Neurometric correlates to sensitive high-frequency sound amplitude modulation detection by mice.

Maurits M. van den Berg, Esmée Busscher, J. Gerard G. Borst, and Aaron B. Wong

Department of Neuroscience, Erasmus MC, University Medical Center Rotterdam, NL-3015 GD Rotterdam, Netherlands

Contents

Neurometric correlates to sensitive high-frequency sound amplitude modulation detection by mice. 1
Abstract 2
Introduction 3
Results 5
Discussion 27
Methods 32
Animals 32
Surgery 32
Amplitude-modulated sound stimulus 33
Behavioral training 33
Auditory Brainstem Response (ABR) 34
Electrophysiological recordings 34
Pure-tone response analysis 35
AM analysis 35
Principal component analysis of response differences 35
Acknowledgements 36
References 37

Abstract

Amplitude modulation (AM) is a common feature of natural sounds, including speech and animal vocalizations. Here, we used operant conditioning and *in vivo* electrophysiology to determine the AM detection threshold of mice as well as its underlying neuronal encoding. Mice were trained in a Go-NoGo task to detect the transition to AM within a noise stimulus designed to prevent the use of spectral side-bands or a change in intensity as alternative cues. Our results indicate that mice, in comparison with other species, detect high modulation frequencies up to 512 Hz exceptionally well, but show poor performance at low frequencies. Our *in vivo* multielectrode recordings in the inferior colliculus (IC) of both anesthetized and awake mice revealed a few single units with remarkable phase-locking ability to 512 Hz modulation, but not sufficient to explain the good behavioral detection. Using a model of the population response that combined dimensionality reduction with threshold detection, we reproduced the general high-pass characteristics of behavioral detection based on a subset of neurons showing the largest firing rate change (both increase and decrease) in response to AM. Our data thus identify candidate neurons in the IC to explain the high-pass transfer function for AM detection in the mouse.

Introduction

A major goal of neuroscience is to understand how sensory stimuli are encoded in the firing patterns of sensory neurons and how these patterns guide behavior. This is typically studied in animals by comparing neuronal responses with behavioral detection of a suitable stimulus. This comparison of neurometric and psychometric performance has yielded much information about the way relevant sensory features are encoded at different stages of the sensory system, but also about the contribution of other parts of the sensorimotor loop, which include non-sensory factors such as attention, motivation, bias, decision making and the motor response.

In these studies, it is a common finding that psychometric performance is poor compared to neurometric performance (Stüttgen et al. 2011). In sensory detection tasks, individual neurons may already outperform the animal in detecting the stimulus or a change in the stimulus (Kettner and Thompson 1985; Newsome et al. 1989; Malone et al. 2010; Johnson et al. 2012). At the same time, stimulation of even a single neuron may already influence sensory detection (Tanke et al. 2018). Then what might account for this apparent discrepancy? A possibility is that the activity of different sensory neurons are correlated, as these so-called noise correlations limit the information these neurons can carry (Azeredo da Silveira and Rieke 2021). However, noise correlations are often small (Renart et al. 2010; Bartolo et al. 2020), and a recent, detailed study on the precision of visual discrimination afforded by visual cortex neurons, showed that the threshold based upon the population was two orders of magnitude more sensitive than the performance of the mouse, indicating that there are little or no limitations due to noise correlations within the population responses in the sensory precision that can be attained (Stringer et al. 2021). This could mean that under these conditions the gap between neurometric and psychometric factors is filled by non-sensory factors (Kettner and Thompson 1985; de Lafuente and Romo 2006; Guo et al. 2014; Goris et al. 2017; Stringer et al. 2021), but we argue here that sensory performance in some behavioral tasks may also be easily overestimated.

Here, we compare the responses of neurons in the mouse inferior colliculus (IC) to amplitude modulations (AM) in sounds with their behavioral ability to detect these modulations. Amplitude modulations form an important cue in communication sounds such as speech, but also in the identification of environmental sounds, segregating sound sources during auditory scene analysis, pitch perception or musical perception (reviewed in (Joris et al. 2004; Rees and Langner 2005). One way or another, the firing of each auditory neuron is modulated by sound amplitude. AM tuning is typically characterized by presenting sinusoidally amplitude modulation (SAM) of a carrier sound at different frequencies to measure how well neurons phase lock to the modulation frequency, yielding the temporal modulation transfer function (tMTF), or how vigorously they fire in response to different modulation frequencies, which is called the rate modulation transfer function (rMTF). This characterization also yields the modulation frequency at which phase locking is best (the temporal best modulation frequency, tBMF) or the modulation frequency at which the firing frequency is highest (rBMF). The representation of AM stimuli has been studied in detail at various stages within the auditory system. Between the cochlear nucleus and the IC, the tMTF most noticeably changes. The tuning becomes sharper, but in general the response bandwidth is reduced as well (Joris et al. 2004). At the same time the ability of IC neurons to phase-lock to high modulation frequencies is reduced compared to the cochlear nucleus. Presumably, the convergence of excitatory and inhibitory inputs contributes to tuning changes observed at the level of the IC (Casseday et al. 1994; Leary et al. 2008; Geis and Borst 2009; Ono et al. 2017). For the rMTF some of the mechanisms have been elucidated in whole-cell

recordings; it was found that both the timing of inhibitory and excitatory synaptic inputs and intrinsic cellular properties related to spike generation could contribute to the rate tuning of IC neurons to SAM stimuli (Geis and Borst 2009). The decrease in the upper limit of modulation frequency to which neurons phase-lock further continues towards the auditory cortex (Joris et al. 2004), which has led to the suggestion that AM sounds are encoded by a rate code rather than a temporal code for behavioral detection (Lu and Wang 2004; Bendor and Wang 2007; Dong et al. 2011; Niwa et al. 2012b). Other studies have shown that temporal codes are often found to more accurately represent sounds for both pure AM and natural vocalizations with large AM components (Schnupp et al. 2006; Wang et al. 2007; Walker et al. 2008). Even though there are many studies that have studied either the behavioral detection of AM or the electrophysiological encoding of AM (reviewed in (Joris et al. 2004; Rees and Langner 2005), the question which features in the AM tuning have behavioral relevance is therefore still far from settled.

How individual neurons contribute to AM recognition in sounds is much less well understood. We therefore compare SAM tuning of single units from multielectrode recordings both in anesthetized and awake mice with the performance in a behavioral SAM detection task. When comparing behavioral and neural performance, we find that despite the large diversity and the complex tuning of many IC neurons to AM noise, only a small fraction of the units is sufficient to explain most of the variance that is relevant for the task. By also taking into account the relation between modulation depth and response latency, we could provide an adequate description for the behavioral performance based on the neural responses. Moreover, a comparison of neuronal responses in anesthetized and awake animals indicated that the latter approximates the behavioral responses better. Our data thus provide a partial explanation for the gap that exists between neurometric and psychometric performance.

Results

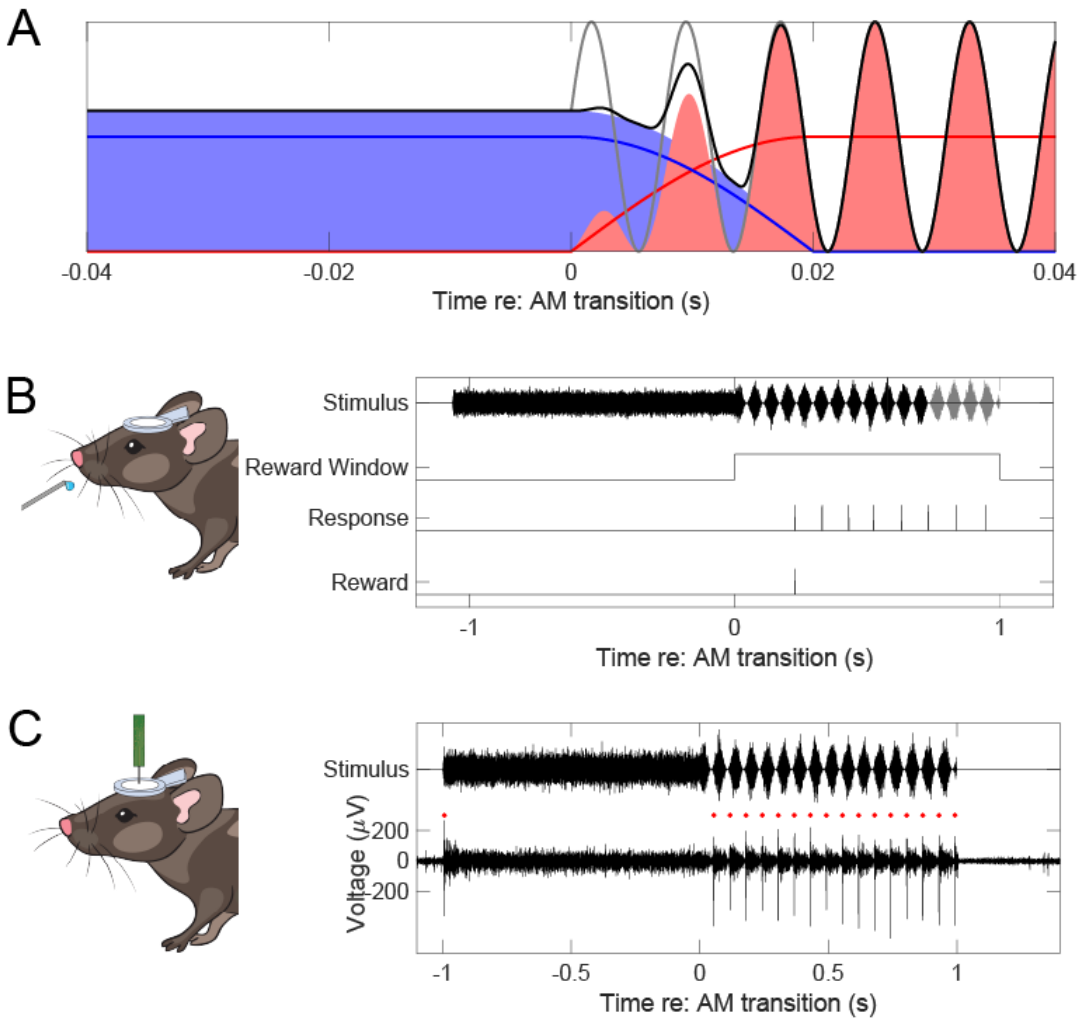


Figure 1 Experimental design. A) Trials consisted of unmodulated (UM) noise stimuli that transitioned to amplitude-modulated (AM) noise stimuli with matched intensity and spectral properties. (blue area: envelope of unmodulated noise with cross-fading, red area: envelope of the modulated noise with cross-fading, blue & red lines: multiplicative factors for cross-fading, grey line: modulated envelope without cross-faded transition, black line: combined envelope with cross-faded transition) B) Behavioral test trial structure: During the UM period of random 1-2 s duration, the mouse had to withhold licking; during the reward window, which started at the transition to AM, licking was rewarded with a water drop. A premature response would end the trial (black stimulus trace) 500 ms after the rewarded response, compared to the planned stimulus duration (grey stimulus). C) Electrophysiology trial structure, the UM and AM periods were both fixed at 1 s. An example trace from an individual channel is shown. Red markers: spikes from a single unit.

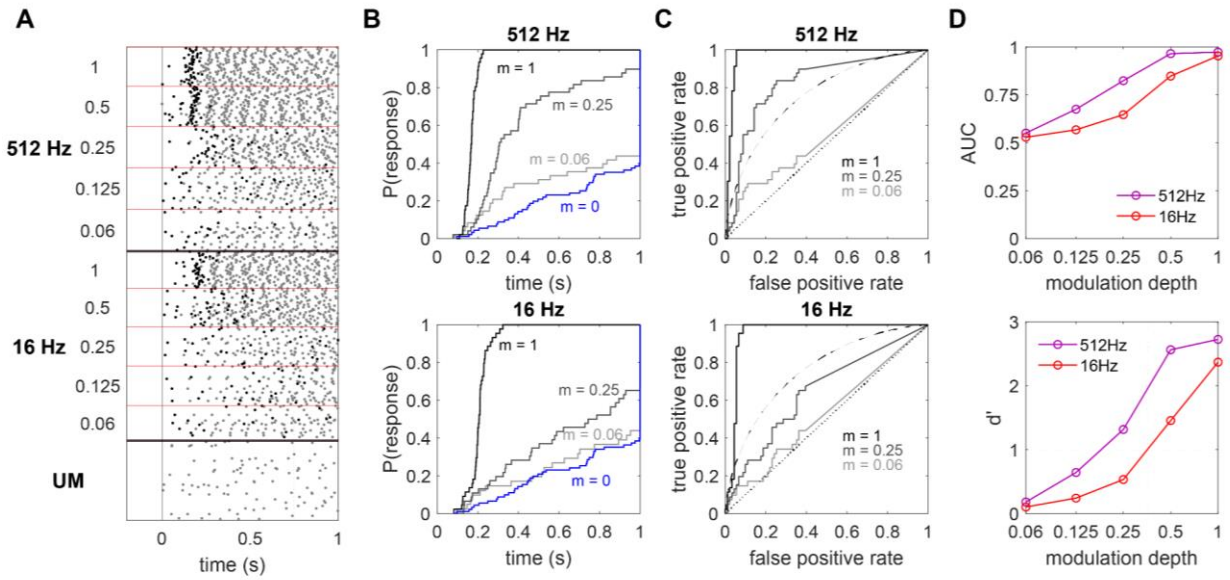


Figure 2 Behavioral detection of a low and a high frequency AM stimulus. (A) Dot raster of lick responses to AM transitions for an example mouse at a low and a high modulation frequency; Responses at 512 and 16 Hz modulation frequency were sorted from top to bottom for decreasing modulation depths. Black dots represent the rewarded licks; grey dots the other licks. The UM modulation frequency is the unrewarded catch trial condition. (B) Cumulative distribution of response latencies after the AM transition for the two modulation frequencies at various modulation depths m. Response latencies shorter than 75 ms were excluded. (C) Receiver-Operator-Characteristics (ROC) analysis for response latency, plotting cumulative probability of a response to AM stimuli versus catch trials as the criteria (response window) changes. Dashed curve is the theoretical ROC curve for Sensitivity index (d') = 1. (D) Top: Area-under-the-curve (AUC) of the ROC curves in C plotted as a function of modulation depth. Bottom: Sensitivity index (d') converted from AUC values according to $d' = \sqrt{2} * \Phi^{-1}(AUC)$*, where* Φ^{-1} *is the inverse cumulative distribution function of the standard normal distribution. All responses are from the same animal.*

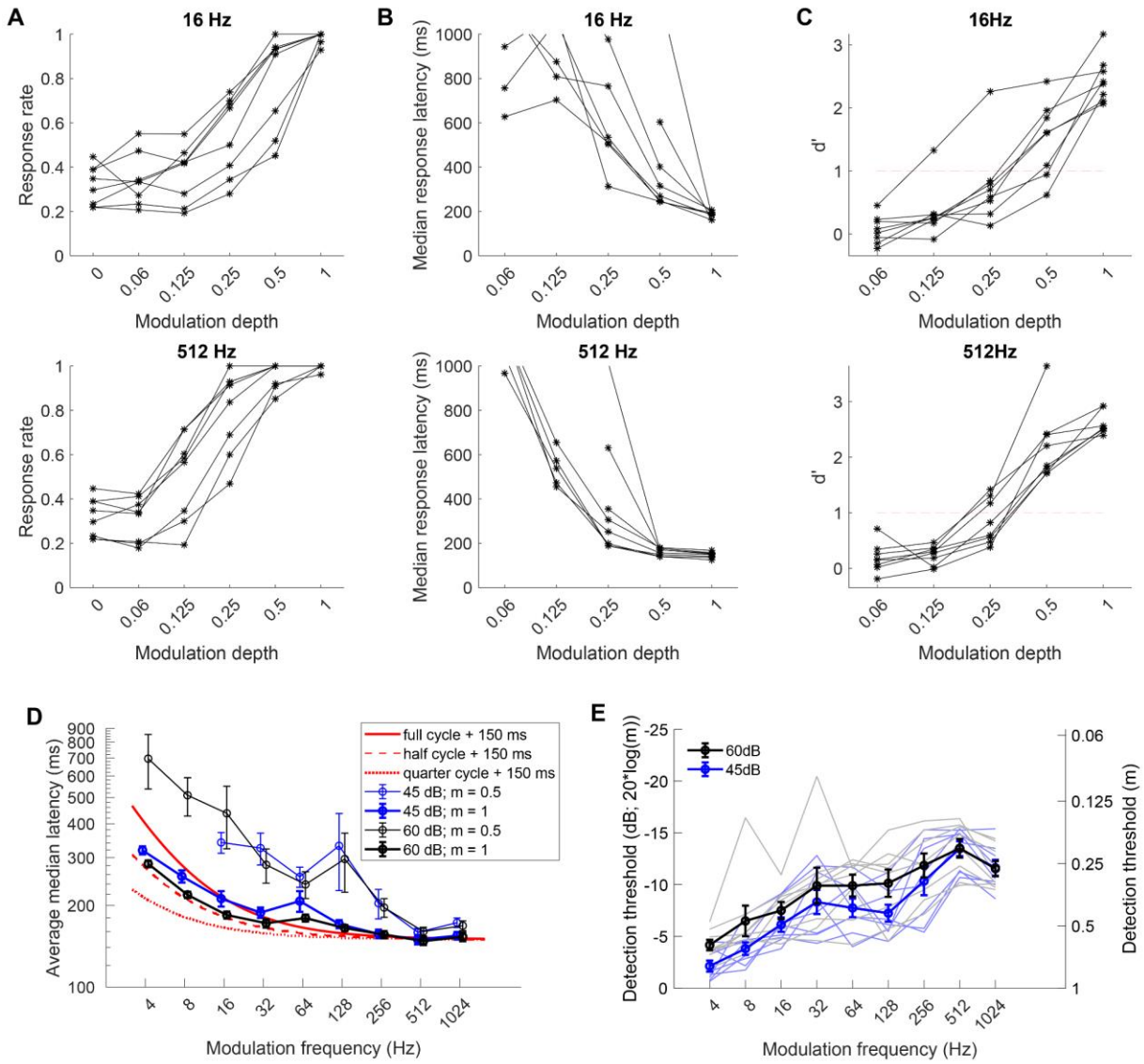


Figure 3 Behavioral detection thresholds at different AM frequencies show a high pass pattern. (A) Probability of a behavioral response within 750 ms after transitioning to 16 (upper) or 512 Hz (lower) AM at different modulation depths. (B) Median response latencies of animals at 16 and 512 Hz, showing faster responses at higher modulation depths. (C) Response latencies at 16 (upper) and 512 (lower) Hz were converted to d' using ROC analysis to account for response bias. Red, horizontal dashed lines indicate d'=1, our definition of the detection threshold. (D) Median response latencies (mean ± s.e.m.) to AM sounds as a function of modulation frequency for 60 dB (black) and 45 dB (blue) noise. At full modulation (m = 1; thick lines), the response latencies showed high consistency across animals, and could be well approximated by a fixed delay (150 ms) plus half a modulation cycle (red dashed line). (E) Frequency- and intensity-dependence of the AM detection threshold of mice plotted as mean ± s.e.m. (thick lines). Thin lines are responses of the individual mice. Significant frequency and intensity dependence was observed. (Linear mixed effect model; frequency: $p < 0.0001$, intensity: $p < 0.0001$).

Eight mice were trained in a Go-NoGo task to respond to the transition from an unmodulated to an amplitude modulated (AM) broadband noise stimulus (Figure 1B). The stimulus was designed to specifically test for amplitude-modulation, while controlling for spectral and intensity differences induced by the amplitude modulation (see Figure 1A and Methods). Individual detection thresholds

were obtained by testing a total of nine different modulation frequencies ranging from 4 to 1024 Hz at a total of five different modulation depths and two sound intensities. We used response latency as the main readout of how readily each animal detected each stimulus. Example responses to a low (16 Hz) and a high (512 Hz) modulation frequency at 60 dB are shown in Figure 2. The mice learned the task well, except, as is often the case with a Go-NoGo paradigm, false positive rates during catch trials were relatively high. However, they did not differ for 45 dB (32.8 ± 13.1%; range 12.2-46.1%) and 60 dB stimuli (31.7 ± 8.9%; range 21.8-44.7%; paired t-test, $p = 0.68$).

To correct for response bias, response latencies during target and unmodulated catch trials were compared using ROC analysis (Figure 2B,C). Areas under the curve (AUC) were calculated (Figure 2D) and subsequently converted to the sensitivity index *d'* (Figure 2D). Higher d' values were observed at a modulation frequency of 512 Hz than at 16 Hz, indicating that this animal was better at detecting the former. For each mouse and each modulation frequency, the detection threshold was defined as the modulation depth at which d' crossed 1.

Figure 3 summarizes the results of all 8 animals. As expected, response probability for all animals increased with increasing modulation depth for a wide range of modulation frequencies (Figure 3A). Moreover, at all modulation frequencies animals responded faster to higher modulation depths (Figure 3B), probably reflecting the certainty of the animal's detection and subsequent response. The average median response latency to the fully modulated stimuli could be approximated by adding half a period of the modulation to a fixed delay of about 150 ms, suggesting that under these conditions detection approached the physical limit at which sinusoidal amplitude modulations can be detected.

Similar to the example shown in Figure 2, we converted the response latencies of all animals into the sensitivity index d' to estimate their detection threshold (Figure 3C). Behavioral detection of AM showed a clear high-pass pattern with detection thresholds in the lower range (8 Hz and lower) at a modulation depth of about 0.5, whereas in the higher range (256 Hz and higher) it was about 0.25 (Figure 3E). The detection threshold worsened at 1024 Hz, possibly due to phase-locking limitations at the level of the auditory nerve (Taberner and Liberman 2005).

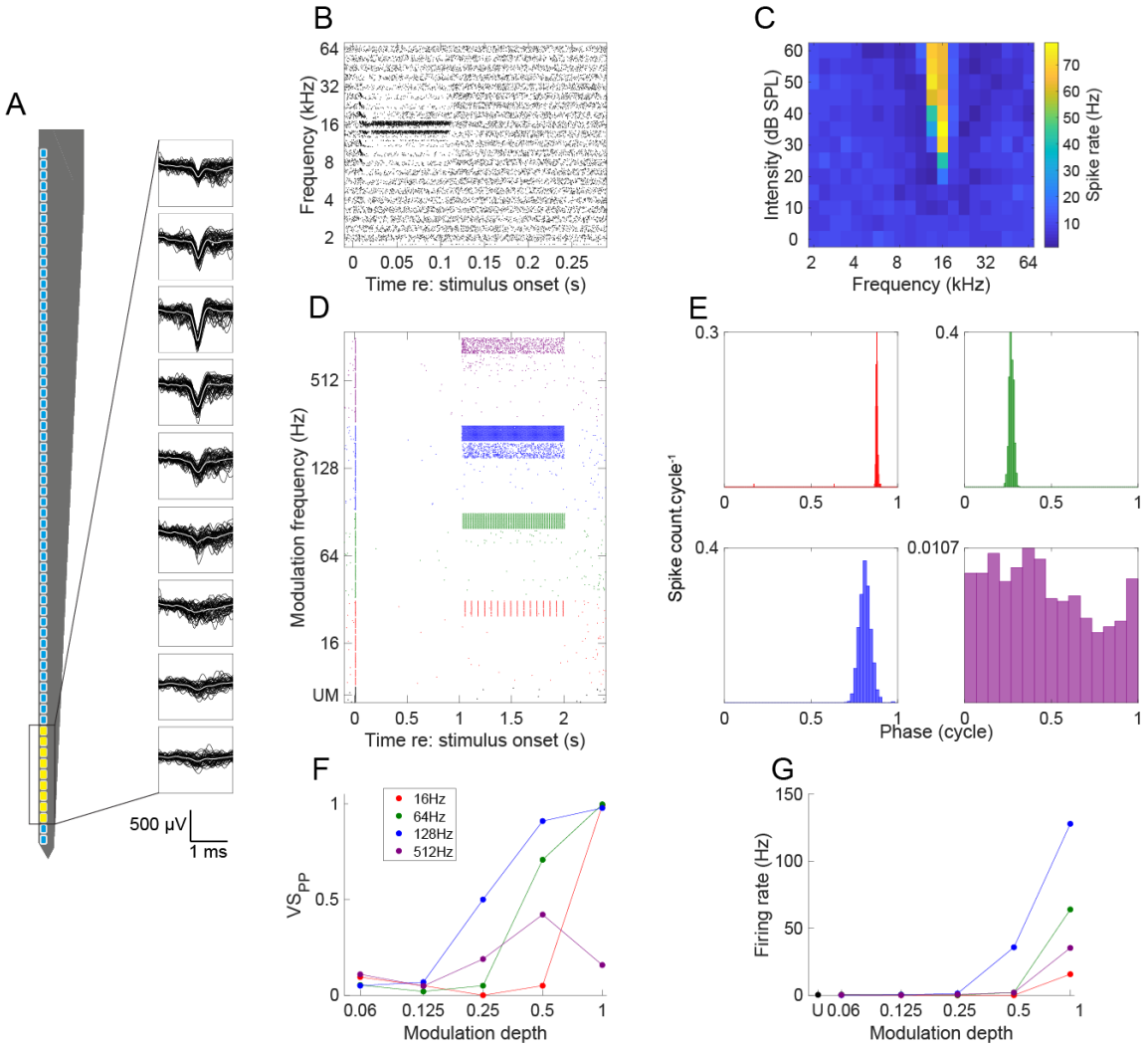


Figure 4 Example unit from the central nucleus of the IC with an onset response and relatively good phase locking. A) Spike waveforms detected on adjacent channels. B) Dot raster plot of the pure tone responses with the ordinate sorted by frequency followed by intensity. C) Responses to 100 ms pure tones displayed as a heatmap show sharp tuning with a CF of 16 kHz and a minimum threshold of 20 dB. D) Dot raster plot of responses to AM stimuli. Trials are organized by decreasing modulation depth for the different modulation frequencies. Transition from UM to AM occurred at 1 s. E) Cycle histograms of peak phase-locking modulation depth for the same modulation frequencies as shown in D. Bin size is 0.13 ms. F) Temporal coding of AM was quantified by calculating the phase-projected vector strength (VS_{PP}) for the responses during the last 500 ms of the AM period. G) The average firing rate during the last 500 ms of the AM period is used as a measure for rate coding to AM. All AM data shown are for a carrier noise intensity of 60 dB SPL.

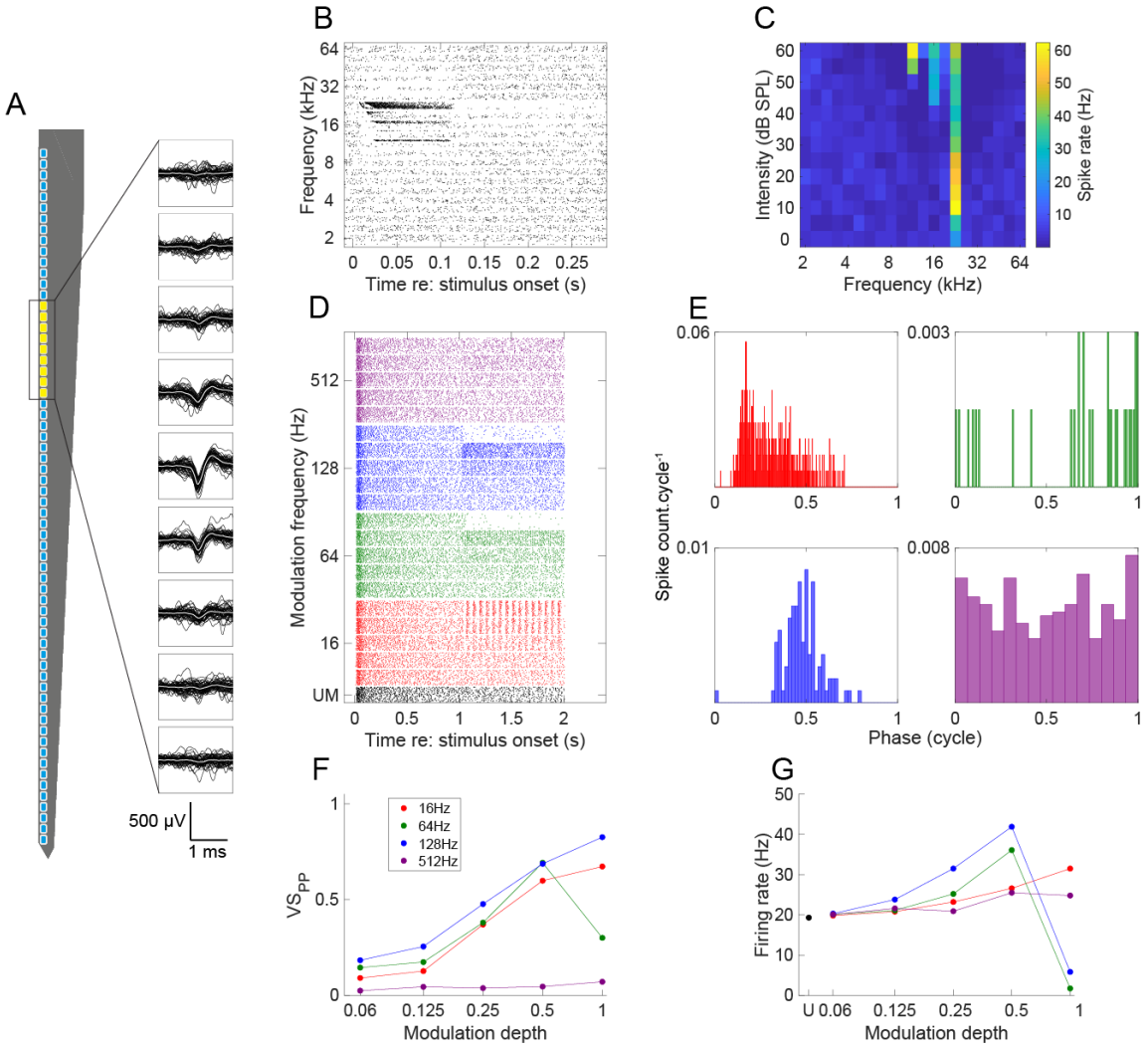


Figure 5 Example unit from the central nucleus of the IC with sustained responses. A) Spike waveforms detected on adjacent channels. B) Dot raster plot of the pure tone responses with the ordinate sorted by frequency followed by intensity. C) FRA of pure tone responses showing sharp tuning with a CF of 22.6 kHz and a minimum threshold of 0 dB. D) Dot raster plot of responses to AM stimuli. Trials organized by decreasing modulation depth for the different modulation frequencies. Transition from UM to AM occurred at 1 s. E) Cycle histograms show phase-locking ability for full modulation stimuli. Color scheme of modulation frequency matches that of D. Bin size ≈ 0.13 ms. F) Temporal coding of AM was quantified by calculating VS_{PP} for the last 500 ms of the AM period. G) The average firing rate during the last 500 ms of the AM period served as a measure for rate coding to AM, and shows a non-monotonic relationship with modulation depth for 64 and 128 Hz modulation frequencies All AM data shown are for a carrier noise intensity of 60 dB SPL.

The unexpectedly high sensitivity to high-frequency modulations prompted us to investigate the underlying encoding mechanism using multichannel electrode recordings in the IC. The IC is a key nucleus in the auditory pathway where tuning for AM responses may emerge (Joris et al. 2004; Rees and Langner 2005). Our aim was to relate the neuronal responses of IC neurons to the perceptual AM responses.

We recorded a total of 112 well isolated single units from the IC of 18 anaesthetized mice using multielectrode silicon probes. We presented the same AM stimuli as in the behavioral task (Figure 1A,C), as well as pure tones to obtain their frequency response areas (FRAs). We calculated both the phase-projected vector strength (VS_{PP}, (Yin et al. 2011) as a measure for phase-locking to the AM stimuli to obtain the temporal modulation transfer function (tMTF) and the average firing rate to obtain the rate modulation transfer function (rMTF).

The units showed large variability in their responses to both AM stimuli and tones. In Figure 4 and 5 we illustrate the responses of two example units from the central nucleus of the IC. Figure 4 shows a unit with very good phase locking. Pure tone firing patterns, visualized as a dot raster, showed strong onset responses across a wide range of frequencies, followed by inhibition (Figure 4B). Sustained firing for the remainder of the tone duration occurred only at a narrow bandwidth centered around the CF. The frequency response area (FRA) further illustrates the sharp tuning with the CF at 16 kHz and a minimum threshold of 20 dB (Figure 4C).

The same unit, when presented with broadband noise, responded to the sound onset with a well-timed single spike response with strong inhibition for the remainder of the UM noise stimulus (Figure 4D). Upon transition to the AM stimulus (1-2 s), a pronounced increase in firing was observed, but mainly at the higher modulation depths. Sharp peaks in the cycle histograms exemplify strong phase-locking at 16 Hz (red), 64 Hz (green) and 128 Hz (blue; Figure 4E). Even at 512 Hz, the firing rate co-modulated with the stimulus period at large modulation depths. VS_{PP} showed a progressive increase for modulation frequencies 16, 64 and 128 Hz, up to a remarkable precision of >0.9 at full modulation depth (m = 1; Figure 4F). The VS_{PP} at 512 Hz also increased to about 0.4 at a modulation depth of 0.5, but decreased at full modulation. In this unit the firing rate also depended strongly on the modulation depth of the AM stimulus (Figure 4G).

Figure 5 illustrates a central nucleus unit with a less well-timed response to AM stimuli. The unit was sharply tuned, and highly sensitive at its CF of 22.6 kHz, with a minimum estimated threshold of 0 dB, the lowest intensity that was presented (Figure 5C). At higher intensities the spectral sensitivity broadened. At CF, tone responses showed the shortest onset delay, and firing was sustained during the tone presentation. Following sustained firing during the tone offset inhibition was observed. Inhibitory sidebands were also present, especially at frequencies above CF (Figure 5B).

The unit showed a sustained response to UM noise with a well-timed onset (Figure 5D). Firing rate generally increased with modulation depth at low- to mid-range modulation frequencies. However, the relationship between firing rate and modulation depth was non-monotonic at modulation frequencies of 64 and 128 Hz, where firing rate initially increased, but then decreased abruptly with increasing modulation depth (Figure 5D,G). A non-monotonic relationship between firing rate and modulation depth was not uncommon among the sampled units. The decrease in firing rate at 64 Hz was so large that the VS_{PP} became very low as well owing to the many trials without a single spike (Figure 5E). This

artificial decrease in VS_{PP} was not observed at 128 Hz, where the decrease in firing rate was less than at 64 Hz.

The units illustrated in Figure 4 and 5 thus illustrate the well-known, large functional heterogeneity of IC neurons (Palmer et al. 2013).

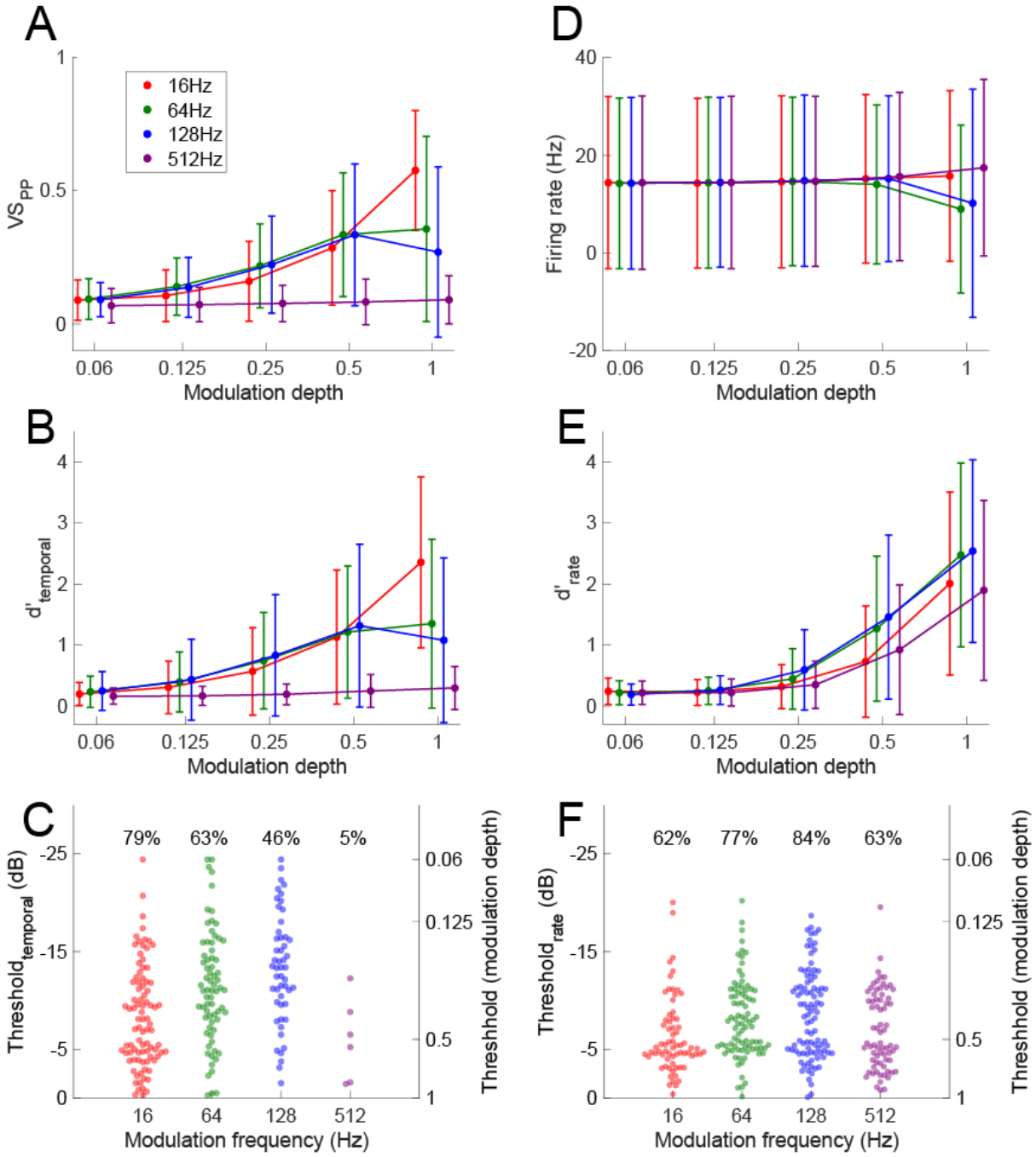
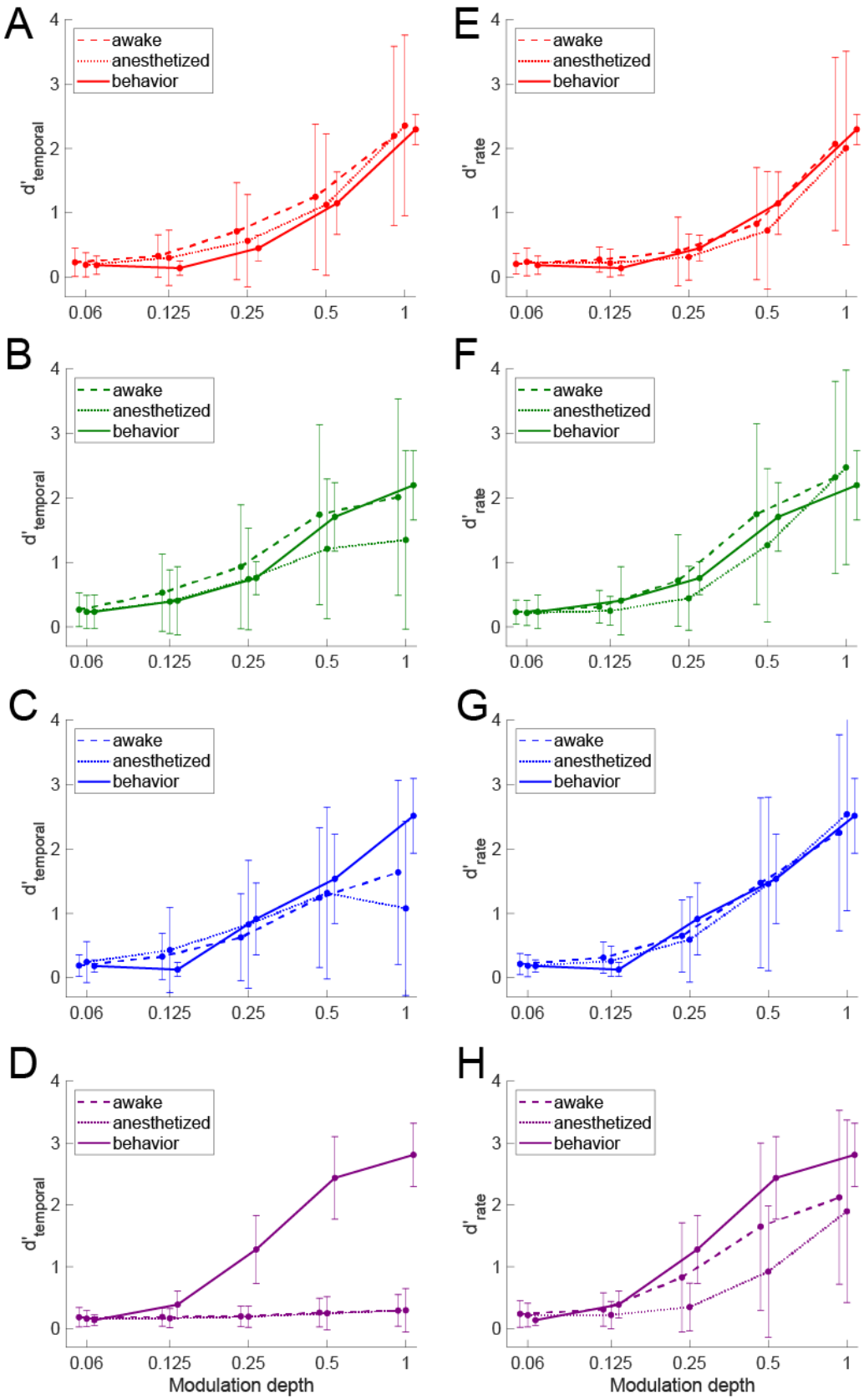


Figure 6 Population AM encoding. (A) Average relation between VS_{PP} and modulation depth. (B) Conversion of VS_{PP} to d' using ROC analysis. (C) AM detection thresholds. Percentages denote the percentage of cells for which a threshold could be found. Thresholds were defined as the modulation depth for which d' = 1. (D-F) As A-C, but for firing rate instead of vector strength. Error bars in the figure show SD. All AM data shown are for a carrier noise intensity of 60 dB SPL.

Because of the heterogeneity of IC neurons, we next looked at the response of the population as a whole. Figure 6 shows the average VS_{PP} (Figure 6A) and average firing rate (Figure 6D) of the population measured in response to the different AM stimuli. Since each stimulus repetition yields a VS_{PP} value, we could apply ROC analysis to convert both VS_{PP} and firing rate to d'. AM detection thresholds were calculated analogously to the behavioral detection thresholds, using a d' of 1 as the cutoff. The population mean curves of d' calculated from the VS_{PP} ($d'_{temporal}$; Figure 6B) closely resembled the population mean curves of VS_{PP} themselves (Figure 6A). VS_{PP} on average increased with increasing modulation depths, except in the case of 64 and 128 Hz where VS_{PP} decreased at the highest modulation depths (i.e. non-monotonic pattern). The phase locking performance for 512 Hz was overall not very good (Figure 6A,B). This pattern was also reflected in the detection thresholds. A detection threshold to 16 Hz modulations could be established in 79% of the units (Figure 6C). This proportion decreased thereafter with increasing modulation frequency, with 64% and 46% of units showing detectable thresholds for 64 and 128 Hz, respectively. This threshold was on average at a lower modulation depth than for 16 Hz, reflecting the overall better phase-locking at these frequencies. In contrast, only 5% of the population showed a threshold for AM at 512 Hz, illustrating the generally poor phase locking at this modulation frequency.

The average firing rate of the IC units showed a generally flat response curve with a large standard deviation (Figure 6D). This variability was due to the large variability in noise-evoked firing rate as well as its modulation by AM. A small increase in the average firing rate can be seen for both 16 and 512 Hz at high modulation depths, whereas for both 64 and 128 Hz the firing rate on average showed a clear decrease at high modulation depths. The strong inhibition of firing rate in some units led to the decrease in mean VS_{PP} at these frequencies (Figure 6A), as a trial with no spikes is defined to have trial VS of 0, which is ultimately averaged across repeated trials to obtain VS_{PP}. Generally, this decrease in VS_{PP} reflects the difficult of assessing phase-locking when firing rate is very low.

Conversion of firing rate to d' by ROC analysis created a more straightforward measure of performance (Figure 6E). At all four tested modulation frequencies, a gradual increase in d' with increasing modulation depth was observed (Figure 6E). A measurable detection threshold was observed in most cases for the mid frequencies 64 and 128 Hz, whereas at the lowest and highest frequencies this percentage was somewhat lower. In the cells with a measurable threshold, this threshold was on average similar across frequencies.


A
B
C
D
E
F
G
H
awake
anesthetized
behavior
d'temporal
d'rate
0
1
2
3
4
0.06
0.125
0.25
0.5
1
Modulation depth

272

Figure 7 Comparing neurometric temporal and rate AM encoding of awake (dashed lines, n = 99 units) and anesthetized (dotted lines, n = 112 units) recordings to behavioral performance (solid lines, n = 8 animals). (A-D) VS-derived d' plotted against modulation depth. (E-F) Firing rate-derived d' plotted against modulation depth. mean ± SD. Modulation frequency color scheme follows that of previous figures: A,E (red) = 16 Hz, B,F (green) = 64 Hz, C,G (blue) = 128 Hz and D,H (purple) = 512 Hz. All AM data shown are for a carrier noise intensity of 60 dB SPL.

As a control for potential effects of anesthesia, we also recorded a total of 99 units in awake, passively-listening animals (n = 4 mice). Similar results were obtained in awake and anesthetized animals both with regard to phase-locking and firing frequencies at different modulation frequencies (Figure 7). Conversion of the neurometric measures of temporal (VS_{PP}) and rate (firing rate) coding to d' made it possible to compare the neural AM detection performance in the anesthetized and awake animals with the psychometric performance. At a modulation frequency of 16 Hz, behavioral performance was similar to the neurometric performance, both for temporal and for rate coding (Figure 7A and E, respectively). Both neurometric measures of AM coding showed similar trends as the psychometric measure at the midrange frequencies 64 Hz (Figure 7B, F) and 128 Hz (Figure 7C, G), with the exception of the drop in temporal coding performance seen at full modulation, as previously described. At 512 Hz, the d' based on VS was much smaller than the behavioral measure (Figure 7D), whereas the d' based on firing rate was more similar (Figure 7H).

The overall performance of the awake units came closer to the behavioral performance. Especially at the mid frequencies, the relation between $d'_{temporal}$ and modulation depth was less nonlinear (Figure 7B,C). At the highest modulation frequency, the d'_{rate} matched the behavioral d' better than its counterpart in anesthetized animals (Figure 7H). However, $d'_{temporal}$ at 512 Hz showed similarly low values as for the anesthetized animals (Figure 7D), indicating that phase locking at the level of the IC unlikely plays a prominent role in encoding such high modulation frequencies.

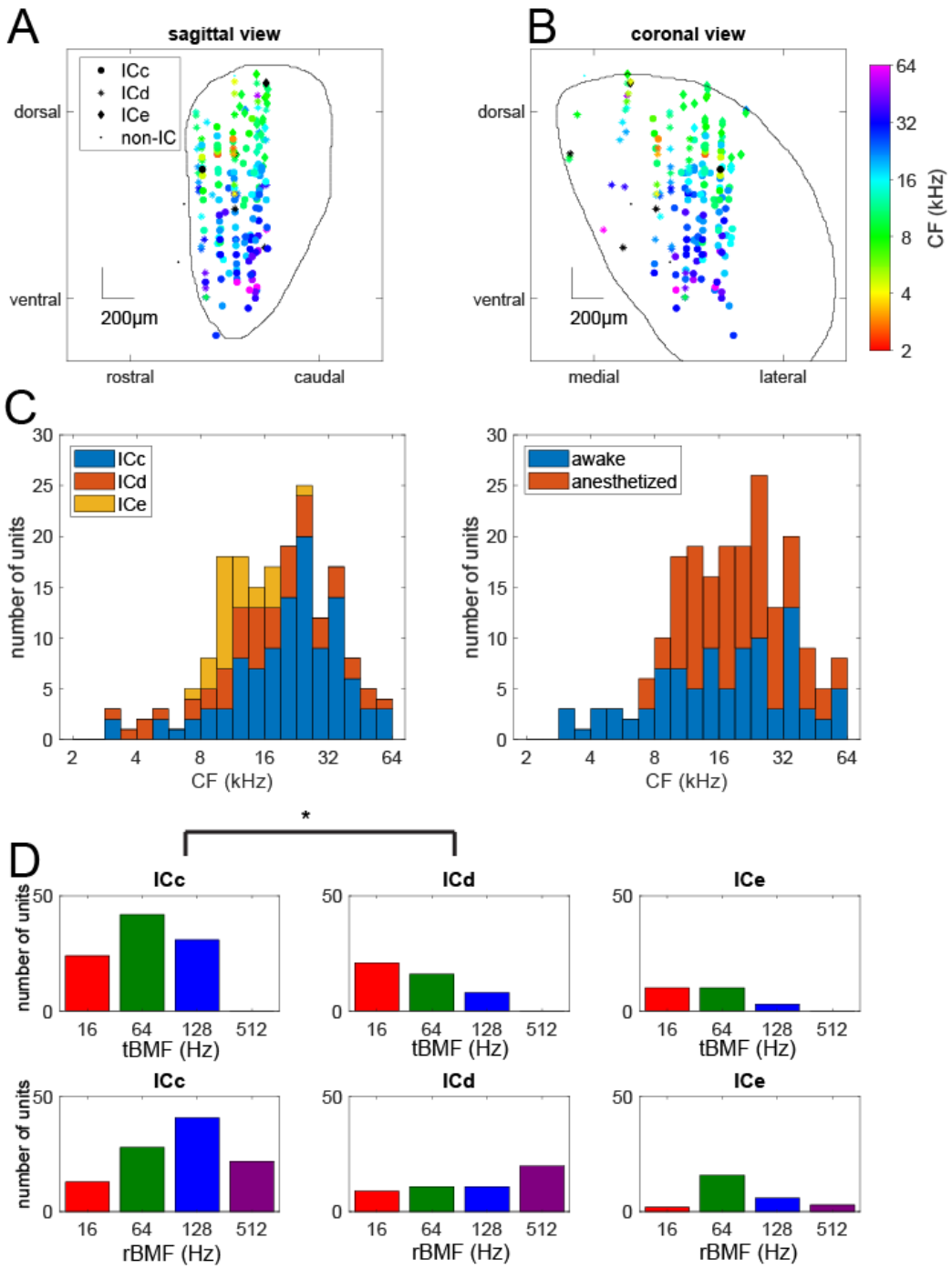


*Figure 8: Spatial organization of pure tone and AM encoding of recorded units. Location of the units was obtained by registering probe tracts to the mouse Allen Brain Atlas. (A) Sagittal view of units in the left IC. Color of symbols represents CF and shape represents location of units within subnuclei of the IC. (B) as (A), but in coronal view. (C) Left: Distribution of CFs in the different subnuclei of the IC. Right: Distribution of CFs in recordings from anesthetized and awake recordings. (D) Top: distribution of temporal best modulation frequencies (tBMF) in the different subnuclei of the IC. * denotes a*

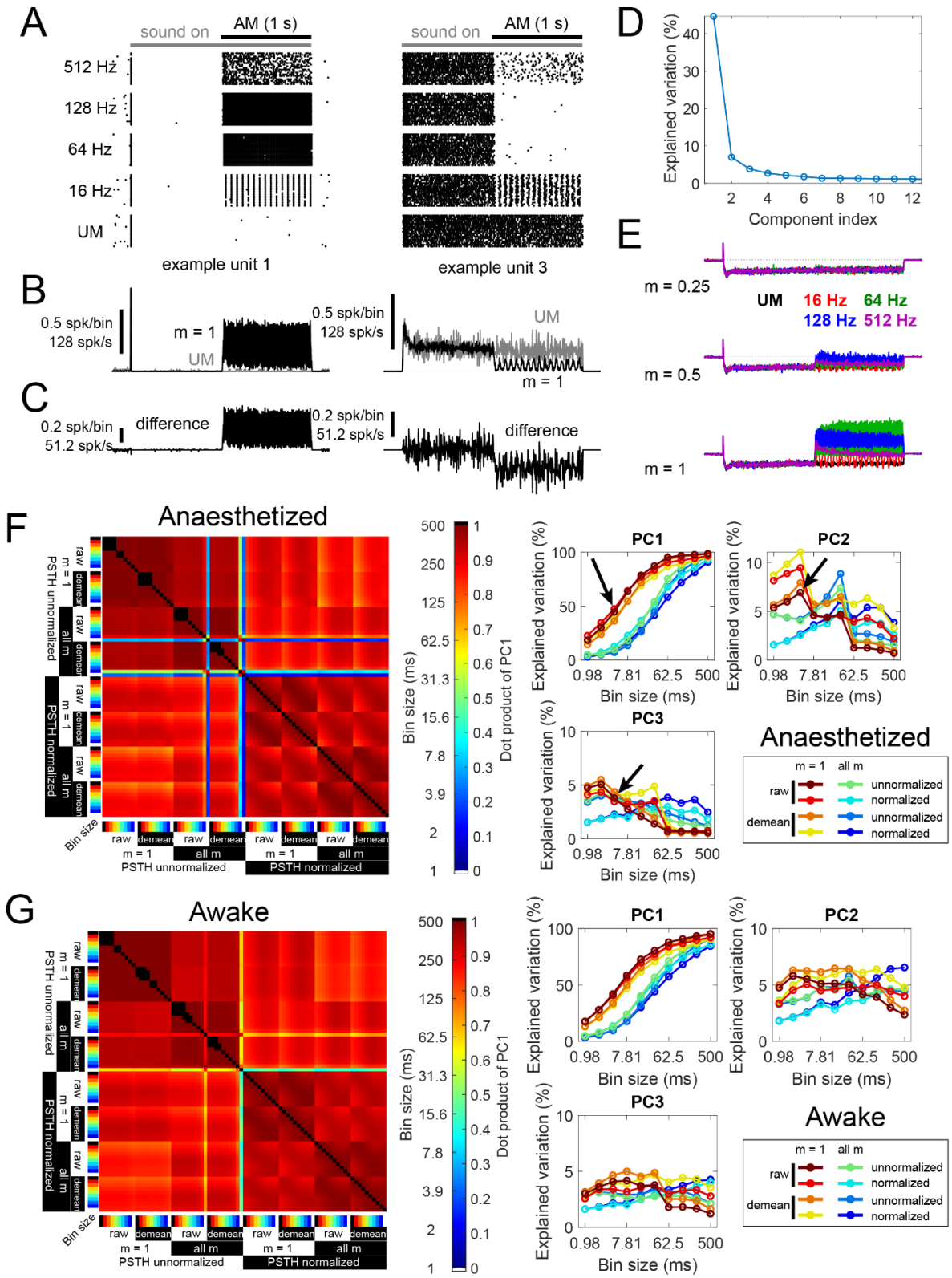
significant difference in tBMF distribution between ICc and ICd (Kruskal-Wallis test followed by Bonferroni post hoc test; p = 0.025). Bottom: distribution of rate best modulation frequencies (rBMF) in the different subnuclei of the IC.

In order to investigate potential differences between the subdivisions of the IC in the encoding of AM stimuli, the recording electrodes were coated with the fluorescent dye DiO or DiI, and the probe tracks were reconstructed following postmortem histology. The location of individual units could be determined by aligning the histological sections to the Allen Brain Atlas CCFv3 (Wang et al. 2020). All single units obtained from anesthetized recordings (n = 112) and 82% (n = 81/99) of the units from awake recordings could thus be aligned to the reference atlas. The units originated from a large part of the IC except for its most lateral and caudal part. The alignment also allowed us to assign each unit to a subnucleus within the IC based on the borders defined by the Allen Brain Atlas CCF. The proportion of units located within the central (ICc), dorsal (ICd) and external (ICe) subnuclei of the IC were 57%, 27% and 15%, respectively.

Plotting the CF of the units versus their location revealed the tonotopic gradient of the IC along its dorsoventral axis, in both a sagittal (Figure 8A) and a coronal view (Figure 8B). The distribution of CFs in the IC central nucleus neurons spanned the entire frequency range of presented stimuli, but peaked between 16 and 32 kHz. Neurons sampled from the dorsal nucleus had a similar distribution of CFs as in the central nucleus, while that of the external nucleus clustered between 8 and 16 kHz (Figure 8C left). The CF distribution of the neurons recorded under anesthesia and awake were generally similar (Figure 8C right).

We did not observe any obvious spatial clustering or gradient in AM encoding, as measured by the d' thresholds calculated from VS and firing rate (results not shown).

The preferred modulation frequencies were compared between the sampled neurons of the subregions of the IC. This was done for both their tBMF and rBMF. The tBMF distribution pointed towards a low-pass pattern in the dorsal and external nuclei, and a band-pass pattern in the central nucleus. This pattern differed statistically significantly between the central and dorsal nuclei (Kruskal-Wallis test followed by Bonferroni post hoc test; p = 0.025).


A
sound on
AM (1 s)
512 Hz
128 Hz
64 Hz
16 Hz
UM
example unit 1
example unit 3
B
0.5 spk/bin
128 spk/s
m = 1
UM
C
0.2 spk/bin
51.2 spk/s
difference
D
Explained variation (%)
Component index
E
m = 0.25
m = 0.5
m = 1
UM
16 Hz
64 Hz
128 Hz
512 Hz
F
Anaesthetized
PSTH unnormalized
PSTH normalized
Bin size
Dot product of PC1
PC1
PC2
PC3
Explained variation (%)
Bin size (ms)
G
Awake
unnormalized
normalized
demean
raw

328

329

Figure 9 Principal component analysis of the difference in responses to AM or UM stimuli. (A) Dot raster plot of two example units in response to fully modulated stimuli at various frequencies and the unmodulated control stimulus (UM). (B) Average peristimulus time histograms (PSTHs) of the same units for all fully modulated stimuli (black; m = 1) and for the unmodulated control stimulus (gray). Bin size: 1/256 Hz ≈ 3.9 ms. (C) Difference between both PSTHs in B ($PSTH_{diff}$), which was used as the input to the principal component analysis (PCA). (D) Scree plot showing the variation explained by each of the principal components in descending order. Only the first 12 components are shown. (E) Average amplitude of PC1 in response to stimuli of various modulation frequencies increased with modulation depth. Bin size: 1/256 Hz ≈ 3.9 ms. (F-G) Robustness of PCA results against parameter choice in both (F) anaesthetized and (G) awake datasets. Left: Dot products of PC1 from PCA using different PSTH bin sizes and parameters: whether the PSTH of each neuron was first normalized ('normalized' vs 'unnormalized'), whether input was the difference between unmodulated response and the full modulation ('m = 1') or stimuli of all modulation depths ('all m'), and whether the mean was subtracted from the PCA input ('raw' vs 'demean'). The generally high value for the dot products of the vectors of weights indicates high similarity. Right: Explained variation plotted for the first three principal components (PC1-3) as a function of PSTH bin size (abscissa) and other parameter choices, as indicated in the boxes. Note the expanded y-scale for PC2 and PC3. Arrows indicate the parameter set shown in D-E and used for the remainder of the article.

Our comparison of behavioral and neurometric thresholds showed some agreements, but also some clear discrepancies, and also showed that a strict rate code or temporal code would not provide a convincing match between the neurometric and the behavioral performance. To further explore how the heterogeneous neuronal responses contribute to the behavioral detection of AM sound, we therefore simulated how an animal would utilize spike information to perform our behavioral task by combining dimensionality reduction with ideas from diffusion decision models (Ratcliff et al. 2016). Since our mice were trained to distinguish AM sounds from unmodulated sounds regardless of their frequencies, we postulate that they must be relying on the difference between the neuronal activity evoked by the different modulation frequencies at the largest modulation depth (m = 1) and the neuronal activity evoked by the unmodulated sound (Figure 9A-C). In order to allow cycle-to-cycle variation and minimize bias towards rate coding, we created peristimulus-time histograms (PSTHs) with a bin size of only 1/256 Hz ≈ 3.9 ms, and calculated the average PSTH difference ($PSTH_{diff}$) between unmodulated and fully modulated trials. We then performed a principal component analysis (PCA) across the neuronal dimension on this difference, and found that the first component (PC1) could already explain 44% of the variation (Figure 9D), with the total variation being the sum of the squared $PSTH_{diff}$. The average amplitude of this PC1 showed a graded increase with increasing modulation depth across all frequencies (Figure 9E). This result was very robust against choices of bin size and data normalization (Figure 9F,G). Normalization and mean-subtraction in general had little effect on the variation explained by PC1. Including all modulation depths generally decreased explained variation. Most prominent, however, was the effect of temporal binning. Larger bins, i.e. smoother data, allowed PC1 to explain more of the overall variability, presumably by averaging out temporal variability in both noise and signal. Despite differences in explained variation, the generally high dot product values indicate that all PC1s have largely similar weights for individual neurons. In the end, we chose not to perform normalization and mean-subtraction to avoid over-representation of low firing rate units, and to focus on the difference in firing rate. Un-normalized input also means the resulting principal components remain in the firing rate scale.

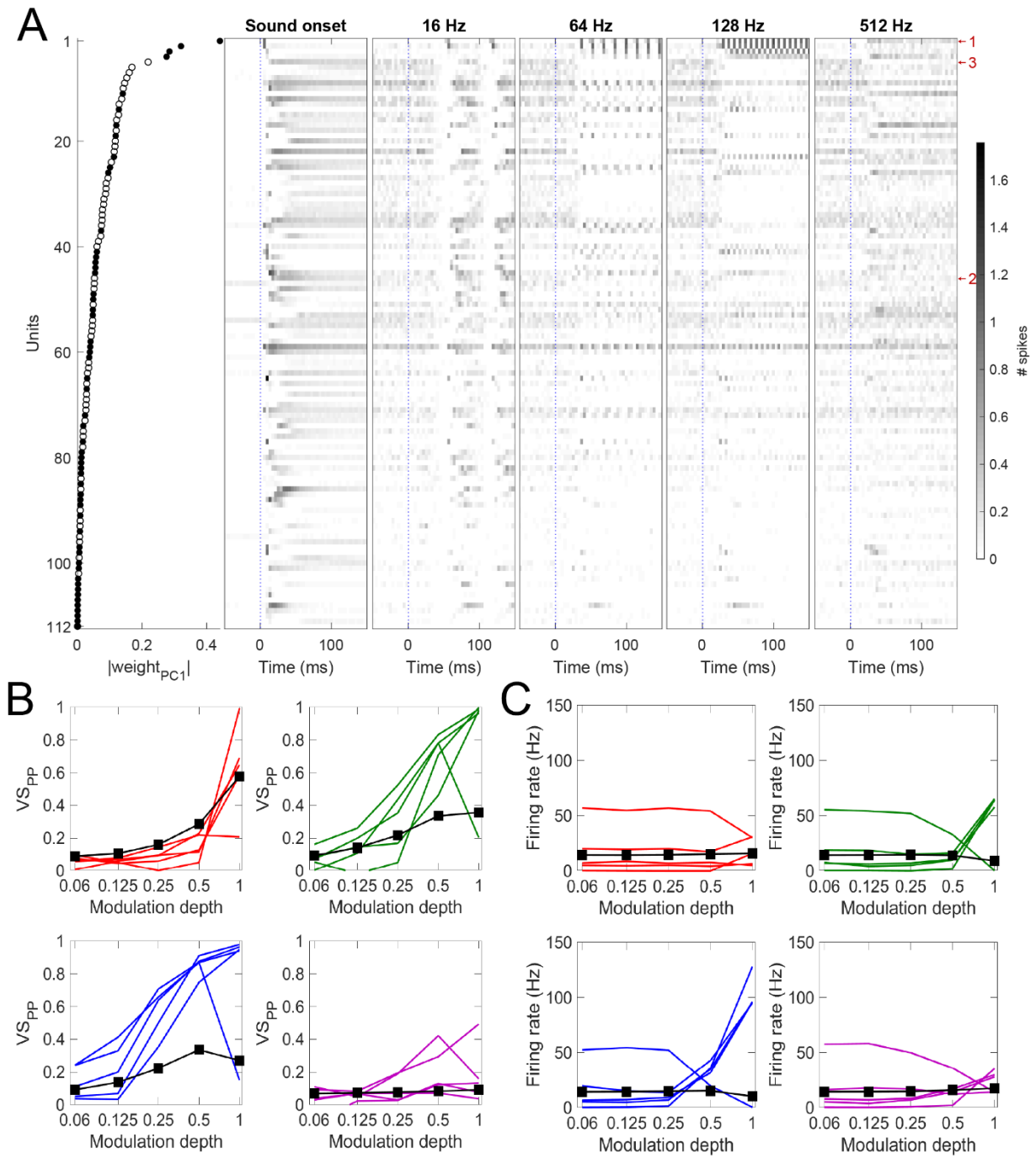


Figure 10. Overview of neurons from anaesthetized recordings ranked according to their contribution to PC1. (A) All neurons from anaesthetized recordings sorted by their contribution to PC1 (left panel, |weight$_{PC1}$|), their average PSTH in response to sound onset (second from left) or to 16, 64, 128 or 512 Hz AM at full modulation. Closed and open symbols indicate positive and negative weights, respectively. Time is relative to either sound onset or transition to AM noise (blue, vertical dotted lines). Bin size: 1/256 Hz ≈ 3.9 ms. Small red arrows with numbers indicate example units. (1: Figure 4, 9A-C; 2: Figure 5; 3: Figure 9A-C) (B) VS$_{PP}$ of top 5 units as a function of modulation frequency and depth. (C) Firing rate of top 5 units as a function of modulation frequency and depth. Modulation frequency color scheme in (B) and (C) follows that of previous figures: red = 16 Hz, green = 64 Hz, blue = 128 Hz and purple = 512 Hz.

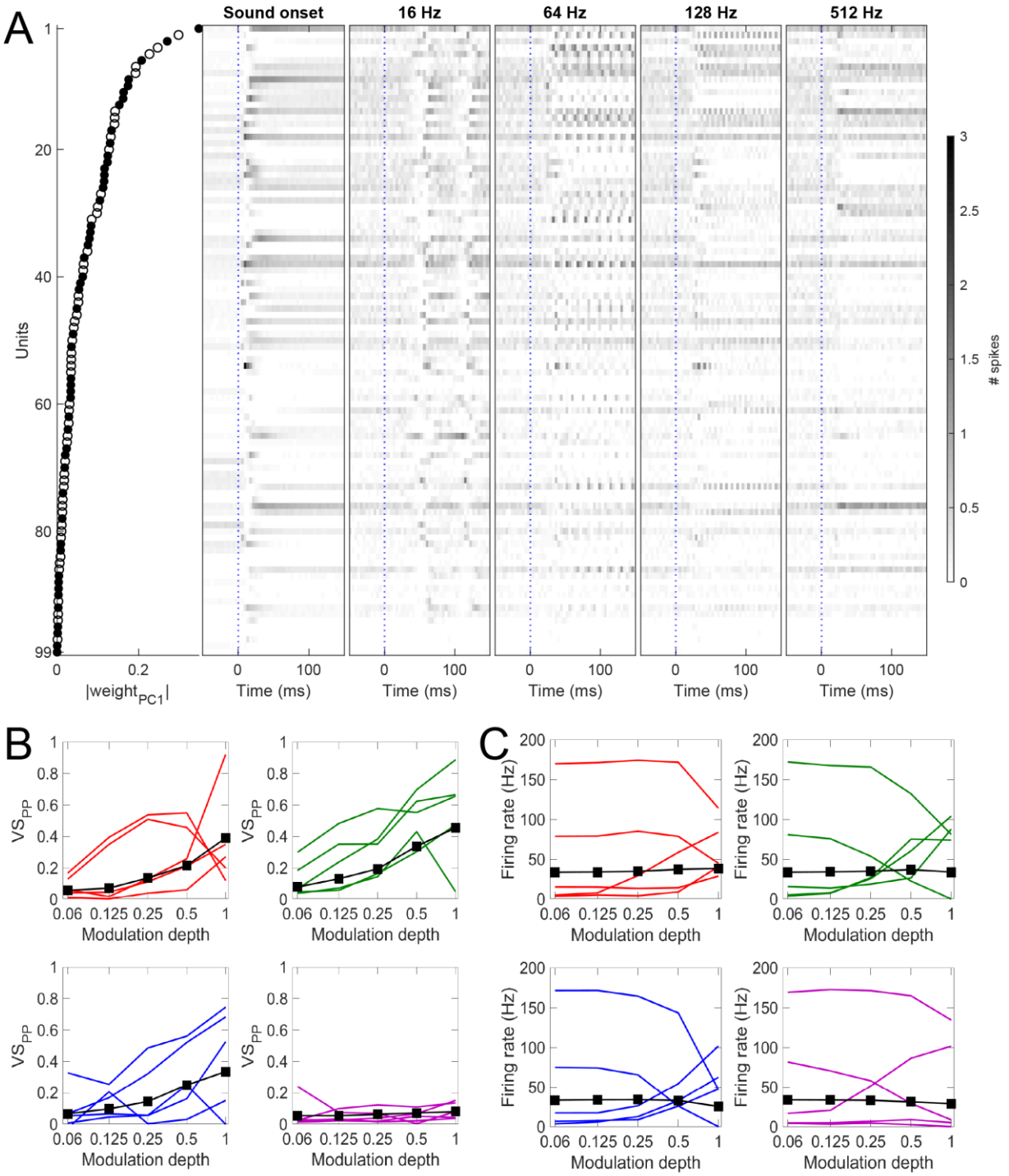


Figure 11. Overview of neurons from awake recordings ranked according to their contribution to PC1. (A-C) As Figure 10A-C, respectively, except recordings were from awake animals.

In Figure 10 we sorted all neurons from anesthetized recordings according to their (absolute) weighting in PC1. The neuron illustrated in Figure 4 had the largest weight, whereas the neuron of Figure 5 came out halfway. Although the top 4 units showed an onset response to the broadband noise carrier (Figures 9A,B, 10A) and off-CF tones (Figure 4), onset response alone did not predict a good response to AM as demonstrated by the numerous onset units poorly represented in PC1. The next few units showed a mixture of sustained and build-up responses to broadband noise. Some lower ranking units were great phase lockers, but showed little change in firing rate upon transition into AM.

In Figure 11 we did the same analysis for the units from awake recordings. Top awake units showed a mixture of onset, sustained and build-up responses to broadband noise. Combining the awake and anesthetized recordings, the distinguishing feature of top neurons appears to be a strong change in firing rate upon AM (Figures 10B, 11B), although many of them are also good phase-lockers (Figures 10C, 11C). On the other hand, the vast majority of the neurons did show modulation in their firing rate to 16 Hz AM.

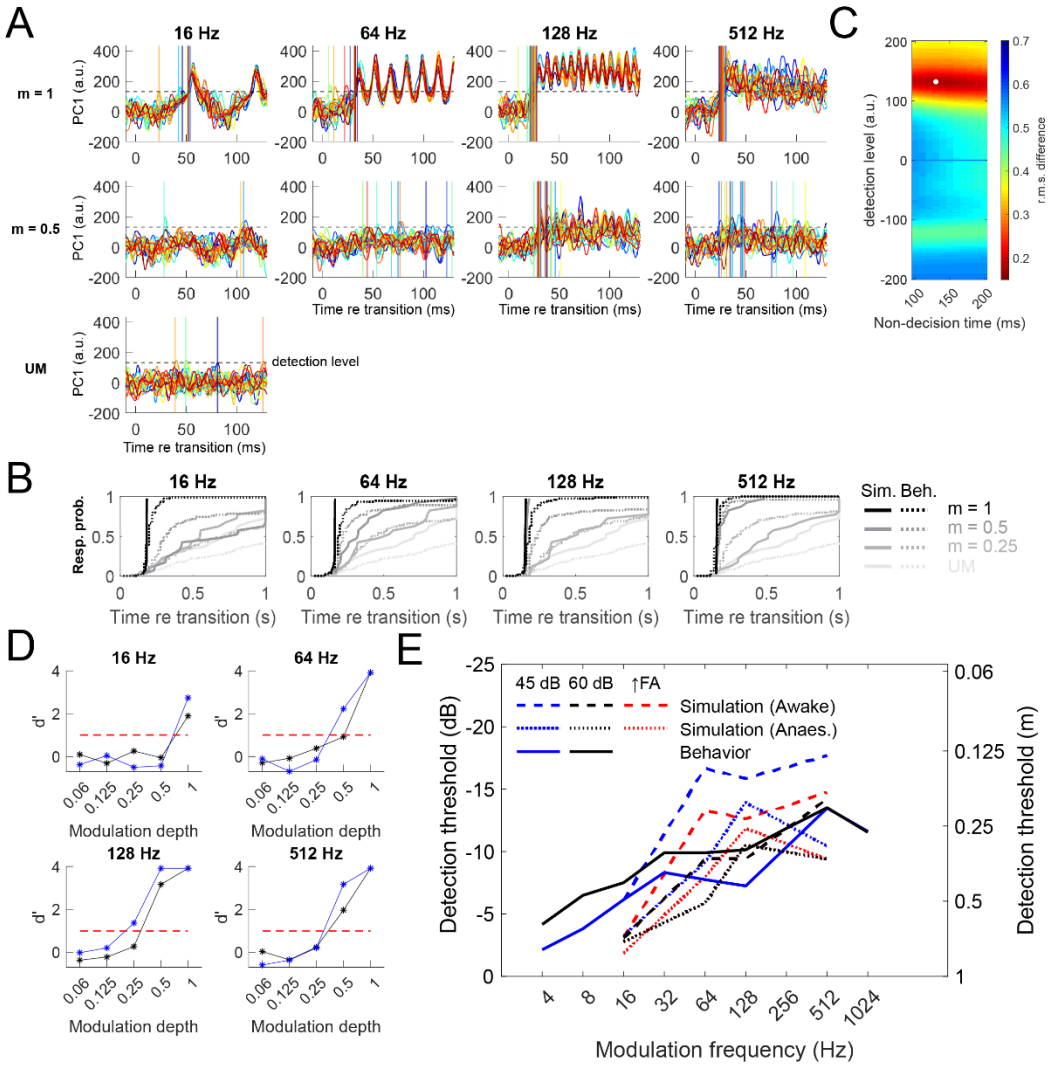


Figure 12 Using PCA of neuronal responses to AM stimuli to infer behavioral thresholds. (A) Single trial evolution of PC1 over time for selected modulation frequencies and depths at 60 dB for the data from the anaesthetized mice. Traces were created from kernel spike density functions with a Gaussian kernel (sigma: 2 ms). Horizontal lines represent the optimal "detection level"

of the model that best fitted behavioral response latencies combined across all mice, and vertical lines represent the point at which PC1 exceeded this level, with colors corresponding to the PC traces. The model adds an additional non-decision time of 130 ms. (B) Comparison of CDF of simulated and behavioral response latency at different modulation frequencies and depths. (C) Root-mean-square error to response latency CDF as a function of parameters. White dot indicates location of the minima (non-decision time: 130 ms; detection level: 132 a.u.) (D) d' calculated from simulated response latencies for 45 dB (blue) and 60 dB (black). (E) AM detection thresholds derived from the model compared to those from behavior. Red broken lines represent simulation thresholds with increased false alarm rate for 45 dB.

While we performed the PCA with the trial-averaged response, analogous to the training experience of the animals, during the behavioral task they only received information from single trials for each response decision. We simulated this by projecting the single trial responses onto PC1 (Figure 12A), followed by measuring how long it took to cross a decision level at each trial. The decision level was obtained from a fit to the data. This threshold crossing readily captured the modulation frequency dependence in response latencies (Figure 3D) at full modulation depth (m = 1; Figure 12B). Interestingly, at a modulation depth of 0.5, the phase-locked fluctuation of PC1 at 16 Hz quite often did not reach the threshold, despite the vector strength being comparable to the 64 and 128 Hz stimuli (Figure 6A). It is possible that the longer modulation period at 16 Hz led to greater variability in spike time among neurons, making synchronous firing less likely.

To compare the model with behavior, we added an additional latency for downstream processes such as motor response time, resulting in a set of 20 simulated response latencies for each stimulus (Figure 12B). We similarly calculated *d'* values from these latencies and estimated detection thresholds for different modulation frequencies (Figure 12D). The thresholds estimated from simulated detection latencies for the 60 dB stimuli captured the high pass pattern in the behavioral response, with 16 Hz being poorly distinguished from unmodulated noise, but progressively better discrimination for increasing modulation frequencies (Figure 12E). The sensitivity at 512 Hz modulation, however, did decrease slightly for the anaesthetized data, as was also apparent in the population firing rate d' (Figures 6E). A slightly lower simulated false alarm rate at 45 dB (especially with awake data) led to an overestimation of AM sensitivity, which was corrected when false alarm rate at 60 dB was used instead to calculated *d'* (red broken lines in Figure 12E; see Discussion). Overall, we conclude that our model provided a good fit of the detection thresholds of the behavioral data.

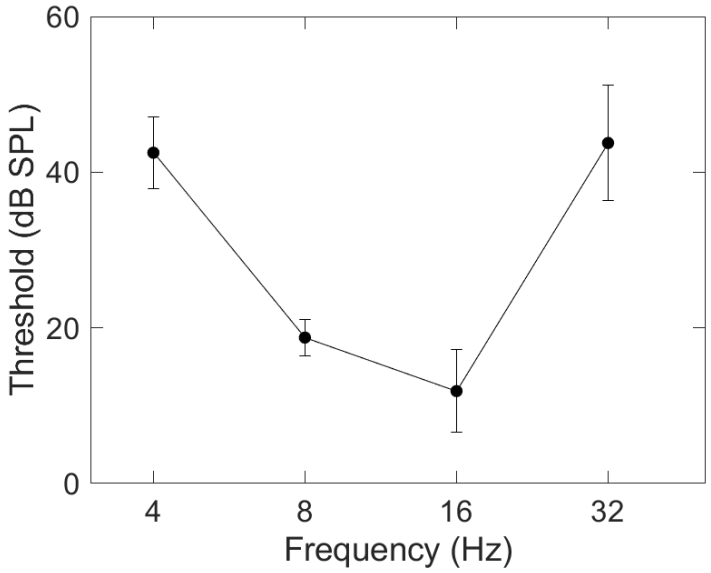


Figure S1. Hearing threshold of the mice. ABR thresholds (mean ± sd) of the B6CBAF1/JRj mice used in the behavioral experiments in this study (n = 8).

Discussion

We compared psychometric and neurometric detection of amplitude modulations in noise in the mouse. We used a new stimulus, which precluded a contribution of spectral broadening as a cue for the AM detection. We found that AM detection by the mouse shows high-pass characteristics, with relatively poor performance at the lowest tested frequency (4 Hz) and very good performance at a high frequency (512 Hz). We compared these behavioral thresholds with neurometric thresholds based on the AM responses of single units in the IC. The poor phase locking at 512 Hz and good phase locking at 16 Hz already indicated that a purely temporal code could not provide an adequate description of the behavioral results. In contrast, a PCA based threshold method that took latencies into account did provide an adequate fit, especially in awake animals. The PCA-based comparison suggested that a few neurons had a disproportionate effect on the overall performance. These neurons showed large and rapid changes in firing rate upon the transition to AM. Because their responses were well timed, their contribution to the population response should be substantial, especially for high modulation frequencies, where response jitter was low.

<u>Advantages and disadvantages of the method used.</u>

Sinusoidal amplitude modulation of a tone results in the addition of a sideband both below and above the carrier frequency. We used noise as a carrier, meaning that amplitude modulation resulted in spectral broadening. For low modulation frequencies, this spectral broadening is small and thus unlikely to be detected. However, at the high modulation frequencies at which the mouse excelled, the spectral broadening becomes substantial, raising the question whether the mouse used not only amplitude, but also spectral cues. We therefore used a new method to compensate for the addition of sidebands due to AM. With these stimuli, we are confident that the AM detection thresholds we measured were not contaminated by spectral or intensity cues. The method described in our paper should be useful to retest the measured sensitivities at high modulation frequencies in other studies, to exclude a possible contribution of spectral broadening as a cue for AM detection. A possible disadvantage of our method is the presence of an unmodulated side band, which may contribute to the overall behavioral performance.

For the detection of AM, we used a Go-NoGo test in which the mouse had to detect a change from noise to AM. Similar paradigms have been used in a variety of species to detect psychophysical AM modulation thresholds for different modulation frequencies (Klump and Okanoya 1991; Moody 1994; Dent et al. 2002; Bremen et al. 2017). The paradigm is relatively easy to learn. Its main disadvantage is that it is relatively hard to control the false-positive lick rate, which was about 30%.

<u>Comparison of behavioral AM detection with earlier study in the mouse and with other species</u>

The mice were remarkably good at detecting AM at high modulation frequencies. At 512 Hz, their detection threshold of 21% (-13.5 dB) was lower than for other species measured previously (Dent et al. 2002; Carney et al. 2013). Only rats, with a threshold of ~23% at 500 Hz had a similar sensitivity as the mouse at high-frequency modulations (Kelly et al. 2006). Our study is in general agreement with a recent behavioral study in mice, in which a band-pass characteristic was also found for the behavioral detection, with the most sensitive detection of ~13% at 160 Hz, and detection at 640 Hz at ~21% (Cai and Dent 2020). Discrepancies with our study may be related to slight differences in procedures, for example Cai and Dent (2020) presented a single modulation frequency per session whereas we

475 presented them all pseudorandomly within the same session, our method of constant stimuli versus
476 their adaptations of modulation depths, and to differences in the definition of the detection limit. We
477 observed a dip at 128 Hz whereas Cai and Dent (2020) observed the lowest threshold at 160 Hz.

478 The mice were not very good at detecting the lowest modulation frequencies, in sharp contrast with for
479 example humans. Humans possess exquisite sensitivity at 4 Hz, close to the peak of the AM spectrum of
480 speech (Ding et al. 2017) with thresholds as low as 5% (Viemeister 1979). The thresholds of the mice
481 may slightly improve with longer stimulus durations (Sheft and Yost 1990; Moody 1994; Lee and Bacon
482 1997; O'Connor et al. 2011), but this is unlikely to explain their relatively poor low-frequency
483 performance. The relatively low thresholds at high modulation frequencies may enable mice to detect
484 for example the fast amplitude modulations in brief ultrasonic vocalizations (Cai and Dent 2020).

485 Neural responses to tones or AM.

486 The responses to tones were quite heterogeneous, resulting in a large diversity of FRAs, in agreement
487 with previous results (reviewed in (Ehret and Schreiner 2005). We did not attempt to classify the
488 different FRAs, as detailed studies have shown that the FRAs form a continuum, and that clustering into
489 discrete categories is not feasible (Hernández et al. 2005; Palmer et al. 2013).

490 The responses to SAM noise generally matched earlier reports (reviewed in (Joris et al. 2004; Rees and
491 Langner 2005). In mice, responses to AM are generally heterogeneous (Walton et al. 2002; Geis and
492 Borst 2009; Ono et al. 2017). A detailed comparison with earlier reports in mice is not opportune, since
493 we used only a limited set of frequencies, and mostly focused on the effects of modulation depth.
494 However, a few remarks regarding the limits for phase locking and the presence of a non-monotonous
495 relationship between modulation depth and firing rate can be made.

496 The absence of phase locking at 512 Hz in most cells is in agreement with the lack of a phase-locked
497 response in the membrane potential at a modulation frequency of 640 Hz during whole-cell recordings
498 (Geis and Borst 2009). In other species phase locking at even higher frequencies (up to 1200 Hz) has
499 been observed (reviewed in (Rees and Langner 2005). The lack of overall good phase locking at
500 modulation frequencies above 500 Hz does mean that there already must be some degrading of
501 temporal information at the level of the IC, since in mice, phase locking to frequencies >500 Hz is still
502 prominent at the level of the lower brainstem (Pauli-Magnus et al. 2007; Müller et al. 2019). Relatively
503 often, especially in anesthetized animals, we observed a strong decrease in firing rate at full modulation
504 depth at the intermediate frequencies (64 and 128 Hz) after an initial increase with increasing
505 modulation depth. In the rMTF, which is typically based on stimuli at full modulation depths, these
506 would be band-reject neurons. Such responses are also relatively frequently observed in rat and gerbil
507 (Zhang and Kelly 2003; Krebs et al. 2008). Interestingly, these neurons showed a more band-pass-like
508 character at low modulation depth. Onset inhibition may underlie the decrease in firing rate at high
509 modulation depth (Geis and Borst 2009), suggesting that anesthesia may favor these well-timed
510 inhibitory inputs.

511 Our use of phase-projected vector strength avoids the spurious high vector strength values at very low
512 spike rates encountered with traditional calculations (Yin et al. 2011). It also provides a trial-by-trial
513 phase locking metric convertible to a d' value through ROC analysis. This allowed us to better compare
514 between temporal and rate coding of AM sound, and to compare these neurometric measures to the
515 psychometric measurements obtained.

516 Effect of sound intensity

517 We found a slight improvement of behavioral AM detection at 60 dB SPL compared to 45 dB SPL. In
518 other species the effects have not been very pronounced either. There was no clear effect of sound
519 intensity on behavioral AM detection in parakeets or chinchillas (Dooling and Searcy 1981; Salvi et al.
520 1982). Owls and starlings and humans detect AM slightly better at higher sound intensities (Klump and
521 Okanoya 1991; Kohlrausch et al. 2000; Dent et al. 2002), similar to what was observed here, whereas
522 macaques have slightly worse performance at higher sound intensities (Moody 1994). The mechanisms
523 underlying intensity-dependent changes are complex, and will depend among others on the tuning
524 properties of the units that are differentially recruited (Rees and Møller 1987; Wang et al. 2021).

525 Effect of anesthesia

526 Interestingly, the data obtained in awake animals approached the high-pass characteristics (Figure 7)
527 better than the data obtained under anesthesia. This was also the case for the d' of the VS or of the
528 firing rate, suggesting that anesthesia may have had a depressing effect on neuronal responses. In
529 previous studies, anesthesia had variable impact on the auditory system. Ketamine-xylazine may
530 increase minimum threshold and interpeak latencies of ABRs in mice (van Looij et al. 2004). Only a
531 limited number of studies addressed the impact on responses to SAM stimuli. In the rat medial
532 geniculate body, urethane anesthesia did not significantly affect firing rate or VS for modulation
533 frequencies between 2 and 128 Hz (Cai et al. 2016). Only mild effects of pentobarbital plus ketamine
534 anesthesia were observed on responses to AM stimuli in the gerbil IC (Ter-Mikaelian et al. 2007). In
535 recordings from awake animals, the decrease in firing at large modulation depths that was observed in
536 many units seemed less conspicuous. As this decrease in firing rate also may have contributed to the
537 decrease in phase locking, this decrease in non-linearity is likely to have contributed to the better match
538 of the unanesthetized recordings with the behavioral data. Since well-timed, onset inhibition is likely to
539 contribute to this nonmonotonicity (Geis and Borst 2009), the effect of anesthesia may be to enhance
540 this inhibitory component, but intracellular recordings would be needed to test this.

541 Spatial organization of AM tuning

542 We aligned units isolated from our recordings to the Allen Brain Atlas, which allowed us to pool units
543 across animals. This mapping should also facilitate comparisons with future studies in the mouse. An
544 indication that this procedure was successful was that we observed that our dataset displayed the well-
545 known dorsolateral to ventromedial tonotopic organization (Stiebler and Ehret 1985; Portfors et al.
546 2011). We also studied whether there were other gradients or whether there were differences between
547 subnuclei. We observed differences in the tBMF between different subnuclei, as units in the ICd and ICe
548 tended to phase-lock to lower frequencies than in the ICc. In contrast to some previous studies
549 (Schreiner and Langner 1988; Rodríguez et al. 2010; Baumann et al. 2011; Schnupp et al. 2015), we did
550 not find any obvious evidence for a gradient in tBMF or rBMF or local clustering within the IC. However,
551 we cannot exclude that with finer or more extensive sampling, additional organizational principles
552 would have become apparent.

553 Neurometric vs. psychometric detection of AM

554 A main goal of our study was to try to understand which aspects of the neuronal responses within the IC
555 informed the behavioral detection of amplitude modulations. In general, the mice responded to

556 stronger and higher modulation frequency stimuli with shorter response delays. A coding scheme that
557 directly uses phase-locking would not provide a good prediction of the behavioral data, since VS was
558 generally low at the well detected 512 Hz (Figure 7D). Moreover, the short detection latencies observed
559 for the fully modulation stimuli indicated that for these stimuli a modulation in firing that can occur
560 within a cycle must underly the detection mechanism. PSTH-based classification methods (Foffani and
561 Moxon 2004; Malone et al. 2010) including those that uses dimensionality reduction will not shed light
562 on the decision process, as they include information from the whole stimulus duration, which are not
563 available to the animal when it responds early.

564 In light of this, we opted for a two-step process that followed the behavioral training of the animals in
565 our task: Spike responses of the whole stimulus are available to "train" the model to detect AM (PCA-
566 step), but a threshold crossing enables the model to respond at individual trials as soon as enough
567 information is available, while balancing false alarm rates. In particular, as the animals are only
568 concerned with detection of AM, but not to distinguish between different modulation frequencies, we
569 performed a dimensional reduction (PCA) on the difference between the responses to the averaged
570 modulated (AM) and unmodulated (UM) stimuli. This approach was similar to what has been used in
571 demixed PCA (Kobak et al. 2016). Note that the PCA was performed on the neuronal dimension with
572 time points as different "samples". The first principal component already explained >40% of the
573 variance, allowing us to use only this first component to predict the behavioral responses from the IC
574 population data. We used a linear dimensionality reduction method instead of newer, non-linear
575 methods (such as t-SNE(van der Maaten and Hinton 2008), UMAP(McInnes et al. 2018) or PHATE(Moon
576 et al. 2019) for better interpretability. The linear combination of neuronal activities can be regarded as
577 the input to a perceptron (Rosenblatt 1958), a neurologically plausible way for the animal to use this
578 information. As alternatives, PC1 generated in our first PCA step can be substituted with a vector
579 provided by linear classification methods such as linear discriminant analysis or (linear) support-vector
580 machine. However, the position of the classifier (i.e., detection level) will be agnostic to the intrinsic
581 response bias of the animal, and will need to be adjusted in the subsequent step.

582 Because of the clear relationship between modulation depth and frequency with response delay (Figure
583 3), we used a simple level-crossing model that used the size of the population response along the first
584 PCA dimension to obtain a threshold for behavioral detection. Many studies looking at AM encoding do
585 not take the exact timing of behavioral responses into account (Malone et al. 2010; Dong et al. 2013).
586 For our model only two parameters needed to be estimated. A strong point is therefore that we thus
587 base our population analysis on both the magnitude and the timing of the population response. Despite
588 its simplicity, the model overall provided a decent fit of the behavioral data, and it captured the high-
589 pass aspect of the behavioral detection. At 512 Hz, the large changes in firing rates of the top units
590 ensured good performance. At 16 Hz, the predicted performance was much lower, despite decent phase
591 locking at these low frequencies. At low modulation frequencies the envelope amplitude changes are
592 slower than at high modulation frequencies, since with SAM noise, modulation period and rate of
593 amplitude change covary. Even with units of strong phase-locking abilities, the larger temporal spread of
594 their spikes to 16 Hz modulation led to less effective summation across units (Figure 9A) and may
595 explain the relatively poor performance at 16 Hz.

596 Another advantage of dimensionality reduction is that we can rank neurons by their contribution to the
597 principal component and compare their properties. The neurons that contributed most generally had

low spontaneous firing rate, a strong and generally well-timed response to AM (Figures 10B-C, 11B-C), but showed a mixed (onset, adapted or sustained) response to the noise carrier.

Some limitations of the model were the following. We did not explicitly include information about the trial's temporal structure available to the animal. In the simulation, the false positive response in catch trials were driven entirely by random fluctuations of the principal component, i.e. synchronous activity from the component units. This can be triggered by temporal fine structure in the noise carrier or by co-modulation of non-auditory signals in the awake animal. As the noise was not frozen across different recordings, this may explain the "better" detection threshold at 45 dB, where lower firing rate in general decreases chances of PC crossing the detection level for the catch trial. The lack of noise correlation across awake recordings sessions may exacerbate this problem and lead to a further underestimation of false alarm rate at 45 dB. In fact, when we recalculated the simulated threshold at 45 dB using the catch trials at 60 dB, a better fit was obtained (red lines in Figure 12E). An explicit false alarm rate independent of neuronal activity may improve the model. A further limitation is that neurometric and psychometric data were not collected simultaneously. The neuronal AM discriminability may improve in a task-engaged animal compared to our passively listening animals (Niwa et al. 2012a; Dong et al. 2013; Niwa et al. 2015; von Trapp et al. 2016).

Our results agree with the lower envelope principle, which states that decisions to lick were based on a small number of the most sensitive neurons, whereas the many other, less sensitive neurons are largely disregarded (Barlow 1972). It is also in agreement with the idea that the decision decoder that is responsible for the behavioral detection works optimally, since a characteristic of an optimal decoder is that it weighs sensory neurons by their sensitivity (Jazayeri and Movshon 2006; Haefner et al. 2013). A study in gerbils suggested that indiscriminate pooling, rather than a most-sensitive-neuron approach, best explains improved AM tone detection in adults relative to juveniles (Rosen et al. 2010). However, to more firmly establish causal relations and to discriminate between different decoder models, it would be necessary to assess the impact of manipulating the firing of IC neurons, for example using optogenetics (Marshel et al. 2019) or juxtacellular stimulation (Tanke et al. 2018).

Methods

Animals

Wildtype B6CBAF1/JRj mice were obtained from Janvier (n = 8 for behavior; n = 18 for the acute electrophysiology recordings, 4 of which were from behavioral training and n = 4 for awake electrophysiology recordings). During the experimental period, the mice were housed with 2 to 4 mice per cage, enriched with a running wheel, an acrylic house and nest building material. Ethical approval was granted prior to the start of the experiments from the national authority (Centrale Commissie Dierproeven, The Hague, The Netherlands; license no. AVD2016789) as required by Dutch law, and all experiments were performed according to institutional, national, and European Union guidelines and legislation as overseen by the Animal Welfare Board of the Erasmus MC.

Surgery

For mice used in behavioral experiments and awake electrophysiology recordings, a titanium head plate was placed on the skull of the animals for head-fixation during the experiments. The animals were anesthetised through respiratory delivery of isoflurane and maintained at surgical level of anesthesia, assessed through the hind limb withdrawal reflex. A heating pad with rectal feedback probe (40-90-8C; FHC, Bowdoinham, ME, USA) was used to maintain body core temperature at 36-37 °C. Eye ointment (Duratears®; Alcon Nederland, Gorinchem, The Netherlands) was used to keep the eyes moist during surgery. Boluses of buprenorphine (0.05 mg/kg; Temgesic®, Merck Sharp & Dohme, Inc., Kenilworth, NJ, USA) and carprofen (5 mg/kg; Rimadyl®, Zoetis, Capelle a/d IJssel, The Netherlands) were injected subcutaneously at the beginning of surgery. The skin overlying the IC was incised. Lidocaine (Xylocaine® 10%; AstraZeneca, Zoetermeer, The Netherlands) was applied before removing the periosteum and cleaning the skull. After etching the bone surface with phosphoric acid gel (Etch Rite™; Pulpdent Corporation, Watertown, MA, USA), the titanium head plate was glued to the dorsal surface of the skull using dental adhesive (OptiBond™ FL; Kerr Italia S.r.l., Scafati, SA, Italy) and further secured with dental composite (Charisma®; Heraeus Kulzer GmbH, Hanau, Germany).

For acute anesthetized experiments, prior to headplate attachment, anesthesia was induced with ketamine (10%, Alfasan, Woerden, The Netherlands) and xylazine (Sedazine®, ASTfarma, Oudewater, The Netherlands) mixture (130 $mg.kg^{-1}$ and 13 $mg.kg^{-1}$, respectively) by IP injection. An intraperitoneal cannula (Butterfly Winged Infusion Set, 25G x ¾”, NIPRO EUROPE N.V., Zaventem, Belgium) was placed to allow for a maintenance solution of ketamine/xylazine to be injected. This was done with a syringe pump (Fusion 200, Chemyx, Stafford, TX, USA) providing a continuous rate infusion of between 35 - 45 mg/kg/h of ketamine and 1.4 – 1.8 mg/kg/h of xylazine. Adequacy of anesthetic depth was evaluated by toe-pinch, and the infusion rate was adjusted when needed. Headplate attachment was performed as described above. A craniotomy was performed overlying the left IC for recording. In acute experiments a second craniotomy was performed overlying the frontal lobe for the ground screw. For mice used in repeated awake recordings, a silver wire was tucked underneath the skull through the craniotomy for grounding. To protect the craniotomy and minimize shift of the brain, the craniotomy of the acute mouse was covered with 2% agarose dissolved in rat Ringer's solution (148 mM NaCl, 5.4 mM KCl, 1 mM $MgCl_2$, 1.8 mM $CaCl_2$, 5 mM HEPES), while those of the awake recordings were covered in Dura-Gel (Cambridge Neurotech Ltd., Cambridge, United Kingdom).

Amplitude-modulated sound stimulus

Sound stimuli were generated in MATLAB R2019a (The MathWorks, Natick, MA, USA) and played back via a TDT System3 setup (a combination of RX6 processor, PA5 attenuator and ED1 electrostatic speaker drivers, or a RZ6 processor) driving two EC1 electrostatic speakers (Tucker Davis Technologies, Alachua, FL, USA). Sound stimuli were presented in open field bilaterally during behavioral and awake electrophysiology experiments, and unilaterally to the right ear in the anesthetized experiments. Sound intensities were calibrated using a condenser microphone (ACO pacific Type 7017; ACO Pacific, Inc., Belmont, CA, USA) connected to a calibrated pre-amplifier and placed at the position of the pinnae.

The base stimulus consists of a broad-band pink (1/f) noise between 2 and 48 kHz. Each trial started with an unmodulated noise with a 10 ms cosine squared ramp. After a certain duration of time, the sound switched into a sinusoidally modulated version of the base stimulus according to:

$$y_{mod} = y \times (1 + m \times sin\ (2\pi f_{mod} t + \varphi)\ )$$

where y and y_{mod} are the base and amplitude-modulated (AM) stimuli, respectively; m is the modulation depth between 0 (no modulation) and 1 (full modulation); f_{mod} is the modulation frequency in Hz; t is time in s; and φ is an offset in phase to accommodate differences in instantaneous intensity (see below). In the behavioral test the animals were trained to detect the transition into amplitude modulation.

Several measures were taken to prevent the animal from detecting possible non-AM cues. First of all, amplitude modulation creates sidebands and increases the overall intensity of the sound. The unmodulated part of each stimulus was generated by first creating another instance of the amplitude-modulated noise of the trial, followed by randomization of the phase of the sideband in the spectral domain via Fourier transform. This created an unmodulated noise which was matched to the amplitude-modulated noise with respect to intensity and spectral properties. Secondly, to avoid an abrupt transition, the unmodulated and amplitude-modulated noise were cross-faded with a 20-ms window following a cosine and a sine function, respectively, between 0 and π/2. The Pythagorean identity $\sin^2\theta + \cos^2\theta = 1$ ensured that the combined instantaneous power of the two uncorrelated noises remained constant. Thirdly, the amplitude modulation started at the phase at which the instantaneous value of the sine function matched the average power of the unmodulated noise: $\varphi = arcsin\ (\sqrt{1 + 0.5\ m} - 1)$, where m is the modulation depth.

Behavioral training

The animals were motivated through water control. They received water only during habituation and training sessions during the week and ad libitum water during the weekends (Friday evening to Sunday afternoon). Each mouse received one session on Monday and two sessions per day for the rest of the week. Before the first training session, mice were habituated for two weeks to head-fixation in the setup with gradually increasing durations. During training sessions, animals were head-fixed and placed in a sound isolated booth between two speakers and in front of a metal lick port for dispensing water rewards. Licking behavior was measured by means of a change in capacitance of the lick port connected to an Arduino Nano utilizing the CapacitiveSensor library (playground.arduino.cc/Main/CapacitiveSensor/). Mice were trained to suppress their licking behavior after the onset of unmodulated (UM) noise until the transition to amplitude modulated noise (16 Hz, full modulation depth, 60 dB SPL), upon which licking was rewarded with a water drop. A timeout (delaying the start of the next trial) was used after licking during the UM period to enforce this paradigm. The

animal was subsequently trained to respond to other modulation frequencies, depths and sound intensities.

The animals were tested with the method of constant stimuli. Modulation frequencies f_{mod} ranged from 4-1024 Hz in octave steps; modulation depths were 0.06, 0.125, 0.25, 0.5 and 1. Additionally, unrewarded catch trials with a transition into unmodulated noise were introduced to gauge the false alarm rate. Any lick response within 750 ms after the onset of the transition was regarded as a hit (normal trials) or false alarm (catch trials). The sensitivity measure *d'* was calculated from AUC from ROC analysis using the formula $d' = |\sqrt{2}\,.\,\Phi^{-1}(AUC)|$where Φ^{-1} represents the MATLAB function (norminv) for the inverse cumulative distribution. To avoid infinities in the d' conversion, the behavioral measures (hit rates and false-alarm rates) were limited to $0.5/N_{trial}$ and $1-0.5/N_{trial}$, respectively, and neurometric measures' AUC values were limited to between 0.0028 and 0.9972.

Auditory Brainstem Response (ABR)

The ABR was used to obtain hearing level thresholds of behavioral animals prior to head plate attachment surgery. Mice were anesthetized with a mixture of ketamine/xylazine (100/10 mg/kg) i.p. and placed 3.5 cm in front of an electrostatic speaker in a sound-attenuated booth. Bilateral needle electrodes were placed subdermally overlying the bulla, while the active electrode was placed at the vertex and a ground electrode placed approximately over the sacrum. Tone pips of 1 ms were presented at 4, 8, 16 and 32 kHz at intensities ranging from 0 to 70 dB sound pressure level (SPL; re 20 µPa) at 5-dB resolution. Tone pips were generated by an RZ6 multi-I/O processor, while recordings were collected on a preamplifier (Medusa4Z, Tucker Davis Technologies, Alachua, FL, USA), all controlled by BioSigRZ (Tucker Davis Technologies, Alachua, FL, USA) software. At each sound intensity tested, a minimum of 500 brainstem evoked responses with artifacts below 30 µV were averaged. The minimum threshold was defined as the lowest intensity at which a reproducible waveform could be identified. After the recording was completed, atipamezole (0.5 mg/kg) (Antisedan®, Zoetis, Capelle a/d IJssel, The Netherlands) was injected i.p. to aid in the recovery from anesthesia.

Electrophysiological recordings

Silicon probe recordings of neuronal responses in the mouse IC to matching AM stimuli were performed both under anesthesia and in awake animals.

A multichannel silicon electrode (64 channel, single linear shank, 1250 µm, H3, Cambridge Neurotech Ltd., Cambridge, United Kingdom) was coated in DiO (3,3'-Dioctadecyloxacarbocyanine perchlorate, catalog number: D275, Invitrogen, ThermoFisher Scientific, Waltham, MA, USA) or DiI (1,1'-dioctadecyl-3,3,3',3'-tetramethylindocarbocyanine perchlorate; catalog number: D282, Invitrogen) saturated in 97% ethanol, and inserted under direct visual inspection during online monitoring of the channels for sound-evoked spikes. The electrode was connected to an acquisition board (RHD USB-interface board, part C3100, Intan Technologies, Los Angeles, CA, United States), after A/D conversion at the headstage amplifier (RHD 64-channel headstage, part C3315, Intan Technologies) and recorded at a sampling frequency of 30 kHz using the supplied recording GUI (RHD data acquisition GUI, Intan Technologies).

Pure-tone responses were recorded in FRA sets (frequency: 2-64 kHz in ¼ octave steps; intensity: 0-60 dB SPL in 5 dB steps; duration: 100 ms; intertrial interval: 200 ms; repetitions: 20, presented in pseudorandom order).

AM recording sets contained amplitude modulated noise stimuli as presented in the behavioral paradigm, except they had a fixed UM duration of 1 s before transitioning into AM, which then played for another 1 s. The intertrial interval (ITI) was set to a minimum of 500 ms. AM parameters were: modulation frequency (16, 64, 128 & 512 Hz), modulation depth (0, 0.06, 0.125, 0.25, 0.5 & 1) and intensity (45 & 60 dB SPL). Trials were presented in pseudorandomized order.

After completion of the recordings, transcardial perfusion under pentobarbital anesthesia was performed with 4% paraformaldehyde (PFA) in 0.12 M phosphate buffer. The brain was dissected out and postfixed in 4% PFA for 30 minutes. After storage in 10% sucrose, the brain was embedded in gelatin, sectioned on a freezing microtome and stained for DAPI. Images of slices (Axio Imager M2, Zeiss, Oberkochen, Germany) were used to reconstruct the location of the probe path in the IC, and aligned to the Allen Brain Atlas Common Coordinate Framework v3 (CCFv3) using Matlab code adapted from the AllenCCF (Shamash et al. 2018). Using the coordinates of the tract and the location of single units along the linear shank of the electrode, the location of individual extracted units within the brain was determined.

Single-unit activity was obtained with spike sorting using Kilosort2 (Pachitariu et al. 2016) and manual curation of the sorted results using Phy2 (https://github.com/cortex-lab/phy).

Pure-tone response analysis

To map the tuning of single units to pure tones the frequency autocorrelation areas (FACA) (Geis et al. 2011) were calculated from the kernel spike density (KSD) of the evoked spike rates. From this tuning curve, the characteristic frequency (CF, the frequency with the lowest intensity to produce a significant response) and the minimum threshold (the lowest intensity to evoke a significant response) were calculated.

AM analysis

Temporal coding was quantified by a measure of neural synchronization to the envelope of AM, phase-projected vector strength (VS_{PP}), as described in Yin et al. (2011). In contrast to traditional vector strength, VS_{PP} tolerates lower firing rates better. Both VS and average firing rate were calculated from the spikes in the last 500 ms of the AM stimulus, in order to eliminate the onset effect of AM.

One challenge in comparing inter-unit performance using the firing rate was the large variations in firing rates between units. Additionally, in order to compare the performance of AM encoding between the temporal and rate modalities, they would need to be calculated on the same scale. To address these challenges, ROC analysis was performed to convert the calculated parameters to a performance score in d' scale. By assigning a cutoff of 1 to the d' curve, a threshold was determined for each modulation frequency through interpolation.

Principal component analysis of response differences

We converted the spike times of all neurons into peristimulus time histograms (PSTHs) with a 3.9 ms bin size ($PSTH_{raw}$; Figure 9A-B) and averaged across the 20 repetitions of each stimulus. We then calculated a response difference between the average response to full modulation stimulus (m = 1) across all modulation frequencies and that to the unmodulated control stimulus ($PSTH_{diff}$; Figure 9C). This difference was combined across neurons into a T x N matrix where N = 112 neurons and T = 4801 time points (0.5 ms sampling between -0.2 to 2.2 seconds relative to stimulus onset) which was used as the input for the principal component analysis (PCA). PCA was performed using the Matlab `pca` function

with the data and the default singular value decomposition algorithm. The PCA found the principal components that best explains the difference between fully modulated and unmodulated responses. This step emulated how an animal learned to distinguish between modulated sound in general from unmodulated noise, without regard to the specific modulation frequency. For evaluation of different parameters, we systematically evaluated the effect of temporal resolution (bin size), PSTH normalization, mean subtraction and inclusion of responses to different modulation depths. Temporal resolution: bin sizes were varied between 0.98 ms (1/1024 Hz) to 500 ms (1/2 Hz). PSTH normalization: SD of each unit's PSTH was normalized to 1. Inclusion of modulation depths: the difference in response is calculated either between average response to all fully-modulated stimuli (m = 1) and averaged response to unmodulated stimuli; or between average response across all modulated stimuli (m = 0.06 to 1; "all mod") and averaged response to unmodulated stimuli. Mean subtraction: the mean of the difference in response ($PSTH_{diff}$) for each unit was subtracted before input to PCA.

We simulated how an animal would make use of this knowledge to perform our behavioral task. Single trial spike times for each neuron were converted into a kernel spike density (KSD) function using a 2 ms Gaussian kernel and sampled at 0.5 ms. The combined response of all neurons was then projected into PC1. Assuming the animals were detecting a change from the unmodulated noise, the average PC1 value in a 500-ms window preceding the AM transition was taken as baseline and was subtracted from the trace. We then checked the time at which the baseline-subtracted PC1 crosses a certain "detection level" value and added an extra 100-200 ms non-decision time to get a simulated response latency, limiting to a maximum of 1 s (stimulus duration). We systematically tested different detection level values (-200 to 200 in 2 unit steps; range of actual PC1 values: [-239.8, 444.6]) and non-decision times (100 to 200 ms in 10 ms steps) and compared them to response latencies collected from behavioral experiments. For each stimulus, empirical cumulative distribution functions of simulated response latencies were calculated from 0 to 1 s in 20 ms bins, and compared to that of the behavioral data using their root-mean-square difference.

Mathematically, this is equivalent to testing when population activity has passed a certain hyperplane in neuronal space, where its orientation was orthogonal to the unit vector of the PC, the origin is the averaged response to unmodulated carrier preceding the transition, and its distance to origin equal to the detection level.

Acknowledgements

We thank Marcel van der Heijden for the concepts of the new AM stimulus; Peter Bremen for assistance in pilot multielectrode recording experiments; Jean Slenter for performing part of the mouse training; Jean Slenter, Sander Kruithof and Erica Sabel-Goedknegt for support in histology.

References

Azeredo da Silveira R and Rieke F (2021). The geometry of information coding in correlated neural populations. *Annu Rev Neurosci* **44**: 403-424 https://doi.org/10.1146/annurev-neuro-120320-082744.

Barlow HB (1972). Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* **1**: 371-394 https://doi.org/10.1068/p010371.

Bartolo R, Saunders RC, Mitz AR and Averbeck BB (2020). Information-limiting correlations in large neural populations. *J Neurosci* **40**: 1668-1678 https://doi.org/10.1523/JNEUROSCI.2072-19.2019.

Baumann S, Griffiths TD, Sun L, Petkov CI, Thiele A and Rees A (2011). Orthogonal representation of sound dimensions in the primate midbrain. *Nat Neurosci* **14**: 423-425 https://doi.org/10.1038/nn.2771.

Bendor D and Wang X (2007). Differential neural coding of acoustic flutter within primate auditory cortex. *Nat Neurosci* **10**: 763-771 https://doi.org/10.1038/nn1888.

Bremen P, Massoudi R, Van Wanrooij MM and Van Opstal AJ (2017). Audio-visual integration in a redundant target paradigm: a comparison between rhesus macaque and man. *Front Syst Neurosci* **11**: 89 https://doi.org/10.3389/fnsys.2017.00089.

Cai H and Dent ML (2020). Best sensitivity of temporal modulation transfer functions in laboratory mice matches the amplitude modulation embedded in vocalizations. *J Acoust Soc Am* **147**: 337 https://doi.org/10.1121/10.0000583.

Cai R, Richardson BD and Caspary DM (2016). Responses to predictable versus random temporally complex stimuli from single units in auditory thalamus: impact of aging and anesthesia. *J Neurosci* **36**: 10696-10706 https://doi.org/10.1523/JNEUROSCI.1454-16.2016.

Carney LH, Ketterer AD, Abrams KS, Schwarz DM and Idrobo F (2013). Detection thresholds for amplitude modulations of tones in budgerigar, rabbit, and human. *Adv Exp Med Biol* **787**: 391-398 https://doi.org/10.1007/978-1-4614-1590-9_43.

Casseday JH, Ehrlich D and Covey E (1994). Neural tuning for sound duration: role of inhibitory mechanisms in the inferior colliculus. *Science* **264**: 847-850 https://doi.org/10.1126/science.8171341.

de Lafuente V and Romo R (2006). Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proc Natl Acad Sci U S A* **103**: 14266-14271 https://doi.org/10.1073/pnas.0605826103.

Dent ML, Klump GM and Schwenzfeier C (2002). Temporal modulation transfer functions in the barn owl (*Tyto alba*). *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* **187**: 937-943 https://doi.org/10.1007/s00359-001-0259-5.

Ding N, Patel AD, Chen L, Butler H, Luo C and Poeppel D (2017). Temporal modulations in speech and music. *Neurosci Biobehav Rev* **81**: 181-187 https://doi.org/10.1016/j.neubiorev.2017.02.011.

Dong C, Qin L, Liu Y, Zhang X and Sato Y (2011). Neural responses in the primary auditory cortex of freely behaving cats while discriminating fast and slow click-trains. *PLoS One* **6**: e25895 https://doi.org/10.1371/journal.pone.0025895.

Dong C, Qin L, Zhao Z, Zhong R and Sato Y (2013). Behavioral modulation of neural encoding of click-trains in the primary and nonprimary auditory cortex of cats. *J Neurosci* **33**: 13126-13137 https://doi.org/10.1523/JNEUROSCI.1724-13.2013.

Dooling RJ and Searcy MH (1981). Amplitude-modulation thresholds for the parakeet (*Melopsittacus undulatus*). *Journal of Comparative Physiology* **143**: 383-388 https://doi.org/10.1007/Bf00611177.

Ehret G and Schreiner C (2005). Spectral and intensity coding in the auditory midbrain. *The inferior colliculus*. J. Winer and C. Schreiner. New York, Springer**:** 312 - 345 https://doi.org/10.1007/0-387-27083-3_11.

Foffani G and Moxon KA (2004). PSTH-based classification of sensory stimuli using ensembles of single neurons. *Journal of Neuroscience Methods* **135**: 107-120 http://www.sciencedirect.com/science/article/pii/S016502700300428X.

Geis H-R and Borst JGG (2009). Intracellular responses of neurons in the mouse inferior colliculus to sinusoidal amplitude-modulated tones. *J Neurophysiol* **101**: 2002-2016 https://doi.org/10.1152/jn.90966.2008.

Geis HRAP, van der Heijden M and Borst JGG (2011). Subcortical input heterogeneity in the mouse inferior colliculus. *J Physiol* **589**: 3955-3967 https://doi.org/10.1113/jphysiol.2011.210278.

Goris RLT, Ziemba CM, Stine GM, Simoncelli EP and Movshon JA (2017). Dissociation of choice formation and choice-correlated activity in macaque visual cortex. *J Neurosci* **37**: 5195-5203 https://doi.org/10.1523/JNEUROSCI.3331-16.2017.

Guo ZV, Li N, Huber D, Ophir E, Gutnisky D, Ting JT, Feng G and Svoboda K (2014). Flow of cortical activity underlying a tactile decision in mice. *Neuron* **81**: 179-194 https://doi.org/10.1016/j.neuron.2013.10.020.

Haefner RM, Gerwinn S, Macke JH and Bethge M (2013). Inferring decoding strategies from choice probabilities in the presence of correlated variability. *Nat Neurosci* **16**: 235-242 https://doi.org/10.1038/nn.3309.

Hernández O, Espinosa N, Pérez-González D and Malmierca MS (2005). The inferior colliculus of the rat: a quantitative analysis of monaural frequency response areas. *Neuroscience* **132**: 203-217 https://doi.org/10.1016/j.neuroscience.2005.01.001.

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

Johnson JS, Yin P, O'Connor KN and Sutter ML (2012). Ability of primary auditory cortical neurons to detect amplitude modulation with rate and temporal codes: neurometric analysis. *J Neurophysiol* **107**: 3325-3341 https://doi.org/10.1152/jn.00812.2011.

Joris PX, Schreiner CE and Rees A (2004). Neural processing of amplitude-modulated sounds. *Physiol Rev* **84**: 541-577 https://doi.org/10.1152/physrev.00029.2003.

Kelly JB, Cooke JE, Gilbride PC, Mitchell C and Zhang H (2006). Behavioral limits of auditory temporal resolution in the rat: amplitude modulation and duration discrimination. *J Comp Psychol* **120**: 98-105 https://doi.org/10.1037/0735-7036.120.2.98.

Kettner RE and Thompson RF (1985). Cochlear nucleus, inferior colliculus, and medial geniculate responses during the behavioral detection of threshold-level auditory stimuli in the rabbit. *J Acoust Soc Am* **77**: 2111-2127 https://doi.org/10.1121/1.391736.

Klump GM and Okanoya K (1991). Temporal modulation transfer functions in the European starling (*Sturnus vulgaris*): I. Psychophysical modulation detection thresholds. *Hear Res* **52**: 1-11 https://doi.org/10.1016/0378-5955(91)90182-9.

Kobak D, Brendel W, Constantinidis C, Feierstein CE, Kepecs A, Mainen ZF, Qi XL, Romo R, Uchida N and Machens CK (2016). Demixed principal component analysis of neural population data. *Elife* **5** https://doi.org/10.7554/eLife.10989.

Kohlrausch A, Fassel R and Dau T (2000). The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. *J Acoust Soc Am* **108**: 723-734 https://doi.org/10.1121/1.429605.

Krebs B, Lesica NA and Grothe B (2008). The representation of amplitude modulations in the mammalian auditory midbrain. *J Neurophysiol* **100**: 1602-1609 https://doi.org/10.1152/jn.90374.2008.

Leary CJ, Edwards CJ and Rose GJ (2008). Midbrain auditory neurons integrate excitation and inhibition to generate duration selectivity: an *in vivo* whole-cell patch study in anurans. *J Neurosci* **28**: 5481-5493 https://doi.org/10.1523/jneurosci.5041-07.2008.

Lee J and Bacon SP (1997). Amplitude modulation depth discrimination of a sinusoidal carrier: effect of stimulus duration. *J Acoust Soc Am* **101**: 3688-3693 https://doi.org/10.1121/1.418329.

Lu T and Wang X (2004). Information content of auditory cortical responses to time-varying acoustic stimuli. *J Neurophysiol* **91**: 301-313 https://doi.org/10.1152/jn.00022.2003.

Malone BJ, Scott BH and Semple MN (2010). Temporal codes for amplitude contrast in auditory cortex. *J Neurosci* **30**: 767-784 https://doi.org/10.1523/JNEUROSCI.4170-09.2010.

Marshel JH, Kim YS, Machado TA, Quirin S, Benson B, Kadmon J, Raja C, Chibukhchyan A, Ramakrishnan C, Inoue M, Shane JC, McKnight DJ, Yoshizawa S, Kato HE, Ganguli S and Deisseroth K (2019). Cortical layer-specific critical dynamics triggering perception. *Science* **365** https://doi.org/10.1126/science.aaw5202.

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

Moody DB (1994). Detection and discrimination of amplitude-modulated signals by macaque monkeys. *J Acoust Soc Am* **95**: 3499-3510 https://doi.org/10.1121/1.409967.

Moon KR, van Dijk D, Wang Z, Gigante S, Burkhardt DB, Chen WS, Yim K, Elzen AVD, Hirn MJ, Coifman RR, Ivanova NB, Wolf G and Krishnaswamy S (2019). Visualizing structure and transitions in high-dimensional biological data. *Nat Biotechnol* **37**: 1482-1492 https://www.ncbi.nlm.nih.gov/pubmed/31796933 PHATE.

Müller MK, Jovanovic S, Keine C, Radulovic T, Rübsamen R and Milenkovic I (2019). Functional development of principal neurons in the anteroventral cochlear nucleus extends beyond hearing onset. *Front Cell Neurosci* **13**: 119 https://doi.org/10.3389/fncel.2019.00119.

Newsome WT, Britten KH and Movshon JA (1989). Neuronal correlates of a perceptual decision. *Nature* **341**: 52-54 https://doi.org/10.1038/341052a0.

Niwa M, Johnson JS, O'Connor KN and Sutter ML (2012a). Active engagement improves primary auditory cortical neurons' ability to discriminate temporal modulation. *J Neurosci* **32**: 9323-9334 https://doi.org/10.1523/JNEUROSCI.5832-11.2012.

Niwa M, Johnson JS, O'Connor KN and Sutter ML (2012b). Activity related to perceptual judgment and action in primary auditory cortex. *J Neurosci* **32**: 3193-3210 https://doi.org/10.1523/JNEUROSCI.0767-11.2012.

Niwa M, O'Connor KN, Engall E, Johnson JS and Sutter ML (2015). Hierarchical effects of task engagement on amplitude modulation encoding in auditory cortex. *J Neurophysiol* **113**: 307-327 https://doi.org/10.1152/jn.00458.2013.

O'Connor KN, Johnson JS, Niwa M, Noriega NC, Marshall EA and Sutter ML (2011). Amplitude modulation detection as a function of modulation frequency and stimulus duration: comparisons between macaques and humans. *Hear Res* **277**: 37-43 https://doi.org/10.1016/j.heares.2011.03.014.

Ono M, Bishop DC and Oliver DL (2017). Identified GABAergic and glutamatergic neurons in the mouse inferior colliculus share similar response properties. *J Neurosci* **37**: 8952-8964 https://doi.org/10.1523/JNEUROSCI.0745-17.2017

Pachitariu M, Steinmetz N, Kadir S and Carandini M (2016). Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. *BioRxiv*: 061481 https://doi.org/10.1101/061481.

Palmer AR, Shackleton TM, Sumner CJ, Zobay O and Rees A (2013). Classification of frequency response areas in the inferior colliculus reveals continua not discrete classes. *J Physiol* **591**: 4003-4025 https://doi.org/10.1113/jphysiol.2013.255943.

Pauli-Magnus D, Hoch G, Strenzke N, Anderson S, Jentsch TJ and Moser T (2007). Detection and differentiation of sensorineural hearing loss in mice using auditory steady-state responses and transient auditory brainstem responses. *Neuroscience* **149**: 673-684 https://doi.org/10.1016/j.neuroscience.2007.08.010.

Portfors CV, Mayko ZM, Jonson K, Cha GF and Roberts PD (2011). Spatial organization of receptive fields in the auditory midbrain of awake mouse. *Neuroscience* **193**: 429-439 https://doi.org/10.1016/j.neuroscience.2011.07.025.

Ratcliff R, Smith PL, Brown SD and McKoon G (2016). Diffusion decision model: current issues and history. *Trends Cogn Sci* **20**: 260-281 https://doi.org/10.1016/j.tics.2016.01.007.

Rees A and Langner G (2005). Temporal coding in the auditory midbrain. *The inferior colliculus*. J. Winer and C. Schreiner. New York, Springer**:** 346 - 376 https://doi.org/10.1007/0-387-27083-3_12.

Rees A and Møller AR (1987). Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. *Hear Res* **27**: 129-143 https://doi.org/10.1016/0378-5955(87)90014-1.

Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, Reyes A and Harris KD (2010). The asynchronous state in cortical circuits. *Science* **327**: 587-590 https://doi.org/10.1126/science.1179850.

Rodríguez FA, Read HL and Escabí MA (2010). Spectral and temporal modulation tradeoff in the inferior colliculus. *J Neurophysiol* **103**: 887-903 https://doi.org/10.1152/jn.00813.2009.

Rosen MJ, Semple MN and Sanes DH (2010). Exploiting development to evaluate auditory encoding of amplitude modulation. *J Neurosci* **30**: 15509-15520 https://doi.org/10.1523/JNEUROSCI.3340-10.2010.

Rosenblatt F (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* **65**: 386-408.

Salvi RJ, Giraudi DM, Henderson D and Hamernik RP (1982). Detection of sinusoidally amplitude modulated noise by the chinchilla. *J Acoust Soc Am* **71**: 424-429 https://doi.org/10.1121/1.387445.

Schnupp JWH, Garcia-Lazaro JA and Lesica NA (2015). Periodotopy in the gerbil inferior colliculus: local clustering rather than a gradient map. *Front Neural Circuits* **9**: 37 https://doi.org/10.3389/fncir.2015.00037.

Schnupp JWH, Hall TM, Kokelaar RF and Ahmed B (2006). Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J Neurosci* **26**: 4785-4795 https://doi.org/10.1523/JNEUROSCI.4330-05.2006.

Schreiner CE and Langner G (1988). Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. *J Neurophysiol* **60**: 1823-1840 https://doi.org/10.1152/jn.1988.60.6.1823.

Shamash P, Carandini M, Harris K and Steinmetz N (2018). A tool for analyzing electrode tracks from slice histology. *bioRxiv*: 447995 https://doi.org/10.1101/447995.

Sheft S and Yost WA (1990). Temporal integration in amplitude modulation detection. *J Acoust Soc Am* **88**: 796-805 https://doi.org/10.1121/1.399729.

Stiebler I and Ehret G (1985). Inferior colliculus of the house mouse. I. A quantitative study of tonotopic organization, frequency representation, and tone-threshold distribution. *J Comp Neurol* **238**: 65-76 https://doi.org/10.1002/cne.902380106.

Stringer C, Michaelos M, Tsyboulski D, Lindo SE and Pachitariu M (2021). High-precision coding in visual cortex. *Cell* **184**: 2767-2778 e2715 https://doi.org/10.1016/j.cell.2021.03.042.

Stüttgen MC, Schwarz C and Jäkel F (2011). Mapping spikes to sensations. *Front Neurosci* **5**: 125 https://doi.org/10.3389/fnins.2011.00125.

Taberner AM and Liberman MC (2005). Response properties of single auditory nerve fibers in the mouse. *J Neurophysiol* **93**: 557-569 https://doi.org/10.1152/jn.00574.2004.

1010 Tanke N, Borst JGG and Houweling AR (2018). Single-cell stimulation in barrel cortex influences
1011 psychophysical detection performance. *J Neurosci* **38**: 2057-2068
1012 https://doi.org/10.1523/JNEUROSCI.2155-17.2018.
1013 Ter-Mikaelian M, Sanes DH and Semple MN (2007). Transformation of temporal properties between
1014 auditory midbrain and cortex in the awake Mongolian gerbil. *J Neurosci* **27**: 6091-6102
1015 https://doi.org/10.1523/JNEUROSCI.4848-06.2007.
1016 van der Maaten L and Hinton G (2008). Visualizing Data using t-SNE. *Journal of Machine Learning*
1017 *Research* **9**: 2579-2605 <Go to ISI>://WOS:000262637600007 t-SNE.
1018 van Looij MAJ, Liem SS, van der Burg H, van der Wees J, De Zeeuw CI and van Zanten BGA (2004). Impact
1019 of conventional anesthesia on auditory brainstem responses in mice. *Hear Res* **193**: 75-82
1020 https://doi.org/10.1016/j.heares.2004.02.009.
1021 Viemeister NF (1979). Temporal modulation transfer functions based upon modulation thresholds. *J*
1022 *Acoust Soc Am* **66**: 1364-1380 https://doi.org/10.1121/1.383531.
1023 von Trapp G, Buran BN, Sen K, Semple MN and Sanes DH (2016). A decline in response variability
1024 improves neural signal detection during auditory task performance. *J Neurosci* **36**: 11097-11106
1025 https://doi.org/10.1523/JNEUROSCI.1302-16.2016
1026 Walker KMM, Ahmed B and Schnupp JWH (2008). Linking cortical spike pattern codes to auditory
1027 perception. *J Cogn Neurosci* **20**: 135-152 https://doi.org/10.1162/jocn.2008.20012.
1028 Walton JP, Simon H and Frisina RD (2002). Age-related alterations in the neural coding of envelope
1029 periodicities. *J Neurophysiol* **88**: 565-578 https://doi.org/10.1152/jn.2002.88.2.565.
1030 Wang L, Narayan R, Graña G, Shamir M and Sen K (2007). Cortical discrimination of complex natural
1031 stimuli: can single neurons match behavior? *J Neurosci* **27**: 582-589
1032 https://doi.org/10.1523/JNEUROSCI.3699-06.2007.
1033 Wang Q, Ding SL, Li Y, Royall J, Feng D, Lesnar P, Graddis N, Naeemi M, Facer B, Ho A, Dolbeare T,
1034 Blanchard B, Dee N, Wakeman W, Hirokawa KE, Szafer A, Sunkin SM, Oh SW, Bernard A, Phillips
1035 JW, Hawrylycz M, Koch C, Zeng H, Harris JA and Ng L (2020). The Allen Mouse Brain Common
1036 Coordinate Framework: A 3D reference atlas. *Cell* **181**: 936-953 e920
1037 https://doi.org/10.1016/j.cell.2020.04.007.
1038 Wang Y, Abrams KS, Carney LH and Henry KS (2021). Midbrain-level neural correlates of behavioral tone-
1039 in-noise detection: dependence on energy and envelope cues. *J Neurosci* **41**: 7206-7223
1040 https://doi.org/10.1523/JNEUROSCI.3103-20.2021.
1041 Yin P, Johnson JS, O'Connor KN and Sutter ML (2011). Coding of amplitude modulation in primary
1042 auditory cortex. *J Neurophysiol* **105**: 582-600 https://doi.org/10.1152/jn.00621.2010.
1043 Zhang H and Kelly JB (2003). Glutamatergic and GABAergic regulation of neural responses in inferior
1044 colliculus to amplitude-modulated sounds. *J Neurophysiol* **90**: 477-490
1045 https://doi.org/10.1152/jn.01084.2002.

1046