Title

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

Abbreviated title

High sensitivity of sound modulation by mice

Authors

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

Affiliation

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

[#] corresponding author: a.wong@erasmusmc.nl.

Contact details

Aaron B. Wong
Email: a.wong@erasmusmc.nl
Postal address: Dept. of Neuroscience, Room EE1202, P.O.Box 2040, NL-3000 CA Rotterdam, Netherlands

Declarations of interest: none.

Author contributions:

Maurits M. van den Berg: Conceptualization, Formal analysis, Investigation. Methodology, Software, Writing - original draft, review & editing.

Esmée Busscher: Investigation.

J. Gerard G. Borst: Conceptualization, Funding acquisition, Methodology, Supervision, Writing - original draft, review & editing.

Aaron B. Wong: Conceptualization, Funding acquisition, Formal analysis, Investigation, Methodology, Software, Supervision, Writing - original draft, review & editing.

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.

Keywords

amplitude modulation; signal detection theory; mouse; inferior colliculus; multielectrode; operant conditioning.

Abbreviations[1]

[1]ABR auditory brainstem response;
AM amplitude modulation / amplitude-modulated;
AUC area under curve;
d' sensitivity index;
IC inferior colliculus;
PSTH peristimulus time histogram;
PCA principal component analysis;
rBMF rate best modulation frequency;
rMTF rate modulation transfer function;
ROC receiver operator characteristics;
SAM sinusoidal amplitude modulation / sinusoidally amplitude-modulated;
tBMF temporal best modulation frequency;
tMTF temporal modulation transfer function;
UM unmodulated;
VS_{PP} phase-projected vector strength

1. 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 sensory detection tasks individual neurons may outperform the animal (Carney et al., 2014; Johnson et al., 2012; Kettner and Thompson, 1985; Malone et al., 2010; Newsome et al., 1989). If many neurons individually outperform the animal, their individual impact on behavior would be expected to be very small. However, this does not seem to be the case, as stimulation of even a single neuron may already impact sensory detection (Tanke et al., 2018). What might account for this apparent discrepancy between the ability of individual sensory neurons to encode a stimulus or impact behavior and the comparatively poor performance of animals in sensory detection tasks? A possibility is that the activity of different sensory neurons are heavily correlated, as these so-called noise correlations would limit the information these neurons can carry (Azeredo da Silveira and Rieke, 2021). However, noise correlations are often small (Bartolo et al., 2020; Renart et al., 2010), as also indicated by the observation that neuronal population responses can be two orders of magnitude more sensitive than the animal's performance, indicating that there are few or no limitations due to noise correlations within the population responses in the sensory precision that can be attained (Stringer et al., 2021). The gap between neurometric and psychometric performance might then be filled by non-sensory factors (de Lafuente and Romo, 2006; Goris et al., 2017; Guo et al., 2014; Kettner and Thompson, 1985; Stringer et al., 2021), but we argue here that sensory performance in some behavioral tasks may also be easily overestimated, for example because the behavioral task may include multiple types of stimuli, making the task more difficult.

Amplitude modulations (AM) forms 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 (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, yielding the rate modulation transfer function (rMTF). The tMTF changes noticeably between the cochlear nucleus and the inferior colliculus (IC). The tuning becomes sharper, and the ability of IC neurons to phase-lock to high modulation frequencies is reduced (Joris et al., 2004). Presumably, the convergence of excitatory and inhibitory inputs contributes to these changes (Casseday et al., 1994; Geis and Borst, 2009; Leary et al., 2008; Ono et al., 2017). For the rMTF, both the timing of inhibitory and excitatory synaptic inputs and intrinsic cellular properties related to spike generation can 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 for behavioral detection AM sounds are encoded by a rate code rather than a temporal code (Bendor and Wang, 2007; Dong et al., 2011; Lu and Wang, 2004;

Niwa et al., 2012b). However, temporal codes may represent sounds for both pure AM and natural vocalizations with large AM components more accurately (Henry et al., 2016; Schnupp et al., 2006; Walker et al., 2008; Wang et al., 2007). Even though there are many studies that have studied either the behavioral detection of AM or the electrophysiological encoding of AM (Joris et al., 2004; Rees and Langner, 2005), the question which mechanisms for encoding AM support behavioral sensitivity is therefore still far from settled.

Here, we performed multielectrode recordings in the IC of both anesthetized and awake mice, and compare SAM responses of single units with the behavioral performance of mice in a SAM detection task. We found 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 taking the relation between modulation depth and response latency into account, 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.

2. Materials and Methods

2.1. Animals

Wildtype B6CBAF1/JRj mice of both sexes were obtained from Janvier (behavior: n = 8 [3 females]; acute electrophysiology recordings: n = 18 [9 females; 4 from behavioral training]; awake electrophysiology recordings: n = 4 [2 females]). 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.

2.2. 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).

2.3. 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 (behavior) or between 2 and 64 kHz (electrophysiology). 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 (Figure S1). First of all, amplitude modulation alters the power spectrum of the noise by the introduction of a pair of sidebands to each of its frequency components. We denote the collective modulation sidebands by the term "sideband spectrum." The unmodulated part of each stimulus was generated by first creating another instance of the carrier noise (Figure S1A), calculating the sideband spectrum required to produce the desired amplitude modulation of the trial (Figure S1B-C), followed by randomization of the phase of these sidebands in the spectral domain (Figure S1D) and finally adding this randomized sideband back to the original carrier noise (Figure S1E). This procedure eliminated any spectral and intensity differences between the modulated and unmodulated parts of the noise stimulus (Figure S1F-H). 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 (Figure S1I-L). 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 (Figure S1I-J).

2.4. 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 AM transition, randomized per trial between 4-7 s at testing stage, was used after licking during the UM period to enforce this paradigm. Additionally, unrewarded catch trials with a transition into unmodulated noise were introduced to gauge the false alarm rate. The stimuli were subsequently expanded to include other modulation frequencies, depths and sound intensities. For each session, stimuli were presented in blocks, in which all possible stimuli were presented in randomized order. Double the number of catch trials were included in each block to allow a more robust estimation of the false alarm rate.

The animals were tested with the method of constant stimuli after stable performance was reached. Modulation frequencies ranged from 4-1024 Hz in octave steps; modulation depths were 0.06, 0.125, 0.25, 0.5 and 1. We reported hit rates and false alarm rates in Figure 3A using a response window within 750 ms after the onset of the AM transition. For each AM stimulus, an ROC analysis was performed to compare its response latencies distribution to that of catch trials in the corresponding intensity. The sensitivity measure *d'* was then calculated from the AUC obtained from ROC analysis using the formula $d' = |\sqrt{2}.\Phi^{-1}(AUC)|$ where Φ^{-1} represents the inverse of the standard normal cumulative distribution function (`norminv` in MATLAB). This method of calculating *d'* was chosen over the classical formula $\Phi^{-1}(hit\ rate) - \Phi^{-1}(false\ alarm\ rate)$ [often written as z(H)-z(F)](Macmillan and Creelman, 2004) to detect differences in the temporal distribution of true and false alarm responses, which may be not be picked up by the classical method with an arbitrarily chosen response window. Furthermore, the AUC method can be extended to physiological data to facilitate comparisons. The method we employed gave more conservative estimates for our dataset than the classical method for calculating *d'* (data not shown). To avoid infinities in d' for cases with perfect separation, we applied the limit $\Phi^{-1}(1 - 1/2N_{trials}) - \Phi^{-1}(1/2N_{trials})$ (Macmillan and Creelman, 2004). For neuronal recordings ($N_{trials} = 20$) this amounted to a maximum *d'* of 3.92.

2.5. 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. All 8 animals had normal hearing thresholds (mean ± s.d. 4 kHz: 42.5 ± 4.29 dB; 8 kHz: 18.8 ± 2.31 dB; 16 kHz: 11.9 ± 5.30 dB; 32 kHz: 43.8 ± 7.44 dB).

2.6. 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). A TTL pulse from the RZ6 indicating the onset and offset of each stimulus was recorded on a digital channel of the Intan interface board for subsequent alignment of extracted spike times relative to the stimuli.

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). Single units were selected based on good signal-to-noise ratio in the waveform with the "template amplitude" easily separable from background multiunit activities, visual inspection of the consistency of spike waveforms across the recording and adherence to a refractory period of 1 ms by inspection of the auto-correlogram.

2.7. 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.

2.8. 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_{PP} and average firing rate were calculated from the spikes occurring in the time window between the transition from UM to AM and the end 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, the calculated parameters were converted to a performance score in d' scale using ROC analysis. For both VS and firing rate, this was done by comparing trial-by-trial measures of VS_{PP} or firing rate between modulated and unmodulated stimuli. AUC was calculated from the ROC curves, and subsequently converted to d' ($d'_{temporal}$ and d'_{rate}, respectively), as described in 2.4 Behavioral training. By assigning a cutoff of 1 to the d' curve, a threshold was determined for each modulation frequency through linear interpolation in dB scale.

The temporal and rate best modulation frequency (tBMF and rBMF) of a unit was determined as the modulation frequency with the lowest $d'_{temporal}$- or d'_{rate}-derived threshold, respectively. Units with identical thresholds for two or more modulation frequencies (temporal: 3/211 units) and those with no detectable threshold at any modulation frequencies (temporal: 21/211 units; rate: 10/211 units) were excluded from further analysis on BMF. Chi-squared tests were used to assess differences in BMF distributions among subregions of the IC. Where overall significance was found ($p<0.05$) –pairwise Chi-squared tests were performed between IC subregion, with the Holm-Bonferroni correction for multiple comparisons.

2.9. 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 neuronal information 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.

3. Results

[Figure 1]

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 (Bremen et al., 2017; Klump and Okanoya, 1991; Moody, 1994). A total of eight mice were trained to respond to the transition from an unmodulated to an amplitude modulated (AM) broadband noise stimulus (Figure 1A). The stimulus was designed to specifically test for the AM cue, while controlling for both spectral and intensity differences induced by the AM (see Figure S1 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, $p = 0.68$).

[Figure 2]

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 and subsequently converted to the sensitivity index d' (see Methods for rationale; Figure 2D). Higher d' values were observed at a modulation frequency of 512 Hz than at 16 Hz, indicating that this animal was

350 better at detecting the former. For each mouse and each modulation frequency, the detection threshold
351 was defined as the modulation depth at which d' crossed 1 (vertical dotted lines in Figure 2D).

352 [Figure 3]

353 Figure 3 summarizes the results of all 8 animals. As expected, response probability for all animals
354 increased with increasing modulation depth for a wide range of modulation frequencies (Figure 3A).
355 Moreover, at all modulation frequencies animals responded faster to higher modulation depths (Figure
356 3B), most likely reflecting the certainty of the animal's detection and subsequent response. The average
357 median response latency to the fully modulated stimuli at 60 dB could be well approximated by adding
358 half a period of the modulation to a fixed delay of about 150 ms (Figure 3D) suggesting that under these
359 conditions, detection approached the physical limit at which sinusoidal amplitude modulations can be
360 detected.

361 Similar to the example shown in Figure 2, we converted the response latencies of all animals into the
362 sensitivity index d' to estimate their detection threshold (Figure 3C). Behavioral detection of AM showed
363 a clear high-pass pattern with detection thresholds in the lower range (8 Hz and lower) at a modulation
364 depth of about 0.5, whereas in the higher range (256 Hz and higher) it was about 0.25 (Figure 3E). The
365 detection threshold worsened at 1024 Hz, in line with the overall low-pass characteristics at high
366 frequencies described in many species, which are likely limited by the membrane time constant of inner
367 hair cells (Palmer and Russell, 1986) which already appears as phase-locking limitations at the level of
368 the auditory nerve (Joris et al., 2004).

369 [Figure 4]

370 [Figure 5]

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

376 We recorded a total of 112 well isolated single units from the IC of 18 anaesthetized mice using
377 multielectrode silicon probes. We presented the same AM stimuli as in the behavioral task (Figure 1B,
378 S1), as well as pure tones to obtain their frequency response areas (FRAs). We calculated both the
379 phase-projected vector strength (VS_{PP}, (Yin et al., 2011) as a measure for phase-locking to the AM stimuli
380 to obtain the temporal modulation transfer function (tMTF) and the average firing rate to obtain the
381 rate modulation transfer function (rMTF).

382 The units showed large variability in their responses to both AM stimuli and tones. In Figure 4 and 5 we
383 illustrate the responses of two example units from the central nucleus of the IC (ICc). Figure 4 shows a
384 unit with excellent phase-locking to AM stimuli. Pure tone firing patterns, visualized as a dot raster,
385 showed a strong onset response across a wide range of frequencies, followed by inhibition (Figure 4B).
386 Sustained firing for the remainder of the tone duration only occurred at a narrow bandwidth centered
387 around the CF. The frequency response area (FRA) further illustrates its sharp tuning with its CF at 16
388 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 (at t = 1 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 depth. In this unit the firing rate also depended strongly on the modulation depth of the AM stimulus (Figure 4G).

Figure 5 illustrates an ICc 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,C).

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 5D, E). 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). The large diversity of FRAs is in agreement with previous results (Ehret and Schreiner, 2005). We did not attempt to classify the different FRAs, as previous clustering attempts were unsuccessful (Hernández et al., 2005; Palmer et al., 2013).

[Figure 6]

We next looked at the response of the population of IC neurons as a whole. Figure 6 shows the average VS_{PP} (Figure 6A) and average UM-corrected 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 a modulation frequency of 16 Hz could be established in 79% of the units (Figure 6C). This proportion decreased with increasing

modulation frequency, with 63% and 51% 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 6% 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 (data not shown). This variability was due to the large variability in noise-evoked firing rate as well as its modulation by AM. To account for this, the firing rates to the UM stimulus at the same intensity were subtracted from the AM firing rates in Figure 6D. 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. All modulation frequencies show an increase in the standard deviation of the AM firing rate as modulation depth increases, illustrating that both increases decreases in firing rate were observed among different units. Note that the strong inhibition of firing rate in some units resulted in a decrease in mean VS_{PP} at these frequencies (Figure 6A), since a trial with no spikes is defined to have a trial VS of 0, which is ultimately averaged across repeated trials to obtain VS_{PP}. Generally, this decrease in VS_{PP} reflects the difficulty of assessing phase-locking when firing rate becomes 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 (77% and 84% respectively), whereas for 16 and 512 Hz this percentage was somewhat lower (62% and 63%, respectively). In the cells with a measurable threshold, these thresholds assumed similar values across frequencies (Figure 6F).

[Figure 7]

As a control for potential effects of anesthesia, we also recorded a total of 99 units in awake, passively-listening animals (n = 4 mice). In awake and anesthetized animals similar results were obtained 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_{PP} 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. A comparison of thresholds between anaesthetized units, awake units and behavior is provided in Figure S2.

[Figure 8]

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) (Figure 8A). 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 parts. 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 dorsolateral to ventromedial tonotopic gradient of the IC (Portfors et al., 2011; Stiebler and Ehret, 1985), in both a sagittal (Figure 8B) and a coronal view (Figure 8C). The distribution of CFs in ICc neurons spanned the entire frequency range of presented stimuli, but peaked between 16 and 32 kHz. Neurons sampled from the ICd had a similar distribution of CFs as in the ICc, while that of the ICe clustered between 8 and 16 kHz (Figure 8D left). The CF distribution of the neurons recorded under anesthesia and awake were generally similar (Figure 8D right).

We did not observe any obvious spatial clustering or gradient in AM encoding, as measured by the d' thresholds calculated from VS_{PP} 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 (Figure 8E top) and rBMF (Figure 8E bottom). ICc, ICd and ICe generally had tBMF at one of the lower modulation frequencies. No significant difference was observed in the proportions of tBMFs among the three subregions when tested using a Chi-squared test (p=0.416). Each subregion of showed a different pattern of proportions of the rBMF, which a Chi-squared test found to be significant between each subregion (p-value: ICd vs. ICe = 0.004, ICc vs. ICe = 0.018 & ICc vs. ICd = 0.038). In contrast to some previous studies (Baumann et al., 2011; Rodríguez et al., 2010; Schnupp et al., 2015; Schreiner and Langner, 1988), we did not find evidence for a gradient in tBMF or rBMF or for local clustering within the IC.

[Figure 9]

Our comparison of behavioral and neurometric thresholds showed better agreement towards rate coding. A limitation of the analysis above is that VS_{pp} and firing rates were calculated for up to the end of the AM period, while an animal engaged in the task often responded earlier with less available information. To further explore how an animal would utilize spike information to perform our behavioral task, we performed a simulation 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 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]

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.

[Figure 11]

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]

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 (one for each repetition) 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.

4. Discussion

We compared psychometric and neurometric detection of AM in the mouse using a new stimulus, which precluded spectral broadening as a cue. AM detection by the mouse showed high-pass characteristics, with comparatively 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. A PCA-based threshold method that took latencies into account did provide an adequate fit, especially in awake animals. The comparison suggested that a few neurons with large and rapid changes in firing rate upon the transition to AM had a disproportionate effect on the overall performance.

4.1. Excluding spectral broadening as a cue for AM detection

AM of a noise carrier results in spectral broadening, which becomes substantial at the high modulation frequencies at which the mouse excelled. To ascertain that the mouse did not use these spectral cues, we compensated for the addition of sidebands, thus excluding a contribution of spectral (or intensity) cues to the measured AM detection thresholds. This method might be useful to retest the measured high-modulation-frequency sensitivities for other species, including humans.

4.2. Comparison of behavioral AM detection with earlier study in the mouse and with other species

The mice were poor at detecting AM at low modulation frequencies, but remarkably good at detecting high modulation frequencies. At 512 Hz, a comparison with other species indicated that the detection threshold of -13.5 dB placed the mice among the best performing species (Carney et al., 2013; Dent et al., 2002; Henderson et al., 1984; Henry et al., 2016; Joris et al., 2004; Kelly et al., 2006). An exact comparison is more difficult as methods differ between studies, and the possible confounding effects of spectral cues at high frequencies have not been dealt with in previous studies. 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 (Cai and Dent, 2020). The relatively low thresholds at high modulation frequencies may enable mice to detect for example the fast amplitude modulations in brief ultrasonic vocalizations (Cai and Dent, 2020).

4.3. Neural responses to AM

The responses to SAM noise generally matched earlier reports (Joris et al., 2004; Rees and Langner, 2005). In mice, responses to AM are generally heterogeneous (Geis and Borst, 2009; Ono et al., 2017; Walton et al., 2002). The absence of phase-locking at 512 Hz in most cells, which was in agreement with previous results using whole-cell recordings (Geis and Borst, 2009), indicates that there already must be some degrading of temporal information at the level of the IC, since in mice, phase-locking to modulation frequencies >500 Hz is still prominent at the level of the lower brainstem (Müller et al., 2019; Pauli-Magnus et al., 2007). In other species, phase-locking to frequencies up to 1200 Hz has been observed in the IC (Rees and Langner, 2005).

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

4.4. Effect of sound intensity

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

4.5. Effect of anesthesia

Interestingly, the neuronal data obtained in awake animals approached the behavioral high-pass characteristics (Figure 7) better than the data obtained under anesthesia. This was also the case for the d' of the VS or of the firing rate, suggesting that anesthesia may have had a depressing effect on neuronal responses. In previous studies, anesthesia had variable impact on the auditory system. Ketamine-xylazine may increase minimum threshold and interpeak latencies of ABRs in mice (van Looij et al., 2004). Only a limited number of studies addressed the impact on responses to SAM stimuli. In the rat medial geniculate body, urethane anesthesia did not significantly affect firing rate or VS (Cai et al., 2016). Only mild effects of pentobarbital plus ketamine anesthesia were observed on responses to AM stimuli in the gerbil IC (Ter-Mikaelian et al., 2007).

Relatively often we observed a strong decrease in firing rate at full modulation depth at the intermediate frequencies (64 and 128 Hz) after an initial increase with increasing modulation depth, similar to recordings in other rodents (Krebs et al., 2008; Zhang and Kelly, 2003). In the rMTF, which is typically based on stimuli at full modulation depths, these would be band-reject neurons, whereas at low modulation depth these neurons showed a more band-pass-like character. In recordings from awake animals the decrease in firing at large modulation depths that was observed in many units seemed less conspicuous than in anesthetized animals. As this decrease in firing rate also may have contributed to the decrease in phase-locking, this decrease in non-linearity is likely to have contributed to the better match of the unanesthetized recordings with the behavioral data. Since well-timed, onset inhibition is likely to contribute to this nonmonotonicity (Geis and Borst, 2009), the effect of anesthesia may be to enhance this inhibitory component, but intracellular recordings would be needed to test this.

4.6. Neurometric vs. psychometric detection of AM

A main goal of our study was to try to understand which aspects of the neuronal responses within the IC informed the behavioral detection of AM. A coding scheme based on phase-locking already failed with the low VS at 512 Hz, as this frequency was behaviorally well detected. PSTH-based classification and many PCA-based methods include information that is not available to the animal when it responds early (Dong et al., 2013; Foffani and Moxon, 2004; Malone et al., 2010).

We therefore opted for a two-step process in which spike responses of the whole stimulus were available to "train" the model to detect AM (PCA-step), whereas a threshold crossing enabled the model to respond at individual trials as soon as enough information was available. Since the animals were trained in a detection task, not to discriminate between different modulation frequencies, we performed PCA on the difference between the responses to the averaged AM and unmodulated stimuli, analogous to demixed PCA (Kobak et al., 2016). The PCA was performed on the neuronal dimension with

time points as different “samples”. The first principal component (PC1) already explained >40% of the variance, allowing us to use only this PC1 to predict the behavioral responses from the IC population data. We used a linear dimensionality reduction method instead of newer, non-linear methods (McInnes et al., 2018; Moon et al., 2019; van der Maaten and Hinton, 2008) for better interpretability.

We obtained simulated response delays with a simple level-crossing model based on the size of the population response along the first PCA dimension. Even though only two parameters needed to be estimated, our model overall provided a decent fit of the behavioral data, capturing the high-pass aspect of the behavioral detection. At 512 Hz, the large changes in firing rates of the top units ensured good performance. At 16 Hz, the predicted performance was much lower, despite decent phase-locking at these low frequencies. At low modulation frequencies, the slow amplitude changes lead to more dispersed spikes that sum less effectively across units, providing a possible explanation for the relatively poor behavioral detection.

Another advantage of dimensionality reduction is that we could rank neurons by their contribution to the principal component to compare their properties. The top contributors generally had low spontaneous firing rates, strong and generally well-timed responses to AM, but showed a diverse (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. Simulated false positive responses were driven entirely by random fluctuations of the principal component, i.e. synchronous activity of 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. Since we did not use frozen noise across different recordings, the lower firing rates at 45 dB will decrease crossings during catch trials, resulting in a “better” detection threshold. The lack of noise correlations across awake recording sessions may further lower false alarm rate at 45 dB. Recalculating the simulated threshold at 45 dB using the catch trials at 60 dB resulted in a better fit. An explicit false alarm rate independent of neuronal activity may therefore 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 (Dong et al., 2013; Niwa et al., 2012a; 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 (Haefner et al., 2013; Jazayeri and Movshon, 2006). 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).

Acknowledgements

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

Funding: This work was supported by the EU Marie Skłodowska-Curie Innovative Training Network LISTEN (H2020-MSCA-ITN #722098) and a ZonMW TOP grant (#91218033).

References

Azeredo da Silveira, R., 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, H.B., 1972. Single units and sensation: a neuron doctrine for perceptual psychology? Perception 1(4), 371-394. https://doi.org/10.1068/p010371.

Bartolo, R., Saunders, R.C., Mitz, A.R., Averbeck, B.B., 2020. Information-limiting correlations in large neural populations. J Neurosci 40(8), 1668-1678. https://doi.org/10.1523/JNEUROSCI.2072-19.2019.

Baumann, S., Griffiths, T.D., Sun, L., Petkov, C.I., Thiele, A., Rees, A., 2011. Orthogonal representation of sound dimensions in the primate midbrain. Nat Neurosci 14(4), 423-425. https://doi.org/10.1038/nn.2771.

Bendor, D., Wang, X., 2007. Differential neural coding of acoustic flutter within primate auditory cortex. Nat Neurosci 10(6), 763-771. https://doi.org/10.1038/nn1888.

Bremen, P., Massoudi, R., Van Wanrooij, M.M., Van Opstal, A.J., 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., Dent, M.L., 2020. Best sensitivity of temporal modulation transfer functions in laboratory mice matches the amplitude modulation embedded in vocalizations. J Acoust Soc Am 147(1), 337. https://doi.org/10.1121/10.0000583.

Cai, R., Richardson, B.D., Caspary, D.M., 2016. Responses to predictable versus random temporally complex stimuli from single units in auditory thalamus: impact of aging and anesthesia. J Neurosci 36(41), 10696-10706. https://doi.org/10.1523/JNEUROSCI.1454-16.2016.

Carney, L.H., Ketterer, A.D., Abrams, K.S., Schwarz, D.M., 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.

Carney, L.H., Zilany, M.S.A., Huang, N.J., Abrams, K.S., Idrobo, F., 2014. Suboptimal use of neural information in a mammalian auditory system. J Neurosci 34(4), 1306-1313. https://doi.org/10.1523/JNEUROSCI.3031-13.2014.

Casseday, J.H., Ehrlich, D., Covey, E., 1994. Neural tuning for sound duration: role of inhibitory mechanisms in the inferior colliculus. Science 264(5160), 847-850. https://doi.org/10.1126/science.8171341.

de Lafuente, V., Romo, R., 2006. Neural correlate of subjective sensory experience gradually builds up across cortical areas. Proc Natl Acad Sci U S A 103(39), 14266-14271. https://doi.org/10.1073/pnas.0605826103.

Dent, M.L., Klump, G.M., Schwenzfeier, C., 2002. Temporal modulation transfer functions in the barn owl (*Tyto alba*). J Comp Physiol A Neuroethol Sens Neural Behav Physiol 187(12), 937-943. https://doi.org/10.1007/s00359-001-0259-5.

Dong, C., Qin, L., Liu, Y., Zhang, X., Sato, Y., 2011. Neural responses in the primary auditory cortex of freely behaving cats while discriminating fast and slow click-trains. PLoS One 6(10), e25895. https://doi.org/10.1371/journal.pone.0025895.

Dong, C., Qin, L., Zhao, Z., Zhong, R., Sato, Y., 2013. Behavioral modulation of neural encoding of click-trains in the primary and nonprimary auditory cortex of cats. J Neurosci 33(32), 13126-13137. https://doi.org/10.1523/JNEUROSCI.1724-13.2013.

Dooling, R.J., Searcy, M.H., 1981. Amplitude-modulation thresholds for the parakeet (*Melopsittacus undulatus*). J Comp Physiol 143(3), 383-388. https://doi.org/10.1007/Bf00611177.

Ehret, G., Schreiner, C., 2005. Spectral and intensity coding in the auditory midbrain, in: Winer, J., Schreiner, C. (Eds.), The inferior colliculus. Springer, New York, pp. 312 - 345.

Foffani, G., Moxon, K.A., 2004. PSTH-based classification of sensory stimuli using ensembles of single neurons. J Neurosci Methods 135(1-2), 107-120. https://doi.org/10.1016/j.jneumeth.2003.12.011.
Geis, H.-R., Borst, J.G.G., 2009. Intracellular Responses of Neurons in the Mouse Inferior Colliculus to Sinusoidal Amplitude-Modulated Tones. J Neurophysiol 101(4), 2002-2016. https://doi.org/10.1152/jn.90966.2008.
Geis, H.R.A.P., van der Heijden, M., Borst, J.G.G., 2011. Subcortical input heterogeneity in the mouse inferior colliculus. J Physiol 589(Pt 16), 3955-3967.
Goris, R.L.T., Ziemba, C.M., Stine, G.M., Simoncelli, E.P., Movshon, J.A., 2017. Dissociation of choice formation and choice-correlated activity in macaque visual cortex. J Neurosci 37(20), 5195-5203. https://doi.org/10.1523/JNEUROSCI.3331-16.2017.
Guo, Z.V., Li, N., Huber, D., Ophir, E., Gutnisky, D., Ting, J.T., Feng, G., Svoboda, K., 2014. Flow of cortical activity underlying a tactile decision in mice. Neuron 81(1), 179-194. https://doi.org/10.1016/j.neuron.2013.10.020.
Haefner, R.M., Gerwinn, S., Macke, J.H., Bethge, M., 2013. Inferring decoding strategies from choice probabilities in the presence of correlated variability. Nat Neurosci 16(2), 235-242. https://doi.org/10.1038/nn.3309.
Henderson, D., Salvi, R., Pavek, G., Hamernik, R., 1984. Amplitude modulation thresholds in chinchillas with high-frequency hearing loss. J Acoust Soc Am 75(4), 1177-1183. https://doi.org/10.1121/1.390767.
Henry, K.S., Neilans, E.G., Abrams, K.S., Idrobo, F., Carney, L.H., 2016. Neural correlates of behavioral amplitude modulation sensitivity in the budgerigar midbrain. J Neurophysiol 115(4), 1905-1916. https://doi.org/10.1152/jn.01003.2015.
Hernández, O., Espinosa, N., Pérez-González, D., Malmierca, M.S., 2005. The inferior colliculus of the rat: a quantitative analysis of monaural frequency response areas. Neuroscience 132(1), 203-217. https://doi.org/10.1016/j.neuroscience.2005.01.001.
Jazayeri, M., Movshon, J.A., 2006. Optimal representation of sensory information by neural populations. Nat Neurosci 9(5), 690-696. https://doi.org/10.1038/nn1691.
Johnson, J.S., Yin, P., O'Connor, K.N., Sutter, M.L., 2012. Ability of primary auditory cortical neurons to detect amplitude modulation with rate and temporal codes: neurometric analysis. J Neurophysiol 107(12), 3325-3341. https://doi.org/10.1152/jn.00812.2011.
Joris, P.X., Schreiner, C.E., Rees, A., 2004. Neural processing of amplitude-modulated sounds. Physiol Rev 84(2), 541-577. https://doi.org/10.1152/physrev.00029.2003.
Kelly, J.B., Cooke, J.E., Gilbride, P.C., Mitchell, C., Zhang, H., 2006. Behavioral limits of auditory temporal resolution in the rat: amplitude modulation and duration discrimination. J Comp Psychol 120(2), 98-105. https://doi.org/10.1037/0735-7036.120.2.98.
Kettner, R.E., Thompson, R.F., 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(6), 2111-2127. https://doi.org/10.1121/1.391736.
Klump, G.M., Okanoya, K., 1991. Temporal modulation transfer functions in the European starling (*Sturnus vulgaris*): I. Psychophysical modulation detection thresholds. Hear Res 52(1), 1-11. https://doi.org/10.1016/0378-5955(91)90182-9.
Kobak, D., Brendel, W., Constantinidis, C., Feierstein, C.E., Kepecs, A., Mainen, Z.F., Qi, X.L., Romo, R., Uchida, N., Machens, C.K., 2016. Demixed principal component analysis of neural population data. Elife 5. https://doi.org/10.7554/eLife.10989.
Kohlrausch, A., Fassel, R., Dau, T., 2000. The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. J Acoust Soc Am 108(2), 723-734. https://doi.org/10.1121/1.429605.
Krebs, B., Lesica, N.A., Grothe, B., 2008. The representation of amplitude modulations in the mammalian auditory midbrain. J Neurophysiol 100(3), 1602-1609. https://doi.org/10.1152/jn.90374.2008.

788 Leary, C.J., Edwards, C.J., Rose, G.J., 2008. Midbrain auditory neurons integrate excitation and inhibition
789 to generate duration selectivity: an *in vivo* whole-cell patch study in anurans. J Neurosci 28(21), 5481-
790 5493. https://doi.org/10.1523/jneurosci.5041-07.2008.
791 Lu, T., Wang, X., 2004. Information content of auditory cortical responses to time-varying acoustic
792 stimuli. J Neurophysiol 91(1), 301-313. https://doi.org/10.1152/jn.00022.2003.
793 Macmillan, N.A., Creelman, C.D., 2004. Detection theory: A user's guide. Psychology press.
794 Malone, B.J., Scott, B.H., Semple, M.N., 2010. Temporal codes for amplitude contrast in auditory cortex.
795 J Neurosci 30(2), 767-784. https://doi.org/10.1523/JNEUROSCI.4170-09.2010.
796 Marshel, J.H., Kim, Y.S., Machado, T.A., Quirin, S., Benson, B., Kadmon, J., Raja, C., Chibukhchyan, A.,
797 Ramakrishnan, C., Inoue, M., Shane, J.C., McKnight, D.J., Yoshizawa, S., Kato, H.E., Ganguli, S.,
798 Deisseroth, K., 2019. Cortical layer-specific critical dynamics triggering perception. Science 365(6453).
799 https://doi.org/10.1126/science.aaw5202.
800 McInnes, L., Healy, J., Melville, J., 2018. UMAP: Uniform Manifold Approximation and Projection for
801 Dimension Reduction. p. arXiv:1802.03426.
802 Moody, D.B., 1994. Detection and discrimination of amplitude-modulated signals by macaque monkeys.
803 J Acoust Soc Am 95(6), 3499-3510. https://doi.org/10.1121/1.409967.
804 Moon, K.R., van Dijk, D., Wang, Z., Gigante, S., Burkhardt, D.B., Chen, W.S., Yim, K., Elzen, A.V.D., Hirn,
805 M.J., Coifman, R.R., Ivanova, N.B., Wolf, G., Krishnaswamy, S., 2019. Visualizing structure and transitions
806 in high-dimensional biological data. Nat Biotechnol 37(12), 1482-1492. https://doi.org/10.1038/s41587-
807 019-0336-3.
808 Müller, M.K., Jovanovic, S., Keine, C., Radulovic, T., Rübsamen, R., Milenkovic, I., 2019. Functional
809 development of principal neurons in the anteroventral cochlear nucleus extends beyond hearing onset.
810 Front Cell Neurosci 13, 119. https://doi.org/10.3389/fncel.2019.00119.
811 Newsome, W.T., Britten, K.H., Movshon, J.A., 1989. Neuronal correlates of a perceptual decision. Nature
812 341(6237), 52-54. https://doi.org/10.1038/341052a0.
813 Niwa, M., Johnson, J.S., O'Connor, K.N., Sutter, M.L., 2012a. Active engagement improves primary
814 auditory cortical neurons' ability to discriminate temporal modulation. J Neurosci 32(27), 9323-9334.
815 https://doi.org/10.1523/JNEUROSCI.5832-11.2012.
816 Niwa, M., Johnson, J.S., O'Connor, K.N., Sutter, M.L., 2012b. Activity related to perceptual judgment and
817 action in primary auditory cortex. J Neurosci 32(9), 3193-3210.
818 https://doi.org/10.1523/JNEUROSCI.0767-11.2012.
819 Niwa, M., O'Connor, K.N., Engall, E., Johnson, J.S., Sutter, M.L., 2015. Hierarchical effects of task
820 engagement on amplitude modulation encoding in auditory cortex. J Neurophysiol 113(1), 307-327.
821 https://doi.org/10.1152/jn.00458.2013.
822 Ono, M., Bishop, D.C., Oliver, D.L., 2017. Identified GABAergic and glutamatergic neurons in the mouse
823 inferior colliculus share similar response properties. J Neurosci 37(37), 8952-8964.
824 https://doi.org/10.1523/JNEUROSCI.0745-17.2017.
825 Pachitariu, M., Steinmetz, N., Kadir, S., Carandini, M., 2016. Kilosort: realtime spike-sorting for
826 extracellular electrophysiology with hundreds of channels. BioRxiv, 061481.
827 https://doi.org/10.1101/061481.
828 Palmer, A.R., Russell, I.J., 1986. Phase-locking in the cochlear nerve of the guinea-pig and its relation to
829 the receptor potential of inner hair-cells. Hearing Research 24(1), 1-15. https://doi.org/10.1016/0378-
830 5955(86)90002-X.
831 Palmer, A.R., Shackleton, T.M., Sumner, C.J., Zobay, O., Rees, A., 2013. Classification of frequency
832 response areas in the inferior colliculus reveals continua not discrete classes. J Physiol 591(Pt 16), 4003-
833 4025. https://doi.org/10.1113/jphysiol.2013.255943.
834 Pauli-Magnus, D., Hoch, G., Strenzke, N., Anderson, S., Jentsch, T.J., Moser, T., 2007. Detection and
835 differentiation of sensorineural hearing loss in mice using auditory steady-state responses and transient

auditory brainstem responses. Neuroscience 149(3), 673-684. https://doi.org/10.1016/j.neuroscience.2007.08.010.
Portfors, C.V., Mayko, Z.M., Jonson, K., Cha, G.F., Roberts, P.D., 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, P.L., Brown, S.D., McKoon, G., 2016. Diffusion Decision Model: Current Issues and History. Trends Cogn Sci 20(4), 260-281. https://doi.org/10.1016/j.tics.2016.01.007.
Rees, A., Langner, G., 2005. Temporal coding in the auditory midbrain, in: Winer, J., Schreiner, C. (Eds.), The inferior colliculus. Springer, New York, pp. 346 - 376.
Rees, A., Møller, A.R., 1987. Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. Hear Res 27(2), 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., Harris, K.D., 2010. The asynchronous state in cortical circuits. Science 327(5965), 587-590. https://doi.org/10.1126/science.1179850.
Rodríguez, F.A., Read, H.L., Escabí, M.A., 2010. Spectral and temporal modulation tradeoff in the inferior colliculus. J Neurophysiol 103(2), 887-903. https://doi.org/10.1152/jn.00813.2009.
Rosen, M.J., Semple, M.N., Sanes, D.H., 2010. Exploiting development to evaluate auditory encoding of amplitude modulation. J Neurosci 30(46), 15509-15520. https://doi.org/10.1523/JNEUROSCI.3340-10.2010.
Salvi, R.J., Giraudi, D.M., Henderson, D., Hamernik, R.P., 1982. Detection of sinusoidally amplitude modulated noise by the chinchilla. J Acoust Soc Am 71(2), 424-429. https://doi.org/10.1121/1.387445.
Schnupp, J.W.H., Garcia-Lazaro, J.A., Lesica, N.A., 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, J.W.H., Hall, T.M., Kokelaar, R.F., Ahmed, B., 2006. Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. J Neurosci 26(18), 4785-4795. https://doi.org/10.1523/JNEUROSCI.4330-05.2006.
Schreiner, C.E., Langner, G., 1988. Periodicity coding in the inferior colliculus of the cat. II. Topographical organization. J Neurophysiol 60(6), 1823-1840. https://doi.org/10.1152/jn.1988.60.6.1823.
Shamash, P., Carandini, M., Harris, K., Steinmetz, N., 2018. A tool for analyzing electrode tracks from slice histology. BioRxiv, 447995.
Stiebler, I., 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(1), 65-76. https://doi.org/10.1002/cne.902380106.
Stringer, C., Michaelos, M., Tsyboulski, D., Lindo, S.E., Pachitariu, M., 2021. High-precision coding in visual cortex. Cell 184(10), 2767-2778 e2715. https://doi.org/10.1016/j.cell.2021.03.042.
Tanke, N., Borst, J.G.G., Houweling, A.R., 2018. Single-cell stimulation in barrel cortex influences psychophysical detection performance. J Neurosci 38(8), 2057-2068. https://doi.org/10.1523/JNEUROSCI.2155-17.2018.
Ter-Mikaelian, M., Sanes, D.H., Semple, M.N., 2007. Transformation of temporal properties between auditory midbrain and cortex in the awake Mongolian gerbil. J Neurosci 27(23), 6091-6102. https://doi.org/10.1523/JNEUROSCI.4848-06.2007.
van der Maaten, L., Hinton, G., 2008. Visualizing Data using t-SNE. J Mach Learn Res 9, 2579-2605.
van Looij, M.A.J., Liem, S.S., van der Burg, H., van der Wees, J., De Zeeuw, C.I., van Zanten, B.G.A., 2004. Impact of conventional anesthesia on auditory brainstem responses in mice. Hear Res 193(1-2), 75-82.

von Trapp, G., Buran, B.N., Sen, K., Semple, M.N., Sanes, D.H., 2016. A decline in response variability improves neural signal detection during auditory task performance. J Neurosci 36(43), 11097-11106. https://doi.org/10.1523/JNEUROSCI.1302-16.2016.

Walker, K.M.M., Ahmed, B., Schnupp, J.W.H., 2008. Linking cortical spike pattern codes to auditory perception. J Cogn Neurosci 20(1), 135-152. https://doi.org/10.1162/jocn.2008.20012.

Walton, J.P., Simon, H., Frisina, R.D., 2002. Age-related alterations in the neural coding of envelope periodicities. J Neurophysiol 88(2), 565-578. https://doi.org/10.1152/jn.2002.88.2.565.

Wang, L., Narayan, R., Graña, G., Shamir, M., Sen, K., 2007. Cortical discrimination of complex natural stimuli: can single neurons match behavior? J Neurosci 27(3), 582-589. https://doi.org/10.1523/JNEUROSCI.3699-06.2007.

Wang, Q., Ding, S.L., Li, Y., Royall, J., Feng, D., Lesnar, P., Graddis, N., Naeemi, M., Facer, B., Ho, A., Dolbeare, T., Blanchard, B., Dee, N., Wakeman, W., Hirokawa, K.E., Szafer, A., Sunkin, S.M., Oh, S.W., Bernard, A., Phillips, J.W., Hawrylycz, M., Koch, C., Zeng, H., Harris, J.A., Ng, L., 2020. The Allen Mouse Brain Common Coordinate Framework: A 3D reference atlas. Cell 181(4), 936-953 e920. https://doi.org/10.1016/j.cell.2020.04.007.

Wang, Y., Abrams, K.S., Carney, L.H., Henry, K.S., 2021. Midbrain-level neural correlates of behavioral tone-in-noise detection: dependence on energy and envelope cues. J Neurosci 41(34), 7206-7223. https://doi.org/10.1523/JNEUROSCI.3103-20.2021.

Yin, P., Johnson, J.S., O'Connor, K.N., Sutter, M.L., 2011. Coding of amplitude modulation in primary auditory cortex. J Neurophysiol 105(2), 582-600. https://doi.org/10.1152/jn.00621.2010.

Zhang, H., Kelly, J.B., 2003. Glutamatergic and GABAergic regulation of neural responses in inferior colliculus to amplitude-modulated sounds. J Neurophysiol 90(1), 477-490. https://doi.org/10.1152/jn.01084.2002.

Figure Captions

Figure 1. Experimental design. (A) Behavioral trials consisted of unmodulated (UM) noise stimuli of random 0.5-3 s duration that transitioned to amplitude-modulated (AM) noise stimuli with matched intensity and spectral properties. During the UM period, any lick from the mouse delayed the start of the AM transition until it did not lick for 4-7 s; the first lick within the 1 s reward window, which started at the transition to AM, was rewarded with a water drop. End of the trial was 500 ms after the response (black stimulus trace) or 1 s after its start (end of grey stimulus), whichever came first. (B) Electrophysiology trial structure. The UM and AM periods were both fixed at 1 s. Example traces from four adjacent channels is shown, and a segment of approximately one modulation period shown in the insert on the right. 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 transition to AM occurred at t = 0 s (solid vertical line). The dotted vertical line marks the end of the response window for calculation of hit rates in Figure 3. UM is the unrewarded catch trial condition. (B) Cumulative distribution of response latencies after the AM transition for 512 Hz (top) and 16 Hz (bottom) at various modulation depths m. Response latencies shorter than 75 ms were considered spontaneous responses and were excluded from analysis. (C) Receiver-Operator-Characteristics (ROC) analysis for response latency, plotting cumulative probability of a response to AM versus UM stimuli as the criterion (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 $\boldsymbol{d'} = |\sqrt{\mathbf{2}} * \boldsymbol{\Phi^{-1}}(\boldsymbol{AUC})|$, where $\boldsymbol{\Phi^{-1}}$is the inverse cumulative distribution function of the standard normal distribution. The vertical dotted lines indicate the estimated modulation depth at which d' = 1. 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 512 (upper) or 16 Hz (lower) AM at different modulation depths. (B) Median response latencies of animals at 512 and 16 Hz, showing faster responses at higher modulation depths. (C) Response latencies at 512 (upper) and 16 (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: $p < 0.0001$).

Figure 4. Example unit from the ICc with an onset response and relatively good phase-locking to AM stimuli. (A) Spike waveforms detected on nine 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 from top to bottom by decreasing

950 modulation depth for the different modulation frequencies. Transition from UM to AM occurred at 1 s.
951 (E) Cycle histograms of peak phase-locking modulation depth for the same modulation frequencies as
952 shown in D. Bin size is 0.13 ms. (F) Temporal coding of AM was quantified by calculating the phase-
953 projected vector strength (VS_{PP}) for the responses during the AM period. (G) The average firing rate
954 during the AM period is used as a measure for rate coding to AM. All AM data shown are for a carrier
955 noise intensity of 60 dB SPL.

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

966 Figure 6. Population AM encoding. (A) Average relation between VS_{PP} and modulation depth. (B)
967 Conversion of VS_{PP} to d' using ROC analysis. (C) AM detection thresholds. Percentages denote the
968 percentage of cells for which a threshold could be found. Thresholds were defined as the modulation
969 depth for which d' = 1. Black-edged markers indicate the mean detection thresholds obtained from
970 behavioral experiments. (D) Average difference between firing rate of AM stimuli and that of the
971 intensity-matched UM stimuli, as a function of modulation depth. (E-F) As B-C, but for firing rate instead
972 of vector strength. Error bars in the figure show SD. All AM data shown are for a carrier noise intensity
973 of 60 dB SPL.

974 Figure 7. Comparing neurometric temporal and rate AM encoding of awake (dashed lines, n = 99 units)
975 and anesthetized (dotted lines, n = 112 units) recordings to behavioral performance (solid lines, n = 8
976 animals). (A-D) VS_{PP}-derived d' plotted against modulation depth. (E-F) Firing rate-derived d' plotted
977 against modulation depth. mean ± SD. Modulation frequency color scheme follows that of previous
978 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
979 shown are for a carrier noise intensity of 60 dB SPL.

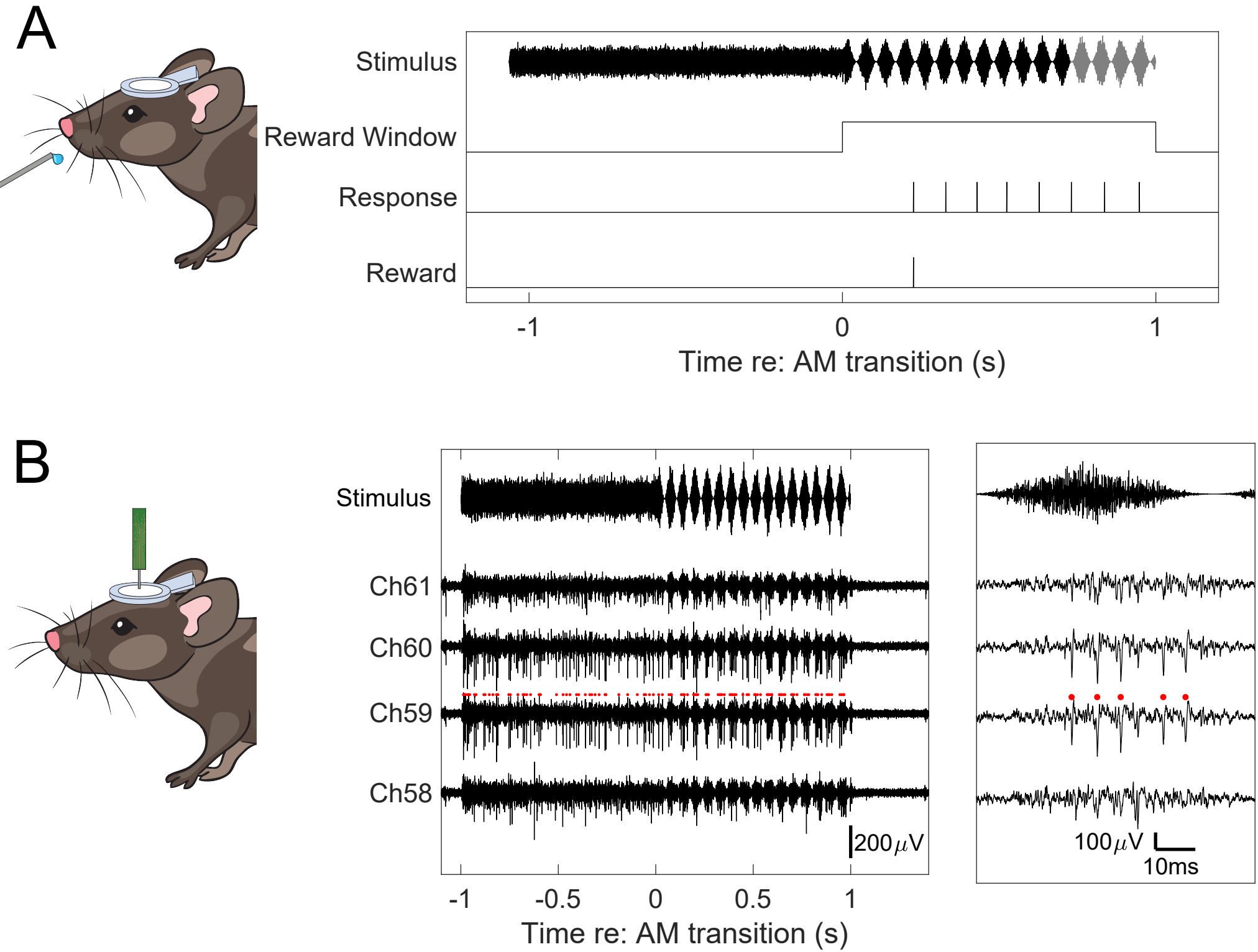
980 Figure 8. Spatial organization of pure tone and AM encoding of recorded units. Location of the units was
981 obtained by registering probe tracts to the mouse Allen Brain Atlas. (A) Right: Coronal section showing
982 the DiO-labelled probe tract. Left: corresponding virtual slice from the Allen Brain Atlas CCFv3 to
983 illustrate the parcellation of the IC. (B) Sagittal view of units in the left IC. Color of symbols represents CF
984 and shape represents location of units within subnuclei of the IC. (C) as (B), but in coronal view. (D) Left:
985 Distribution of CFs in the different subnuclei of the IC. Right: Distribution of CFs in recordings from
986 anesthetized and awake animals. (E) Top: distribution of temporal best modulation frequencies (tBMF)
987 in the different subnuclei of the IC. Bottom: distribution of rate best modulation frequencies (rBMF)
988 within the different subnuclei of the IC. Chi-Squared tests showed significant differences among all the
989 subnuclei in their rBMF distributions.

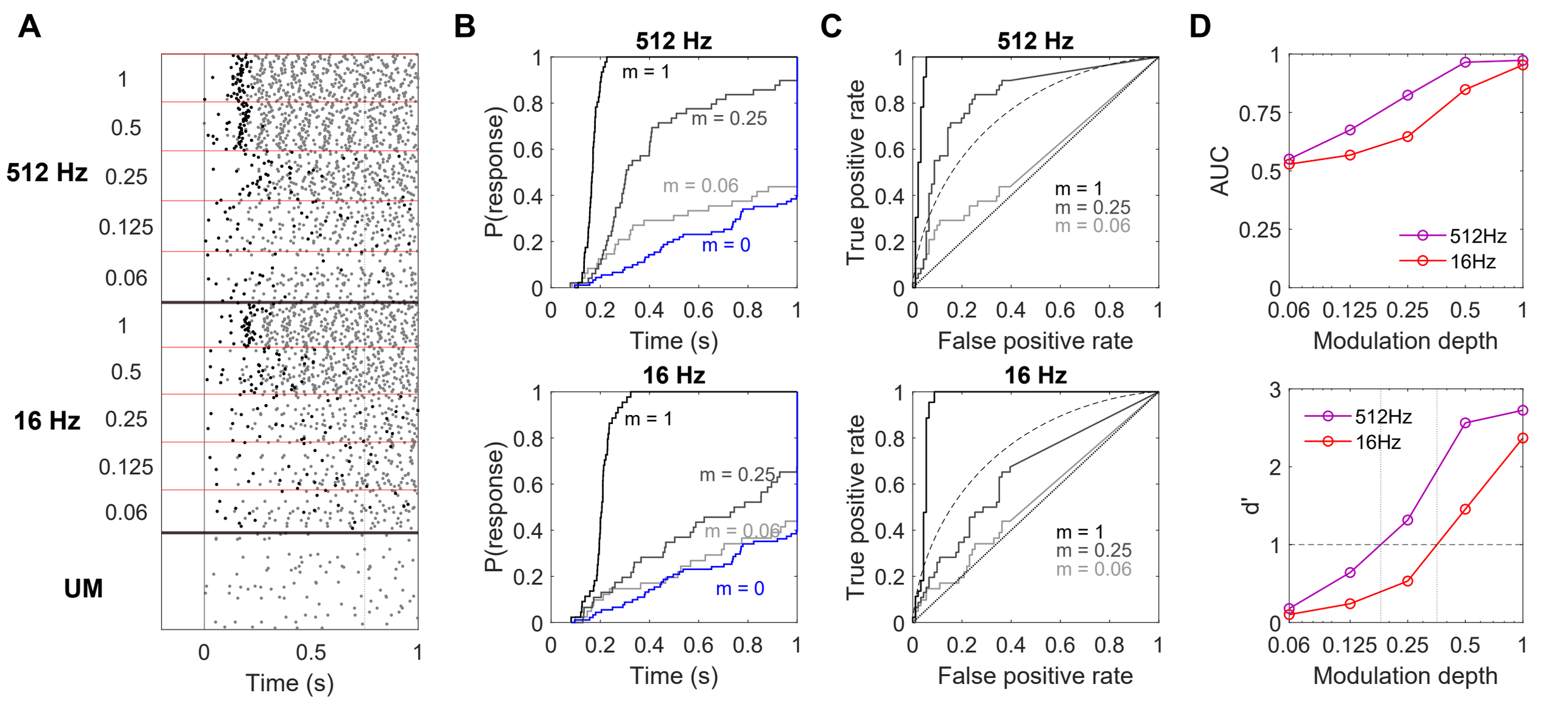
Figure 9. Principal component analysis of the difference in responses to AM or UM stimuli. (A) Dot raster plot of two example units in response to fully modulated stimuli at various frequencies and the unmodulated control stimulus (UM). (B) Average peristimulus time histograms (PSTHs) of the same units 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.

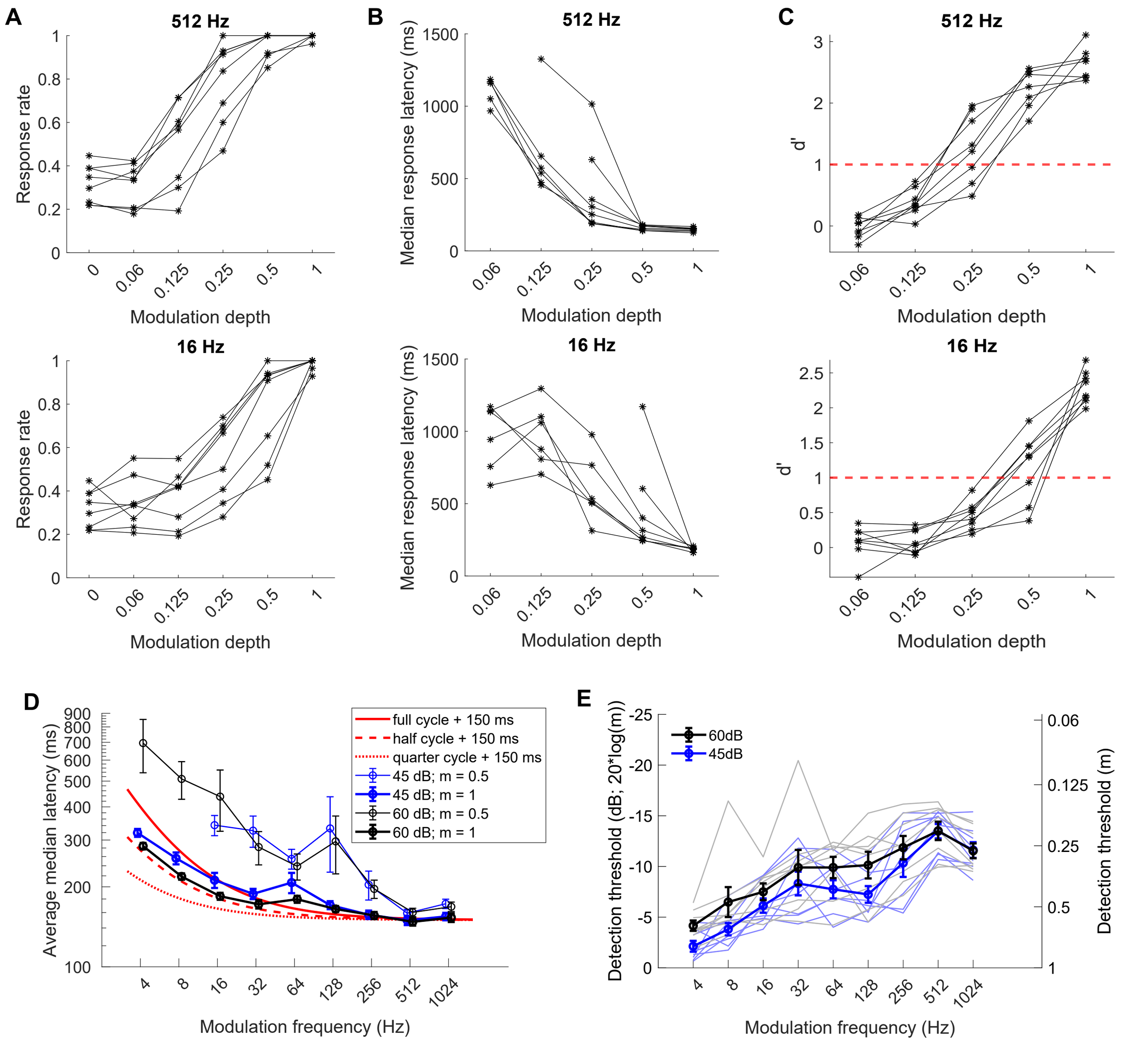
Figure 10. Overview of neurons from anaesthetized recordings ranked according to their contribution to PC1. (A) All neurons from anaesthetized recordings sorted by their contribution to PC1 (left panel, $|weight_{PC1}|$), their average PSTH in response to sound onset (second from left) or to 16, 64, 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.

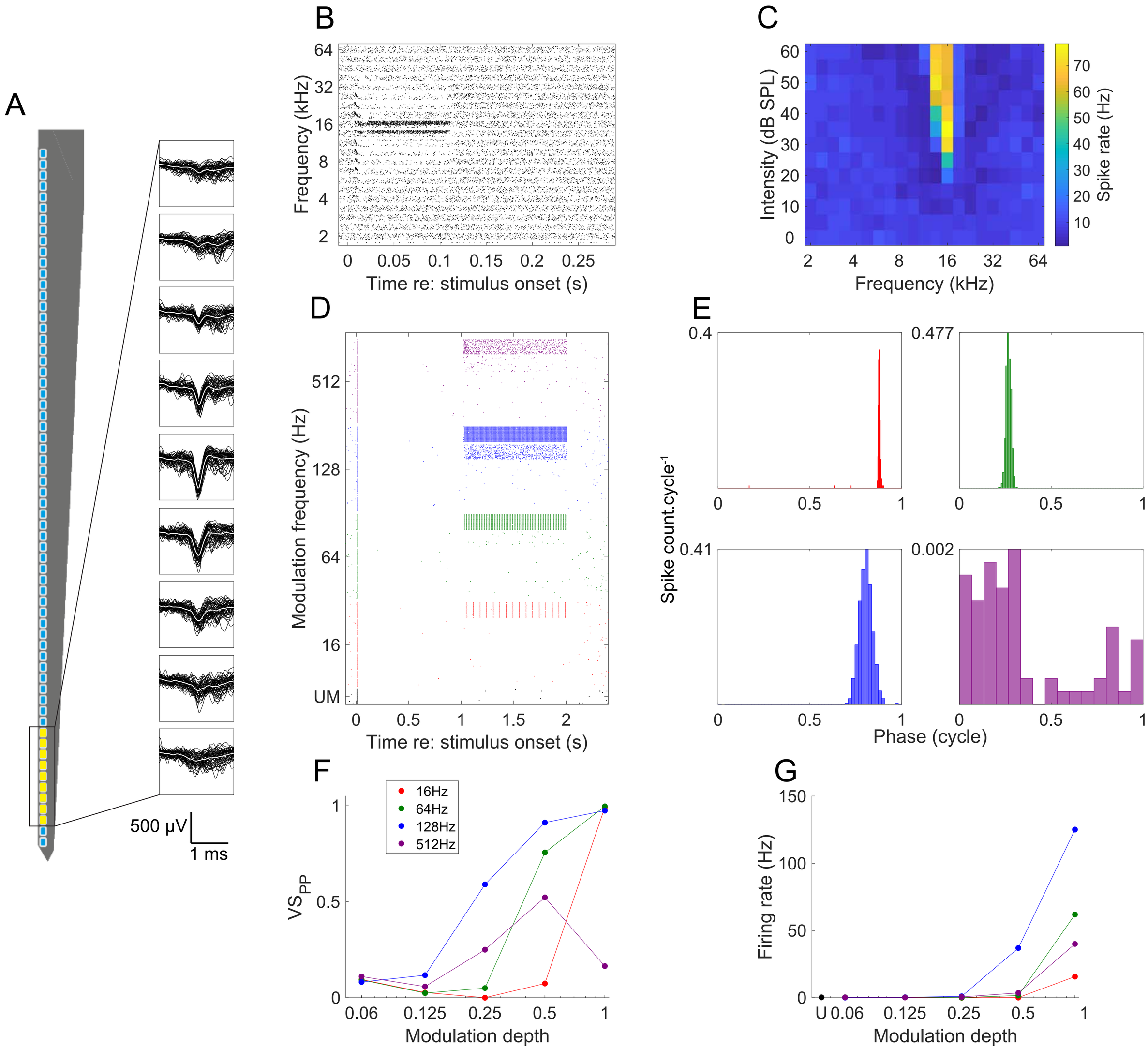
Figure 11. Overview of neurons from awake recordings ranked according to their contribution to PC1. (A-C) As Figure 10A-C, respectively, except recordings were from awake animals.

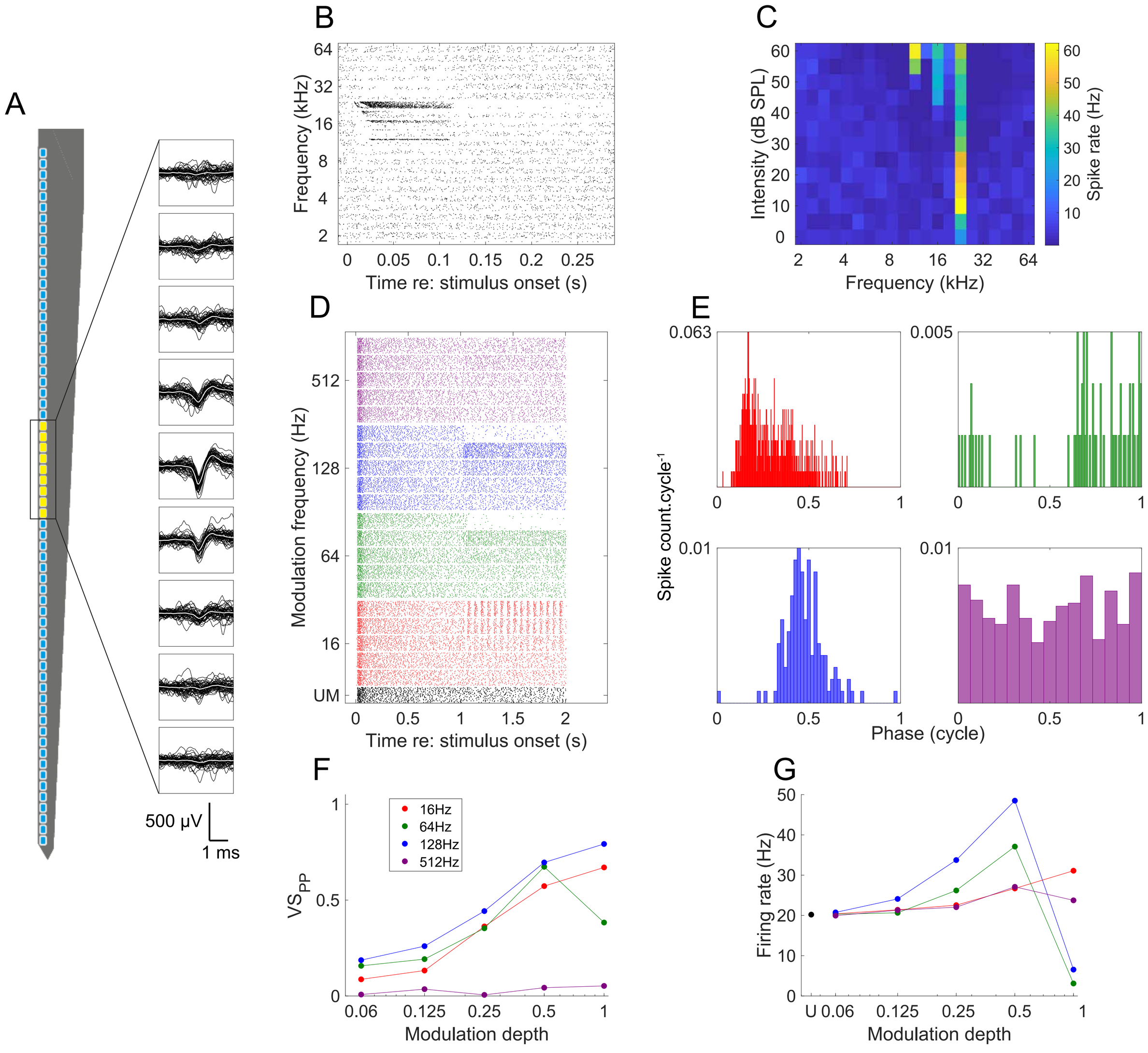
Figure 12. Using PCA of neuronal responses to AM stimuli to infer behavioral thresholds. (A) Single trial evolution of PC1 over time for selected modulation frequencies and depths at 60 dB for the data from the anaesthetized mice. Traces were created from kernel spike density functions with a Gaussian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

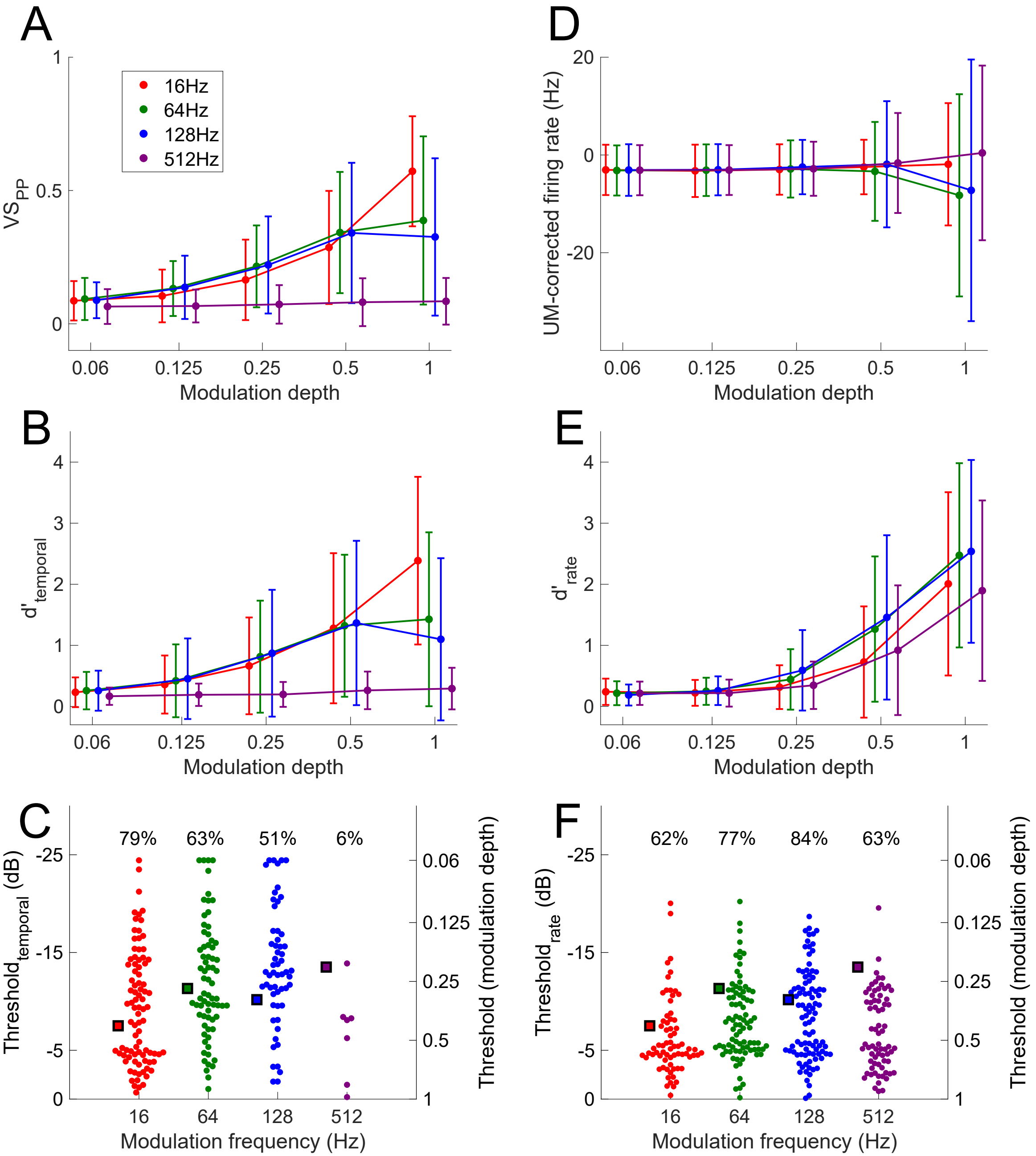

A
Stimulus
Reward Window
Response
Reward
-1
0
1
Time re: AM transition (s)
B
Stimulus
Ch61
Ch60
Ch59
Ch58
200 μV
-1
-0.5
0
0.5
1
Time re: AM transition (s)
100 μV
10ms

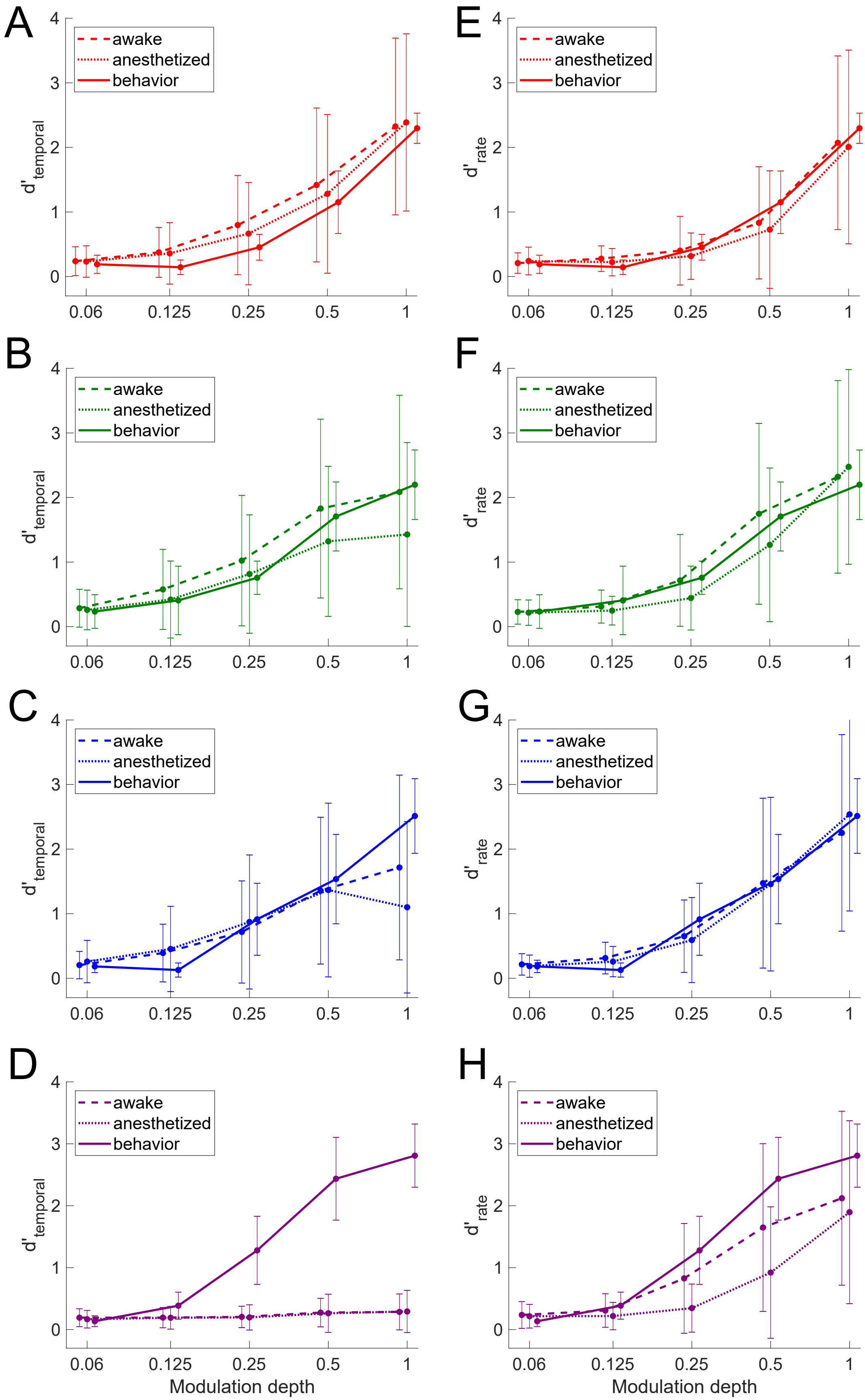

A
512 Hz
16 Hz
UM
1
0.5
0.25
0.125
0.06
0
0.5
1
Time (s)
B
512 Hz
16 Hz
P(response)
m = 1
m = 0.25
m = 0.06
m = 0
Time (s)
C
512 Hz
16 Hz
True positive rate
False positive rate
m = 1
m = 0.25
m = 0.06
D
AUC
d'
512Hz
16Hz
Modulation depth
0.06
0.125
0.25
0.5
1

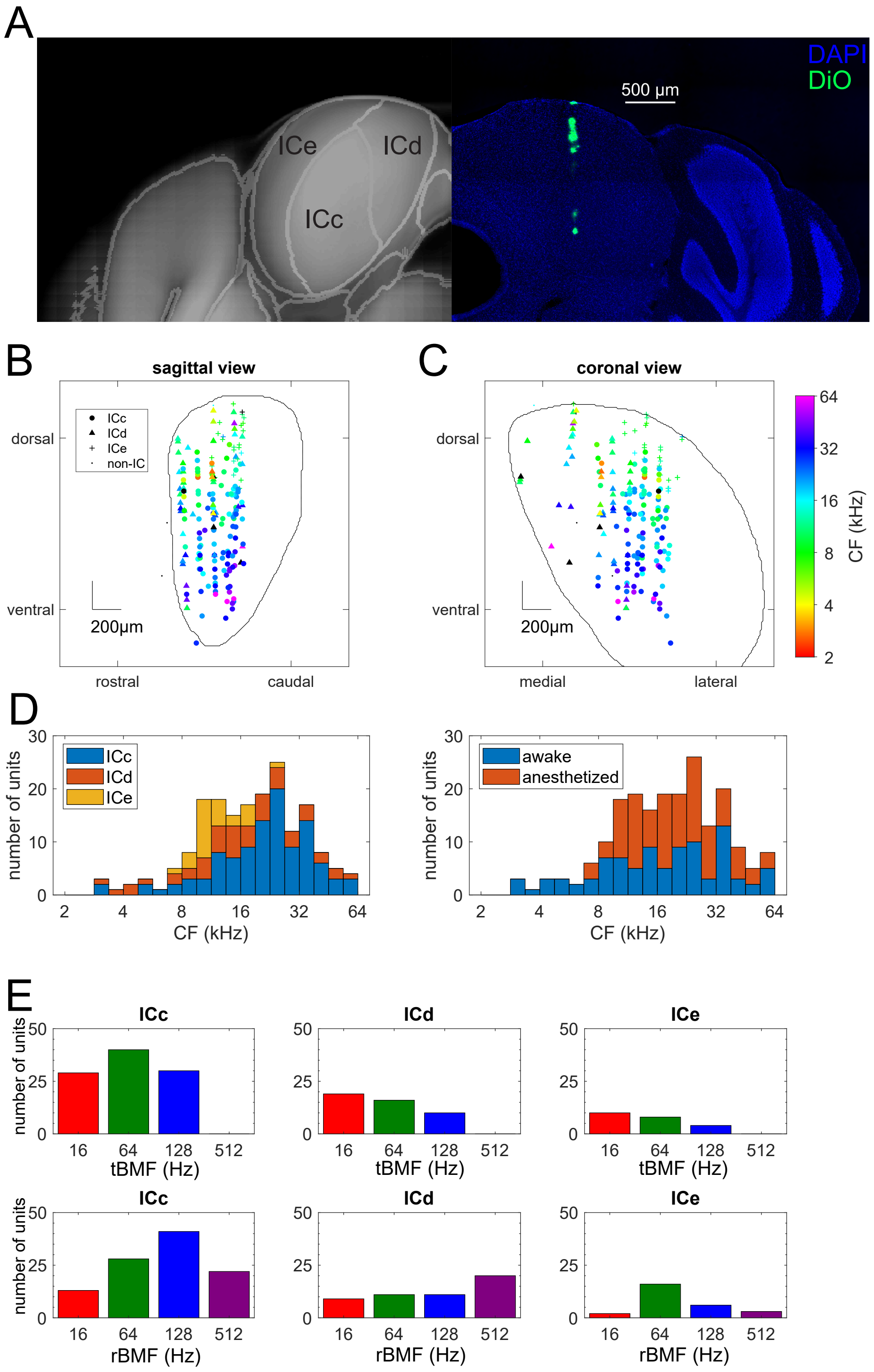

A
512 Hz
Response rate
Modulation depth
16 Hz
B
Median response latency (ms)
C
d'
D
Average median latency (ms)
Modulation frequency (Hz)
full cycle + 150 ms
half cycle + 150 ms
quarter cycle + 150 ms
45 dB; m = 0.5
45 dB; m = 1
60 dB; m = 0.5
60 dB; m = 1
E
Detection threshold (dB; 20*log(m))
Detection threshold (m)
60dB
45dB

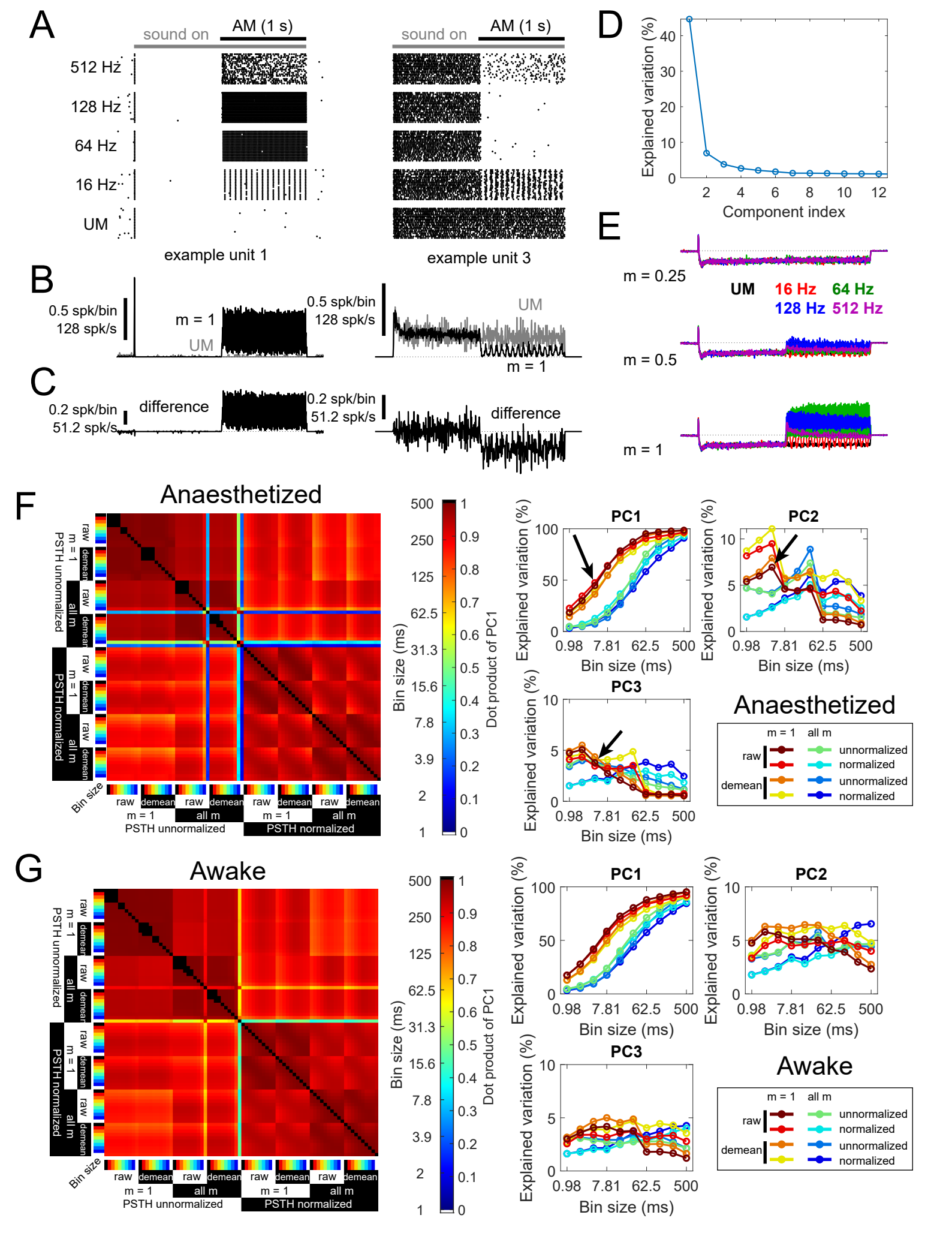

A
500 µV
1 ms
B
Frequency (kHz)
64
32
16
8
4
2
0
0.05
0.1
0.15
0.2
0.25
Time re: stimulus onset (s)
C
Intensity (dB SPL)
60
50
40
30
20
10
0
2
4
8
16
32
64
Frequency (kHz)
Spike rate (Hz)
70
60
50
40
30
20
10
D
Modulation frequency (Hz)
512
128
64
16
UM
0
0.5
1
1.5
2
Time re: stimulus onset (s)
E
Spike count.cycle-1
0.4
0.477
0.41
0.002
0
0.5
1
Phase (cycle)
F
16Hz
64Hz
128Hz
512Hz
VSPP
1
0.5
0
0.06
0.125
0.25
0.5
1
Modulation depth
G
Firing rate (Hz)
150
100
50
0
U
0.06
0.125
0.25
0.5
1
Modulation depth

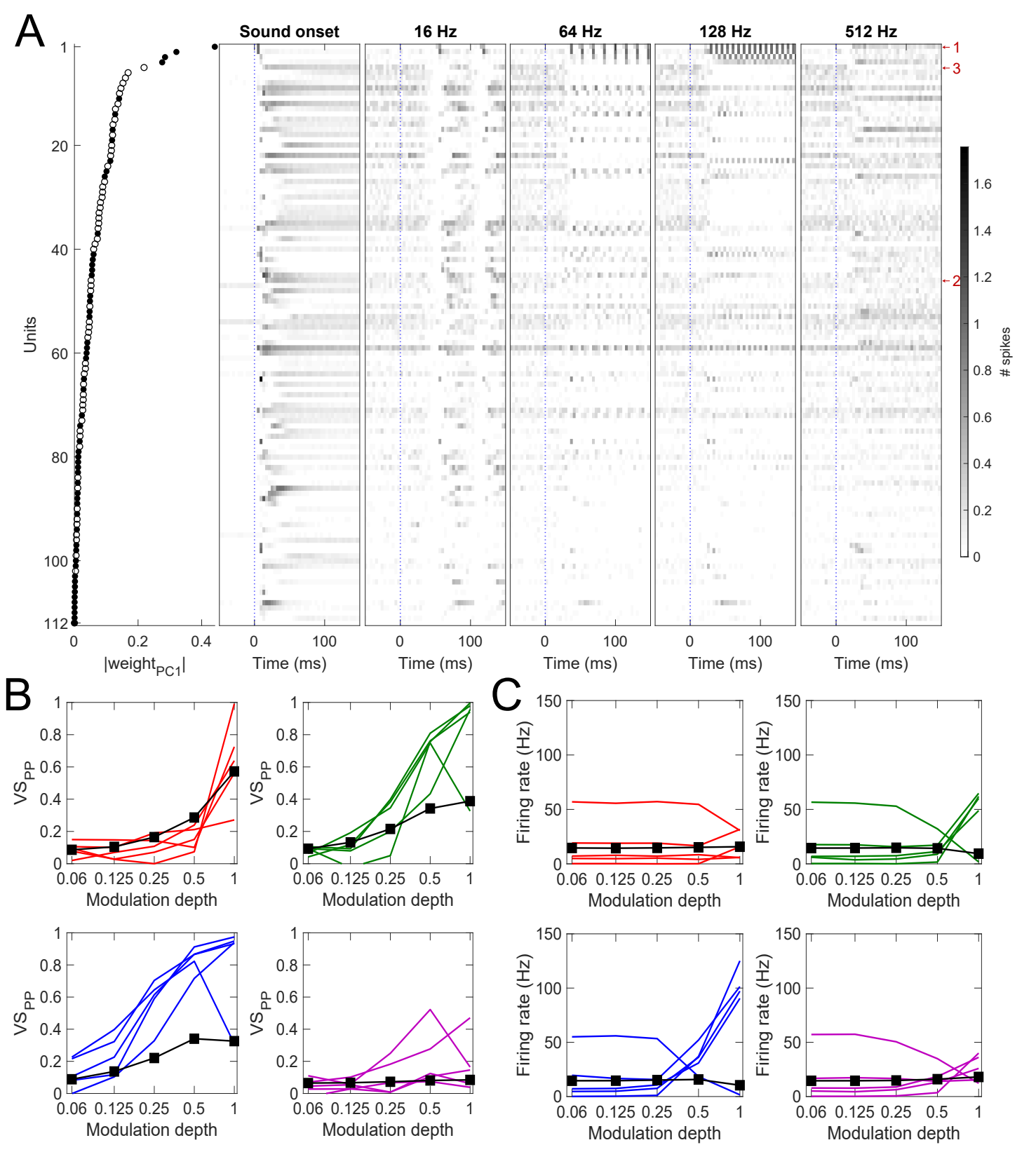

A
500 µV
1 ms
B
Frequency (kHz)
64
32
16
8
4
2
0
0.05
0.1
0.15
0.2
0.25
Time re: stimulus onset (s)
C
Intensity (dB SPL)
60
50
40
30
20
10
0
2
4
8
16
32
64
Frequency (kHz)
Spike rate (Hz)
D
Modulation frequency (Hz)
512
128
64
16
UM
0
0.5
1
1.5
2
Time re: stimulus onset (s)
E
Spike count.cycle-1
0.063
0.005
0.01
0.01
0
0.5
1
Phase (cycle)
F
VSPP
1
0.5
0
16Hz
64Hz
128Hz
512Hz
0.06
0.125
0.25
0.5
1
Modulation depth
G
Firing rate (Hz)
50
40
30
20
10
0
U
0.06
0.125
0.25
0.5
1
Modulation depth

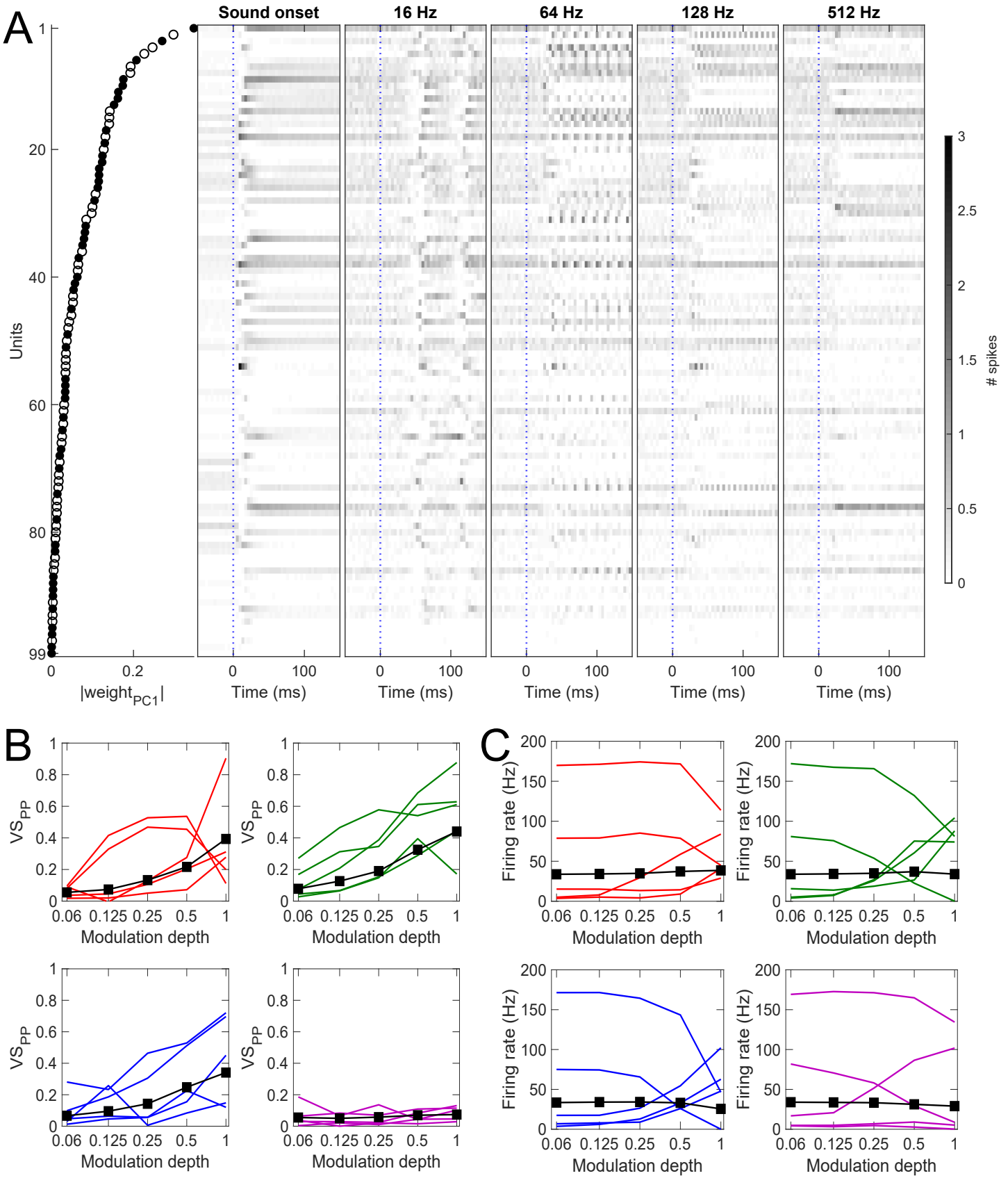

A
16Hz
64Hz
128Hz
512Hz
VS_PP
Modulation depth
B
d'_temporal
Modulation depth
C
79%
63%
51%
6%
Threshold_temporal (dB)
Threshold (modulation depth)
Modulation frequency (Hz)
D
UM-corrected firing rate (Hz)
Modulation depth
E
d'_rate
Modulation depth
F
62%
77%
84%
63%
Threshold_rate (dB)
Threshold (modulation depth)
Modulation frequency (Hz)

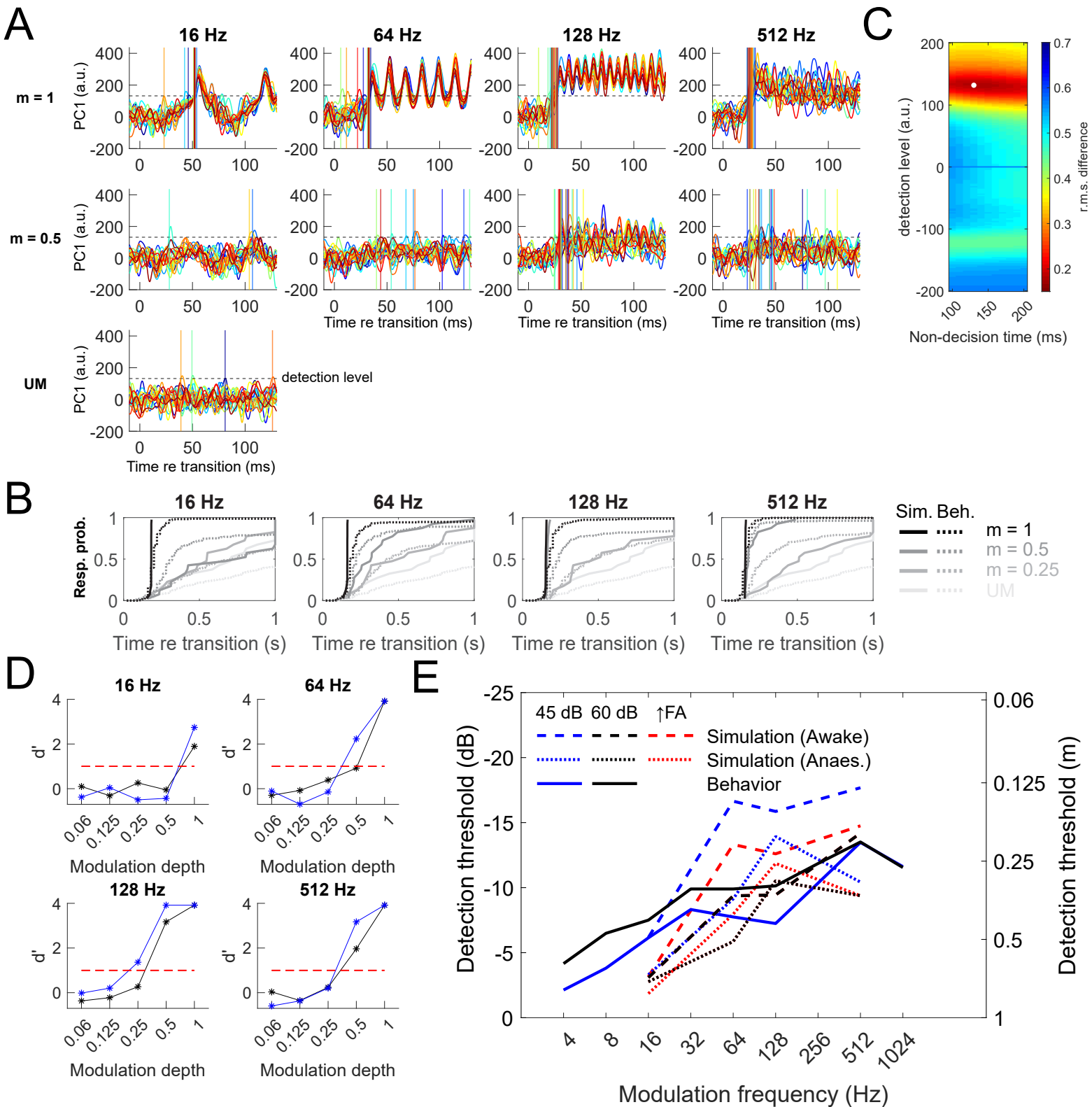

A
B
C
D
E
F
G
H
awake
anesthetized
behavior
d'temporal
d'rate
0
1
2
3
4
0.06
0.125
0.25
0.5
1
Modulation depth


A
ICe
ICd
ICc
500 µm
DAPI
DiO
B
sagittal view
ICc
ICd
ICe
non-IC
dorsal
ventral
200µm
rostral
caudal
C
coronal view
dorsal
ventral
200µm
medial
lateral
64
32
16
8
4
2
CF (kHz)
D
ICc
ICd
ICe
number of units
CF (kHz)
awake
anesthetized
number of units
CF (kHz)
E
ICc
ICd
ICe
number of units
tBMF (Hz)
rBMF (Hz)
16
64
128
512


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
raw
demean
m = 1
all m
Bin size (ms)
Dot product of PC1
PC1
PC2
PC3
Explained variation (%)
unnormalized
normalized
G
Awake


A
Sound onset
16 Hz
64 Hz
128 Hz
512 Hz
Units
|weight$_{PC1}$|
Time (ms)
spikes
B
C
VS$_{PP}$
Firing rate (Hz)
Modulation depth


A
Sound onset
16 Hz
64 Hz
128 Hz
512 Hz
Units
|weight$_{PC1}$|
Time (ms)
spikes
B
C
VS$_{PP}$
Modulation depth
Firing rate (Hz)


A
16 Hz
64 Hz
128 Hz
512 Hz
m = 1
m = 0.5
UM
PC1 (a.u.)
detection level
Time re transition (ms)
C
detection level (a.u.)
Non-decision time (ms)
r.m.s. difference
B
Resp. prob.
Time re transition (s)
Sim. Beh.
m = 1
m = 0.5
m = 0.25
UM
D
d'
Modulation depth
E
45 dB 60 dB ↑FA
Simulation (Awake)
Simulation (Anaes.)
Behavior
Detection threshold (dB)
Detection threshold (m)
Modulation frequency (Hz)