanol), with a drop of saline between the glass and the dura, and dental acrylic cement (Hager Werken Cyano Fast and Paladur acrylic powder). A head fixation bar was also cemented. The open skin wound was also closed with acrylic cement. At the end of the surgery anesthesia was antagonised with a subcutaneous injection of a mixture of 120 µl sterile saline, 800 µl flumazenil (0.1 mg/ml; Fresenius Kabi), and 60 µl antipamezole (5 mg/ml; Pfizer) dosed in 10 µl/g body weight. Mice were single housed after surgery and were allowed to recover for at least 4 weeks before imaging, providing time for Ca^{2+} indicator expression and for the inflammation associated with this surgery to resolve (Holtmaat et al., 2009).

- sTeFo 2P microscopy:

We built a bespoke scanned temporal focusing (sTeFo) 2 photon microscope for fast volumetric in vivo Ca^{2+} imaging of large IC neuronal populations from awake mice. Technical details and detailed working principle are extensively described elsewhere (Prevedel et al., 2016). The main difference with respect to the microscope design published by Prevedel et al. (2016) was a higher repetition laser (10 MHz, FemtoTrain, Spectra Physics). Laser power while scanning was kept below 191 mW, measured after the objective, to avoid heating the brain tissue excessively (Prevedel et al., 2016). 470 μm^2 fields of view were imaged at 128 px^2 resolution and spaced in 15 µm z steps to cover 570 µm in depth (38 z steps) at a volume rate of 4.01 Hz. Artefacts produced by objective piezo z drive flyback were excluded from analysis.

In a typical imaging session, cranial window implanted mice were briefly (< 1 min) anesthetized with 5% isoflurane in O_2 for quick head fixation at a custom stage, slightly restraining their bodies inside a 5 cm diameter acrylic tube, and positioned under our custom 2P microscope. Mice fully recovered from the brief isoflurane anesthesia, showing a clear blinking reflex, whisking and sniffing behaviours and normal body posture and movements, immediately after head fixation. A pilot group of mice were habituated to the head fixed condition in daily 20 min sessions for 3 days, however we did not observe a marked contrast in the behaviour of habituated versus unhabituated mice during our relatively short 25 min imaging sessions. In consequence imaging sessions never surpassed a maximum of 25 min, after which the mouse was returned to its home cage. Typically, mice were imaged a total of 2-11 times (sessions), one to three times a week.

- Sound Stimulation:

A custom sound-attenuating chamber was incorporated to the microscope setup, allowing only the objective to access the inside, to attenuate room and microscope scanner noise reaching the mouse. All possible sound reflecting surfaces were lined with 1 cm or 0.5 cm thick foam to preserve free-field conditions. SPL was measured using a free-field prepolarized measurement microphone (PCB Piezotronics, 378C01, with signal conditioner 482A21) placed on the setup stage in the position of the head of the mouse. Background R.M.S. SPL measurements during imaging without sound stimuli delivery were (in dB): 20 kHz: 68.1; 20-40 kHz: 79.5. For sound stimuli to be salient, their R.M.S. SPL was at least 9 dB above background R.M.S. SPL.

Sound stimuli consisted of 20-40 kHz broadband white noise at 96.2 dB R.M.S. SPL and a 20 kHz pure tone at 77.2 or 94.9 dB R.M.S. SPL, 500ms in total duration including 10 ms up and down ramps. Sound stimuli were delivered through an electrostatic speaker (Tucker Davis Technologies, ES1 coupled to a ED1 speaker driver) mounted on an arm at an 8.5 cm distance away from the head of the mouse, which was mounted on a NEMA17 bipolar stepper motor (26 Ncm, 1.8 deg/step) placed under the head of the mouse to position the speaker around the frontal hemifield. The motor was controlled using an Arduino UNO microcontroller, running GRBL v0.9 (G-code interpreter, https://github.com/grbl/grbl) and coupled with the Arduino CNC Shield V3.10 (https://blog.protoneer.co.nz/arduino-cnc-controller/) running A4988 motor drivers. To control and program speaker positioning, we used the universal G-code sender (https://winder.github.io/ugs_website/). Synchronisation of speaker positioning and sound stimulus delivery to data acquisition was done through a second Arduino UNO board receiving a scan start TTL from the microscope to trigger stimulation programs.

500 ms audio stimuli (including 10 ms up and down ramps) were synthesised using MATLAB (0.5 MHz sampling rate). Sound delivery was performed using a DAQ board (National Instruments, PXIe-6341) interfaced to MATLAB using the DAQmx wrapper from scanimage (https://vidriotechnologies.com/scanimage/).

A typical stimulation protocol consisted of a series of presentations of a sound stimulus in which azimuth angles varied randomly from 13 different frontal positions (frontal 180° split into 15° steps). Stimuli were presented every 5 seconds and each azimuth angle was presented on average 14 times per session. Minimum number of same-azimuth trials collected was 8.

- Ca^{2+}-imaging signal analysis:

Volumetric imaging data was visualised and rendered using FIJI (Schindelin et al., 2012). Motion correction was performed using NoRMCorre (Pnevmatikakis and Giovannucci, 2017) on individual imaging planes from the volumetric datasets. jRGECO1a signal was extracted, denoised and neuropil corrected from each

individual imaging plane using CaImAn (Giovannucci et al., 2019). Spike probabilities were estimated using MLSpike (Deneux et al., 2016).

- Statistical analyses:

Statistical analyses were performed using MATLAB functions and custom scripts. For hypothesis testing, a p value < 0.05 was considered significant. nS/N ratios were evaluated as the inverse of the coefficient of variation of the recordings. Linear discriminant analysis was performed using MATLAB LDA function developed by Will Dwinnell (2022): (https://www.mathworks.com/matlabcentral/fileexchange/29673-lda-linear-discriminant-analysis). IC population response classification was performed using published linear decoders (MLD, MCD) (Day and Delgutte, 2013; Franco et al., 2004; Jazayeri and Movshon, 2006) implemented in MATLAB, or using the following functions from the statistics and machine learning toolbox from MATLAB: fitcknn, fitcnet, fitcecoc and fitcensemble using decision tree learners and adaptive boosting as an aggregation method. For the latter functions, bayesian hyperparameter optimization was employed to fit the models. For all classifiers, a “leave one out” cross-validation strategy was implemented, in which the models were fitted to the population responses collected in all trials except one (for each imaging session) and the one population response left out was decoded using the fitted model. This procedure was iterated until all population responses collected were decoded (for each imaging session). Decoding errors from N.S. tuned neuron subsamples of equal size to tuned neuron samples and CNIC subsamples of equal size to DCIC samples are the average measures calculated from 10 independent resampling iterations per mouse. Decoding errors from randomised, scrambled responses are the average measures calculated from 100 independent resampling iterations per mouse. For noise correlation analyses, negative values resulting mostly from subtracting the mean response to baseline single trial responses (or lack of responses) of the corresponding azimuth, were not considered.