Selective effects of arousal on population coding of natural sounds in primary auditory cortex

Charles R. Heller[1,2], Zachary P. Schwartz[2] Daniela Saderi[2], Stephen V. David[2*]

[1] Neuroscience Graduate Program, Oregon Health and Science University

[2] Oregon Hearing Research Center, Oregon Health and Science University

[*]Corresponding author (davids@ohsu.edu)

Abstract

In addition to encoding sound stimulus features, activity in primary auditory cortex (A1) is modulated by non-sensory behavioral state variables, including arousal. Here, we investigated how arousal, measured by pupil size, influences stimulus discriminability in A1. To do this, we recorded from populations of A1 neurons in awake animals while presenting a diverse set of natural sound stimuli. In contrast to previous work, the large stimulus set allowed us to investigate effects of arousal across a wide range of sensory response space. Arousal consistently increased evoked response magnitude and reduced pairwise noise correlations. On average, these changes improved the accuracy of the neural code. However, effects varied across stimuli; neural coding was most improved for areas of the sensory space where noise correlations interfered with the sensory discrimination axis. We also found that first-order modulation of evoked responses and second-order modulation of correlated variability acted on distinct neural populations and timescales, suggesting that arousal interacts with multiple circuits underlying activity in A1.

Introduction

Reliable encoding of sound is critical for many behaviors. It has become increasingly clear, however, that this process is not static. Neural activity in the auditory system dynamically adapts to meet current behavioral demands. For example, engaging in a range of tasks causes auditory neurons to change their overall excitability and/or receptive field properties to enhance the neural representation of task-relevant stimuli.[1–10] However, less is understood about how

the responses of single auditory neurons act in concert to support perception and how coding at the population level is affected by behavioral state changes.[9,11–17]

Across different species and sensory systems, there is mounting evidence that behavioral state changes in arousal, task-engagement, and attention modulate population coding accuracy. One way this is studied is to measure correlated intrinsic activity, or noise correlation, whose strength is generally reduced during behavioral states associated with improved perception.[9,16,18–22] A wealth of theoretical[23–30] and more recently experimental work,[31–33] has shown that correlated activity can fundamentally limit the accuracy of sensory coding by the brain. Thus, reduced correlations may underlie the increased perceptual ability observed during states such as selective attention.[19] Yet, this finding remains the topic of rich debate.

Multiple studies have alternatively shown that correlated noise has no impact on, or can even improve sensory coding.[34–38] Despite these seemingly contradictory results, there is agreement that correlated noise only limits information in the specific case that it aligns with the sensory signal to be decoded.[30] Thus, the impact of correlated noise on sensory coding depends critically on the source of the noise (*i.e.* is the noise carried by the same neurons that encode the sensory signal). This is still an area of open experimental investigation. These questions have been largely difficult to answer because they require sampling the responses of many neurons to a large number of stimuli across multiple behavioral states. Ultimately, to gain a complete understanding of the role correlated variability plays in state-dependent population coding, we will need to characterize neural population activity across the entire sensory response space in awake, behaving animals.

In this work, we take a step towards this goal by recording population activity in the primary auditory cortex of awake ferrets while presenting a diverse library of natural sound stimuli. Using our large stimulus set, we were able to show that the role correlated variability plays in population coding can change dramatically across the sensory space. Using established methods for measuring coding accuracy,[28,30–33,39] we found that within the same population of neurons, correlated variability sometimes interferes with coding accuracy and sometimes does not.

Next, we focused on how coding accuracy interacts with behavioral state changes in arousal measured via pupillometry.[18,40] We found that arousal increased first-order single neuron response gain and decreased second-order correlated variability between neurons. Both of these effects contributed to improved neural discriminability of natural sounds during high arousal states. However, they were able to be dissociated; whether or not a pair of neurons

showed arousal-associated modulation of correlated variability could not be predicted from their respective first-order effects. This result is in contrast to previous work on selective attention[22,41] in which attention was found to decrease correlated variability primarily by reducing fluctuations in the first-order attentional gain signal itself. This suggests that arousal may act through different neural circuits. Finally, we found that the arousal dependent reduction in correlated variability occurred via suppression of variability along a low dimensional, high variance axis of non-sensory related activity. The alignment between this axis and the sensory discrimination axis was predictive of improvements in population coding accuracy.

Results

We recorded simultaneous single- and multiunit activity from primary auditory cortex (A1) using single-shank 64-channel, or dual-shank 128-channel linear silicon probes ($n = 371$ SU and $n = 331$ MU, Fig. 1B).[42] Data were obtained from $n = 25$ recording sites in five awake, head-fixed ferrets. During each recording session, we presented a diverse set of randomly interleaved natural sound excerpts[43] in the acoustic field contralateral to the recording hemisphere (Fig. 1A). To monitor spontaneous fluctuations in arousal, pupil size was measured continuously during neural recordings using infrared video[18,40] (Fig. 1A, B).

In ferret A1, changes in pupil size are correlated with monotonic changes in neural firing rate.[18] Thus, to quantify pupil-related changes in activity, we split the neural data in half based on median pupil size (large pupil/high arousal vs. small pupil/low arousal). Qualitatively, we observed that transitions between high and low pupil-indexed arousal were accompanied by changes in both the excitability of individual neurons (first-order statistics) and the correlations between neurons (second-order statistics). During periods when pupil was large, responses to the same sound were generally stronger and more reliable than when pupil was small (Fig. 1D). Additionally, for large pupil trials population activity was desynchronized relative to the low arousal state; coordinated, stimulus-independent fluctuations in the population PSTH were primarily observed during small pupil trials (Fig. 1D).

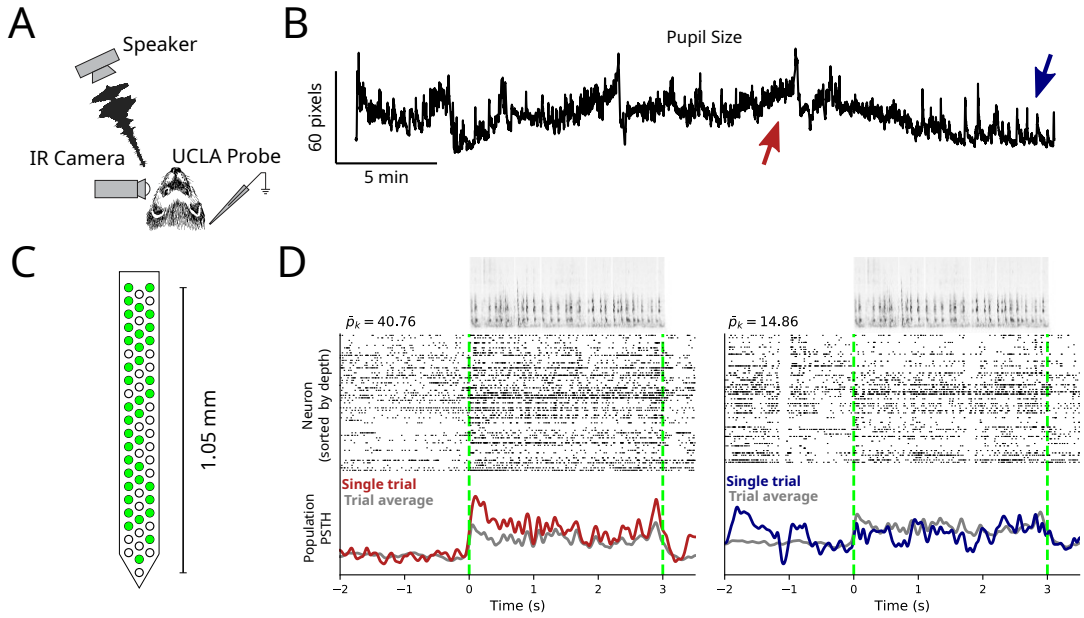

Figure 1. Pupil-indexed arousal modulates A1 population responses to natural sound stimuli. A. Single- and multi-unit activity was recorded from A1 of awake, head-fixed ferrets using laminar electrode arrays during presentation of natural sound stimuli. Pupil size, an index of arousal, was measured simultaneously using infrared video. **B.** Pupil trace from one recording session. Pupil size varied substantially over the course of the recording session, indicating spontaneous transitions between high and low arousal states. **C.** Schematic of 64-channel UCLA laminar probe used to record neural activity. Highlighted circles represent electrode channels on which at least one unit was detected during this recording session ($n = 55$ total units). **D.** Top: Spectrogram of one 3 s natural sound excerpt presented multiple times during the recording session. Frequencies shown are log-spaced (200 Hz - 20 kHz). Middle: Population raster plot of spiking activity by all simultaneously recorded units during a single stimulus presentation when pupil was large (left, red arrow in B) and when pupil was small (right, blue arrow in B). Bottom: Population peri-stimulus time histogram (PSTH) response, averaged across units during the single trial (red / blue) indicated in B and averaged over all repetitions of this stimulus (gray). $\bar{p}_k$ indicates mean pupil size on each respective trial, k.

Arousal increases neural response gain and decreases shared variability in A1

To quantify first-order effects of pupil on spiking activity, we fit a state-dependent generalized linear model (GLM) to each unit's response (Eqn. 1). For each recorded unit, the peri-stimulus time histogram (PSTH) response to each natural sound was averaged across repetitions and provided as input to the GLM. Time varying firing rate during each stimulus presentation was then fit by scaling the PSTH with a multiplicative gain and additive offset proportional to the size of pupil. To account for nonlinearity in the dynamic range of pupil-associated firing rate effects, the pupil trace was passed through a sigmoidal linking function

prior to scaling (Fig. 2A).

To measure model performance, we compared predictions to a purely sensory model with the same design but in which pupil size was shuffled over time. For 105 / 702 units, we saw significantly improved performance with the pupil-dependent vs. sensory-only model. At the population level, this effect was significant ($p \approx 0.0$, Bootstrap test, $n = 25$ recording sessions, 31.9 ± 5.17 units per session) indicating that pupil-indexed changes in arousal account for significant response variability in A1 . To investigate which aspects of the neural response, gain or baseline firing rate, were sensitive to pupil, we compared the gain and offset model fit parameters between large and small pupil conditions (Methods, Eqns. 2, 3). In all units, pupil-dependent gain was greater than zero ($\mu = 0.21 \pm 0.02, p \approx 0.0$, Bootstrap test, $n = 25$ recording sessions, 31.9 ± 5.17 units per session; Fig 2C). Pupil-dependent offset of firing rates was also significantly positive across the population ($\mu = 0.24 \pm 0.05, p \approx 0.0$ Bootstrap test; Fig 2C). These results are largely consistent with previous reports in ferret,[18] mouse,[16,21,40] and primate.[44]

To quantify second-order effects of pupil that modulated correlated activity within the neural population, we measured noise correlation between pairs of simultaneously recorded neurons. Mean noise correlation across neuron pairs is thought to indicate the strength of the dominant stimulus-independent mode of covariation in the population[45] and has been implicated in many studies as evidence of sensory information limiting noise.[19,29,31] On average, noise correlation was significantly reduced when arousal increased ($\rho_{small} = 0.047 \pm 0.005$ vs. $\rho_{large} = 0.043 \pm 0.005$, $p = 0.035$, Bootstrap test, $n = 25$ recording sessions, 318 ± 94 unit pairs per session; Figure 2D), consistent with previous reports.[16,18,21] The combined effects of increased gain and decreased noise correlation are consistent with the theory that increased arousal improves the signal-to-noise ratio of sensory responses in A1. Therefore, increased arousal may support accurate discrimination of naturalistic stimuli, a task critical for many behaviors, including speech perception by humans.[46]

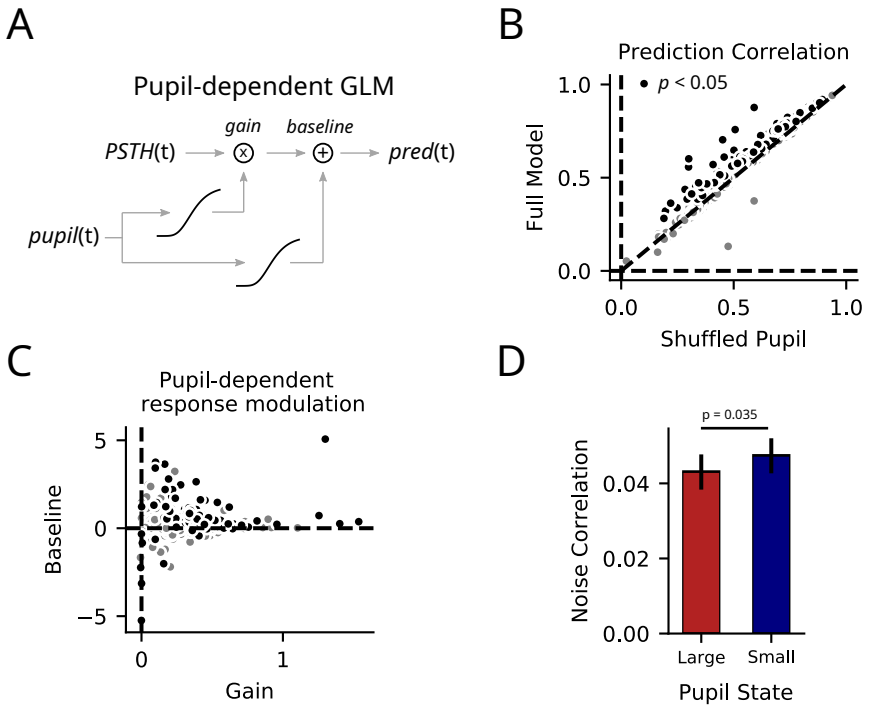

Figure 2. Pupil-indexed arousal increases response gain and decreases pairwise noise correlation in A1. A. GLM for pupil-dependent changes in neural activity. Trial-averaged PSTH response to each stimulus is scaled by a pupil-dependent multiplicative gain and additive baseline to predict the time-varying response on each trial. **B.** Performance of the GLM was evaluated by computing prediction correlation (Pearson's r) between predicted and the actual firing rate. For 105/702 recorded units, the full model (pupil-dependent GLM) significantly outperformed the sensory-only model (Jackknifed t-test, $p < 0.05$). Most other neurons trended toward an improvement, and the effect was significant across the population (full model mean $r = 0.65 \pm 0.02$, shuffled pupil $r = 0.63 \pm 0.02$, $p \approx 0.0$, Bootstrap test, $n = 25$ recording sessions, 31.9 ± 5.17 units per session). **C.** Mean change in gain or baseline between large and small pupil (Eqns. 2, 3). Both values were consistently greater than zero (gain modulation: $\mu = 0.21 \pm 0.02$, $p \approx 0.0$, DC modulation: $\mu = 0.24 \pm 0.05$, $p \approx 0.0$, Bootstrap test, $n = 25$ recording sessions, 31.9 ± 5.17 units per session), demonstrating that pupil increased both baseline and stimulus evoked firing rates across the population. **D.** Mean noise correlation for all simultaneously recorded pairs of units, grouped within recording session, then averaged across sessions ($n = 25$ recording sessions, 318 ± 94 unit-pairs per session), computed over trials when pupil was large versus trials when pupil was small. Error bars represent mean / standard error across recording sessions. Noise correlation was significantly reduced during large pupil trials ($\rho_{small} = 0.047 \pm 0.005$ vs. $\rho_{large} = 0.043 \pm 0.005$, $p = 0.035$, Bootstrap test).

The impact of correlated variability on population coding accuracy varies across sensory stimuli

Next, we wanted to better understand how arousal-associated changes impacted population coding accuracy in A1. A well-established measure of coding accuracy is neural discriminability, d'^2, between pairs of sensory stimuli.[28,30–33,39] The value of d'^2 depends not

only on the mean response evoked by each stimulus, but also on the degree to which trial to trial response variability interferes with sensory discrimination.[30,34,39] To measure d'^2 for each pair of stimuli, we first projected neural responses onto the two-dimensional plane that preserved maximal information about both of these features. This projection, which we call Targeted Dimensionality Reduction (*TDR*, Figure S2), was defined by two axes: the sensory discrimination axis (TDR_1 / $\Delta\boldsymbol{\mu}$) and the global noise axis (see Methods: Glossary). TDR_2 was the axis orthogonal to TDR_1 and in the *TDR* plane. *TDR* minimized over-fitting to single trial noise in d'^2 estimates (Figure S3) and facilitated visualization of the high dimensional neural population data (Figure 3C). Increasing the dimensionality of the *TDR* space by including additional noise components only increased overall d'^2 estimates marginally. In addition, the effect of arousal on d'^2, which is the main focus of this work, was not sensitive to the choice of dimensionality (Figure S5).

Using *TDR*, we measured d'^2 for each pair of sound stimuli (Figure S1) and found that discriminability was strongly stimulus dependent (Figure 3). To investigate the cause of this dependency, we characterized each stimulus pair by two metrics: discrimination axis magnitude and noise interference (Figure 3A). Discrimination axis magnitude referred to the vector magnitude of $\Delta\boldsymbol{\mu}$ and noise interference was defined as the cosine similarity between $\Delta\boldsymbol{\mu}$ and the correlated variability axis ($\mathbf{e}_1$), which was found by computing the first noise principal component in *TDR* space (Figure 3B). These metrics will be used throughout this work. To reference their definitions, see Methods: Glossary.

For pairs of stimuli with large discrimination magnitude and low noise interference, d'^2 was large (Fig. 3A, B.IV) and for pairs with small discrimination magnitude and high noise interference, d'^2 was small (Fig 3A, B.I). To quantify these dependencies, we regressed d'^2 against the per-stimulus noise interference and discrimination axis magnitude (Figure 4E). For all sites, noise interference coefficients were negative and discrimination magnitude coefficients were positive ($\beta_{noise} = -0.39 \pm 0.03, p = 0.000071, U = -3.97, n = 11$ recording sessions, Mann-Whitney U Test; $\beta_{discrimination} = 0.80 \pm 0.02, p = 0.000071, U = 3.97, n = 11$ recording sessions, Mann-Whitney U Test), indicating that this pattern was consistent across experiments. Because d'^2 is calculated using an optimal decoding approach, the optimal decoding axis, $\mathbf{w}_{opt}$, is not only dependent on $\Delta\boldsymbol{\mu}$, but also on the noise structure (compare Figure 3B bottom row and top row). Thus, the projection onto $\mathbf{w}_{opt}$ maximizes the separation between the two stimulus categories while minimizing noise. Therefore, we also included an interaction term in our model to account for this trade-off. This term was significant for all

sites tested ($\beta_{interaction} = -0.18 \pm 0.02, p < 0.000071, U = -3.97, n = 11$ recording sessions, Mann-Whitney U Test).

Our results illustrate that neural discriminability varies systematically across the sensory space. The variation of d'^2 with respect to noise interference emphasizes that correlated variability has very little impact on the ability to discriminate some stimulus pairs, but is critically important for others.

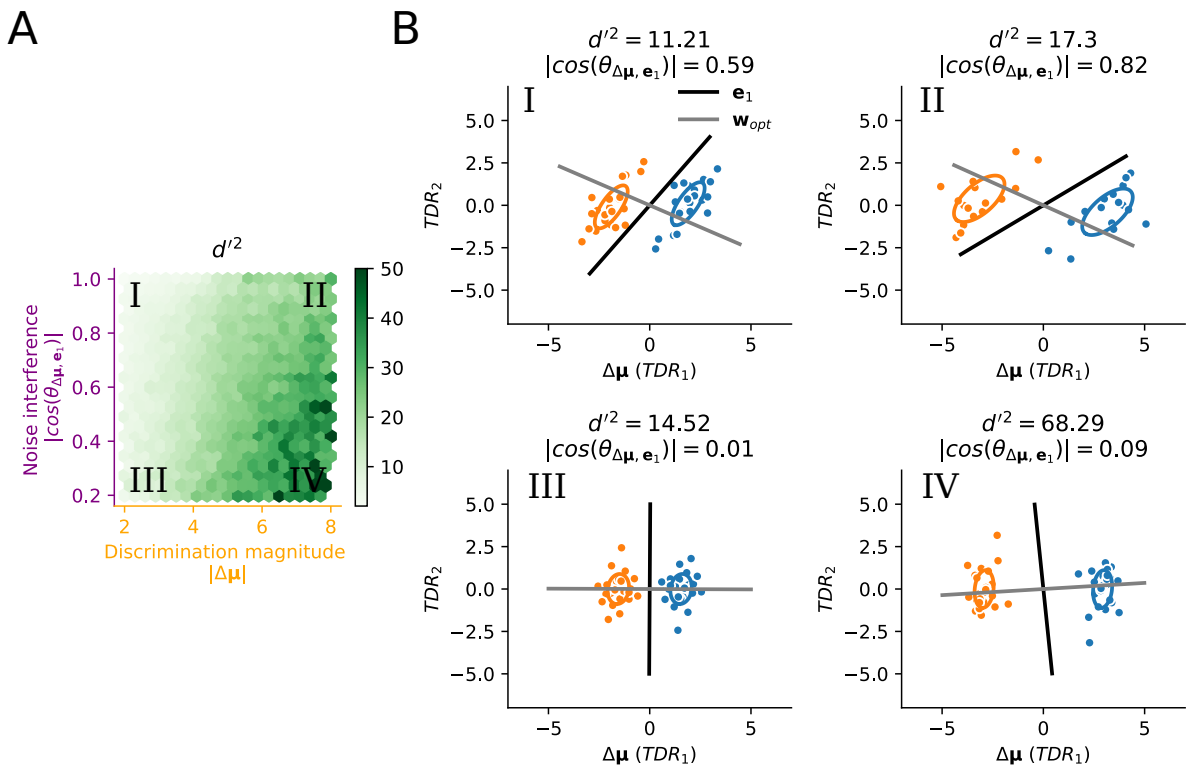

Figure 3. Natural stimulus discriminability in A1 populations. A. Heat map indicates neural discriminability, d'^2, averaged across stimulus pairs (pairs of 250 ms natural sound segments) and recording sites. Values were binned according to discrimination magnitude (x-axis) and noise interference (y-axis). See Figure S4 for the distribution of stimulus pairs across this space. **B**I-IV. Single-trial responses to example stimulus pairs in the reduced *TDR* space. Dots represent single trial responses and ellipses indicate standard deviation across trials. Example stimulus pairs shown were all collected from the same neural population during a single recording session. Each panel shows a pair of stimuli from a different quadrant (I - IV) in A. $\mathbf{e}_1$ represents the first eigenvector of the noise in the *TDR* space and $\mathbf{w}_{opt}$ represents the optimal decoding axis separating the two stimulus classes (blue vs. orange).

Arousal improves neural discrimination of natural sound stimuli

We used the framework for characterizing discriminability defined above to address our central question: Do arousal-dependent changes in gain and covariability (Figures 1-2) affect

population coding accuracy in A1? To answer this, we computed the decoding axis for each pair of stimuli across all repetitions in all pupil states ($\mathbf{w}_{opt}$, Eqn. 8). We then split the data from each stimulus pair in half based on the median value of pupil across sound repetitions and computed d'^2 separately for small and large pupil conditions. The use of a single decoding axis prevented biasing d'^2 toward being optimal for one or the other condition. Averaging across stimulus pairs, we found that discriminability of natural sounds was improved during high arousal states for most recording sites (Figure 4A). This change was significant across the population ($p = 0.02, W = 7, n = 11$ recording sites, Wilcoxon signed-rank test). Moreover, the magnitude of the effect was correlated with the amount of normalized pupil variance within an experiment ($r = 0.63, p = 0.029, n = 11$ recording sessions, permutation test), indicating that the change is indeed driven by pupil-indexed fluctuations in arousal (Figure S6).

We took advantage of our large stimulus set to further investigate if these changes in discriminability were the same across all pairs of stimuli or, alternatively, if the changes were stimulus dependent, like overall discriminability which varied greatly between stimulus pairs (Figure 3). To compare the relative magnitude of arousal effects across all pairs, we defined a metric to capture the fraction change in discriminability, $\Delta d'^2$ (Eqn. 11). For each stimulus pair, $\Delta d'^2$ was defined as d'^2_{large} minus d'^2_{small} divided by the sum of these two quantities. Thus, $\Delta d'^2$ ranged from -1 to 1 and could be directly compared across both stimulus pairs and recording sites.

For each recording site, we binned $\Delta d'^2$ according to the discrimination axis magnitude and noise interference (as in Figure 3) and averaged across sites (Fig. 4B). For stimulus pairs with high noise interference, where the discrimination and noise axes were aligned, discriminability in both high and low arousal states was low (Fig. 4C, left). In these cases, increased arousal led to a greater relative improvement in discriminability (Fig. 4D, left). Overall discriminability was also dependent on the discrimination axis magnitude (Fig. 4C, right), but arousal-dependent changes did not show a clear pattern (Fig. 4D, right). We quantified the axis-dependent changes in $\Delta d'^2$ using linear regression. We observed a positive linear relationship between $\Delta d'^2$ and noise interference across in 8/11 recording sites ($p < 0.05$, t-test), and this trend was significant across the entire dataset ($\beta_{noise} = 0.102 \pm 0.028, p = 0.001, U = 3.25, n = 25$ recording sessions, Mann-Whitney U Test; Figure 4F). We did not observe a consistent relationship between discrimination magnitude and $\Delta d'^2$ ($\beta_{discrimination} = 0.004 \pm 0.030, p = 0.718, U = 0.36, n = 11$ recording sessions, Mann-Whitney U Test; Figure 4F).

Collectively, these results demonstrate that arousal-associated improvements in stimulus

discriminability are largest where correlated variability aligns with the discrimination axis ($\Delta\boldsymbol{\mu}$). We propose two possible explanations for this result:

(1) Previous work suggests that trial by trial fluctuations in first-order gain are a major source of correlated variability in neural populations.[22,37,41,47,48] It is possible that such fluctuations are aligned with the principal axis of covariability identified during *TDR*. If so, improvements in d'^2 would be largest where noise interference is high, due to increased gain helping to separate stimuli along the discrimination axis.

(2) Alternatively, our results could be explained by a second-order decrease in variance along the correlated variability axis when arousal is high. Distinct from increased first-order gain magnitude, this could be the result of an arousal-dependent reduction in the variability of shared gain across the population, as has been suggested previously.[22,37,41] Alternatively, it could be due to modulation of an additional noise source, for example, afferent sensory noise or top-down modulatory input.

Although either first- or second-order effects could explain our results, it is also possible that both contribute to changes in discriminability. Below we focus on which of these effects contribute to the discriminability changes.

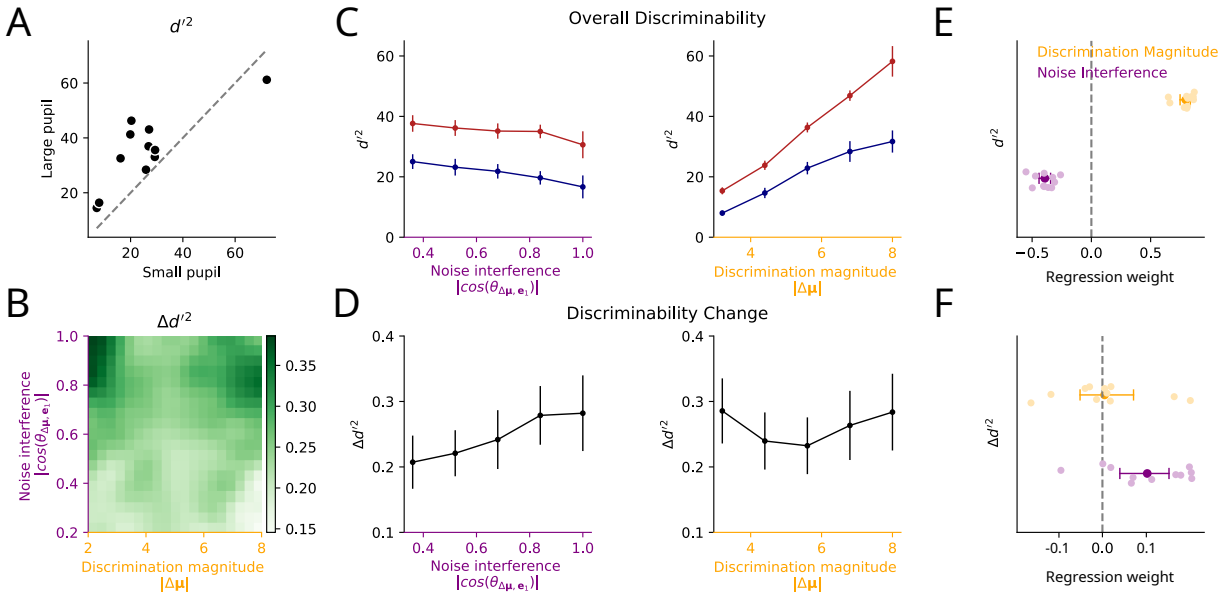

Figure 4. Arousal improves discrimination of natural sounds. A. Scatter plot compares mean discriminability (d'^2) across natural sound pairs for small (low arousal) versus large (high arousal) pupil trials. Each point indicates the mean across all stimuli for a single recording session. Mean discriminability was nearly always greater during large pupil trials ($p = 0.02, W = 7, n = 11$ recording sessions, Wilcoxon signed-rank test). **B.** Heatmap shows the relative change in discriminability for large vs. small pupil ($\Delta d'^2$) as a function of discrimination axis magnitude ($|\Delta\boldsymbol{\mu}|$) and noise interference ($|cos(\theta_{\Delta\boldsymbol{\mu},\mathbf{e}_1})|$. Results are smoothed with a Gaussian filter of width two bins. **C.** d'^2 for large and small pupil trials. d'^2 is grouped into five evenly spaced bins by noise interference (left) or discrimination magnitude (right). Points / error bars indicate the mean / standard error of d'^2 across recording sessions. **D.** Arousal-related change in discriminability ($\Delta d'^2$) for each session, plotted as in C. **E** Each point represents the regression coefficient per recording session for linear predictions of d'^2 across all pupil conditions. Error bars show bootstrapped 95% confidence interval across sessions. $\beta_{Noise} = -0.39 \pm 0.03, p = 0.000071, U = -3.97, \beta_{Discrimination} = 0.80 \pm 0.02, p = 0.000071, U = 3.97, n = 11$ recording sessions, Mann-Whitney U test. **F.** Same as in (E) for predictions of $\Delta d'^2$. Error bars show bootstrapped 95% confidence interval across sessions. $\beta_{Noise} = 0.102 \pm 0.029, p = 0.001, U = 3.25, \beta_{Discrimination} = 0.004 \pm 0.030, p = 0.718, U = 0.36, n = 11$ recording sessions, Mann-Whitney U test.

Arousal improves discriminability through a combination of increased response gain and reduced correlated variability

The changes we observed in either first- or second-order response statistics (Fig. 2) could contribute to enhanced stimulus discriminability in the large pupil condition. To dissociate these possibilities, we adapted the approach used by Cohen et al., 2009[19] to simulate two datasets. In the first, we isolated first-order arousal effects by fixing the covariance of the simulated data between large and small pupil trials but matching the mean and variance of each unit to the values measured in large/small pupil conditions (1st-order simulation). In the second, we

matched both the first- and second-order statistics to those measured in the raw data (1st + 2nd-order simulation). For both simulations, the spike rate response on each trial was drawn from a Gaussian distribution (Eqn. 12).

For the raw data and for each simulated data set, we measured the mean arousal-dependent change in discriminability across all stimulus pairs for each site (Figure 5A). First-order statistics alone accounted for consistent arousal-dependent improvements in decoding for all recording sites. The magnitude of improvement varied across stimulus pairs and was particularly strong for stimulus pairs with high noise interference and low discrimination axis magnitude (Figure 5B, regression coefficients: $\beta_{noise} = 0.0139 \pm 0.025, p = 0.001, U = 3.25$; $\beta_{discrimination} = -0.072 \pm 0.044, p = 0.07, W = -1.81, n = 11$ recording sessions, Mann-Whitney U Test). Adding arousal-dependent second-order statistics yielded an additional increase in $\Delta d'^2$ for some sites, while for others it had minimal or even negative effects (Figure 5A, blue vs. red). This result confirms the theoretical prediction that the impact of changes in correlated variability can vary greatly, depending on tuning of the population for the stimulus pair (Figures 3, 4). Indeed, we found that unique second-order effects (measured by subtracting $\Delta d'^2$ for the first-order simulation from $\Delta d'^2$ for the full simulation) were mostly observed for stimulus pairs with high noise interference and discrimination magnitude (Figure 5C), though these trends were not significant ($\beta_{noise} = 0.05 \pm 0.03, p = 0.27, U = 1.08$; $\beta_{discrimination} = 0.05 \pm 0.02, p = 0.28, U = 1.08, n = 11$ recording sessions, Mann-Whitney U Test). These results demonstrate that arousal-associated modulation of both first- and second-order statistics can contribute to changes in stimulus discriminability and that they may operate in complementary regions of stimulus space. In addition, we observed that mean $\Delta d'^2$ was not significantly different between the raw data and the 1st + 2nd order simulation ($p = 0.29, W = 21, n = 11$ recording sessions, Wilcoxon signed-rank test), suggesting that response statistics beyond second-order do not contribute to the arousal-dependent changes in discriminability that we measured.

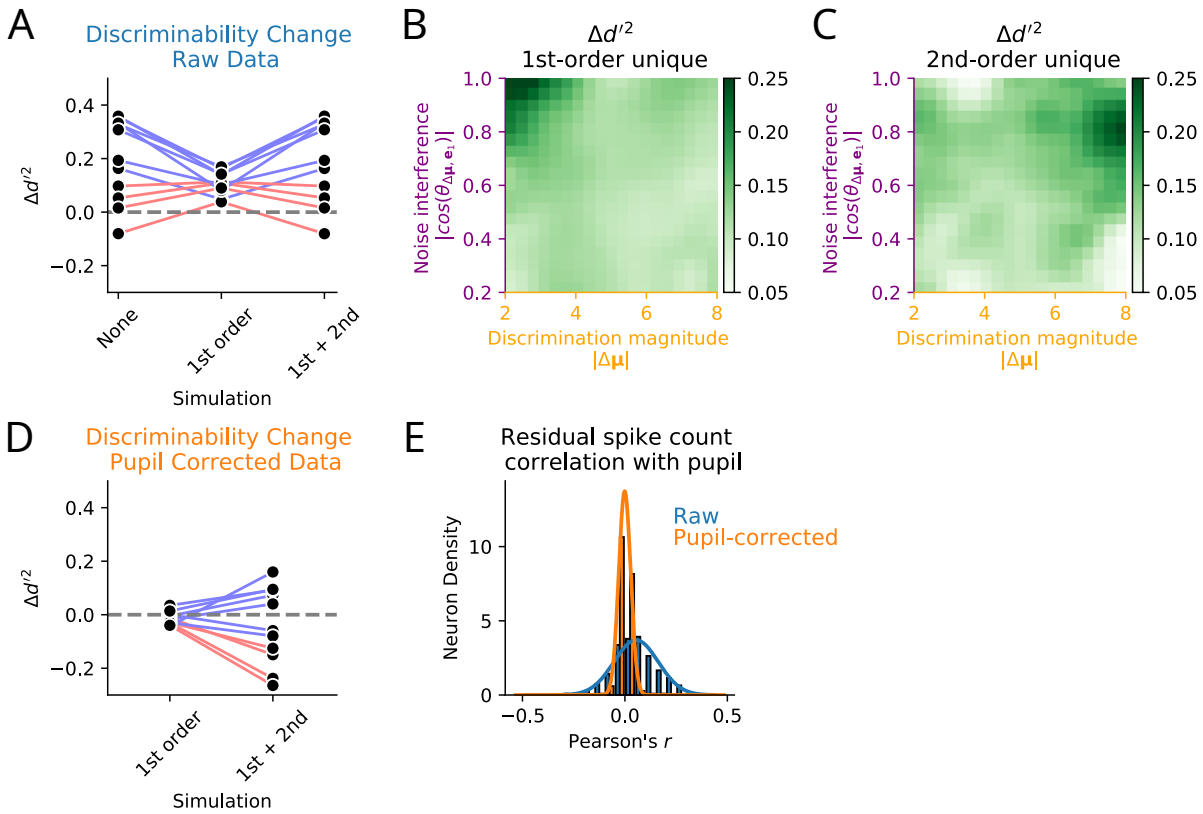

Figure 5. Arousal-dependent changes in first- and second-order response statistics modulate stimulus discriminability. A. Mean change in discriminability, $\Delta d'^2$, for each recording session for the raw data (simulation = None) and for two data sets with simulated pupil-related changes (1st order and 1st + 2nd order). Sessions that trended towards arousal-associated improvements in d'^2 due to second-order effects are indicated in blue for visualization (1st + 2nd > 1st order), others in red. **B.** Heatmap of $\Delta d'^2$ for first-order simulations, binned as in Figures 3-4. **C.** Same as in B for unique 2nd order changes (1st + 2nd order simulation minus 1st order simulation). **D.** Results for the two simulated data sets, as in A, but after first regressing out pupil-dependent firing rate modulations. Sessions are colored according to A. **E.** Pearson's correlation between each unit's residual response (raw response minus stimulus average (blue) or raw response minus GLM model prediction (orange)) and pupil size. Pupil-correction reduced the significant mean unit-correlation with pupil from 0.05 ± 0.01 ($p \approx 0.0$) to -0.0001 ± 0.0022 ($p = 0.67$, Bootstrap test, $n = 11$ recording sessions, 39.45 ± 3.99 units per session). $p-$values correspond to test of if the distribution mean differs from zero.

Arousal-dependent changes in gain and correlated variability operate on distinct timescales

Previous work has shown that selective attention, like arousal, can act to both increase response gain, a first-order statistic, and reduce the magnitude of correlated variability, a second-order statistic.[19] Further work has demonstrated that these reductions in correlated variability are due to a suppression of trial to trial variability in coherent first-order modulation

of single neuron evoked responses.[22,41] These findings suggest that a single, shared circuit mediates both first- and second-order effects of selective attention. We wondered if the same was true for arousal.

If both first- and second-order effects can be explained by pupil-dependent modulation of gain, we reasoned that removing all pupil-explainable variance from the spiking data should eliminate the changes in discriminability. To test this hypothesis, we generated a pupil-corrected data set in which the pupil-dependent GLM prediction (Fig. 2) for each unit was subtracted from its raw response. Next, we computed stimulus discriminability for the pupil-corrected data, using the same procedure as for the raw data above. We found that for the first-order simulation, $\Delta d'^2$ improvements were abolished (Figure 5B), confirming that there was no longer any first-order relationship between pupil size and firing rate (Figure 5E). However, we observed a similar pattern of second-order contributions to $\Delta d'^2$ as for the raw data (compare Figure 5A to 5D). This result suggests that second-order effects cannot be explained simply by coherent first-order effects across neurons. If this were not the case, we would have expected $\Delta d'^2$ for the 1st order simulation and the 1st + 2nd order simulation to be identical in the pupil-corrected data. These results argue that arousal does not modulate second-order correlated variability by reducing shared fluctuations in arousal-dependent gain, as was observed for selective attention in previous studies.[22,41] Instead, they indicate that slow changes in arousal suppress an independent source of correlated variability.

To determine if first- and second-order effects could be further dissociated, we used a bandpass filter to partition the data into temporal frequency bands prior to noise correlation measurements (Fig. 6A, B). This allowed us to identify the timescale over which the second-order effects operate. We found that noise correlation was largest overall in the low-frequency band (< 0.5 Hz), reflecting slow, coordinated fluctuations in the population activity.[38] Pupil-indexed fluctuations in arousal occur at this slow timescale. Thus one would expect effects of coherent first-order modulation to be restricted to this band, and, indeed, comparison of noise correlation before and after pupil correction showed a reduction of noise correlation in this slow temporal band ($p = 0.0003, W = 29, n = 25$ recording sessions, Wilcoxon signed-rank test, Fig. 6A). However, changes in correlation magnitude between high and low arousal states were restricted to higher frequency bands ranging between 0.5 and 25 Hz (Figure 6B, $p = 0.019, 0.002, 0.007, 0.00005, W = 76, 51, 63, 11, n = 25$ recording sessions, Wilcoxon signed-rank test). Therefore, we conclude that the arousal-dependent changes in correlated variability are due to modulation of a process that operates on a distinct, faster timescale than

the change in arousal itself.

If changes in correlated variability are in fact distinct from first-order effects, the two processes may also operate on different populations of neurons. To determine how these effects overlapped in the neural population, we measured the relationship of pairwise noise correlations to the magnitude of first-order effects in each cell. First-order effects were quantified by the pupil-modulation index (MI) for each unit (Eqn. 4). A value of MI= 1 indicated a unit that only responded when pupil was large, and MI= 0 indicated no change between large and small pupil. While MI was able to predict overall noise correlations above chance, it contained no information about the arousal-dependent reduction in noise correlations (Figure 6C, D). Thus, in addition to acting on distinct timescales, first- and second-order effects of changing arousal operate on non-overlapping populations of neurons in A1.

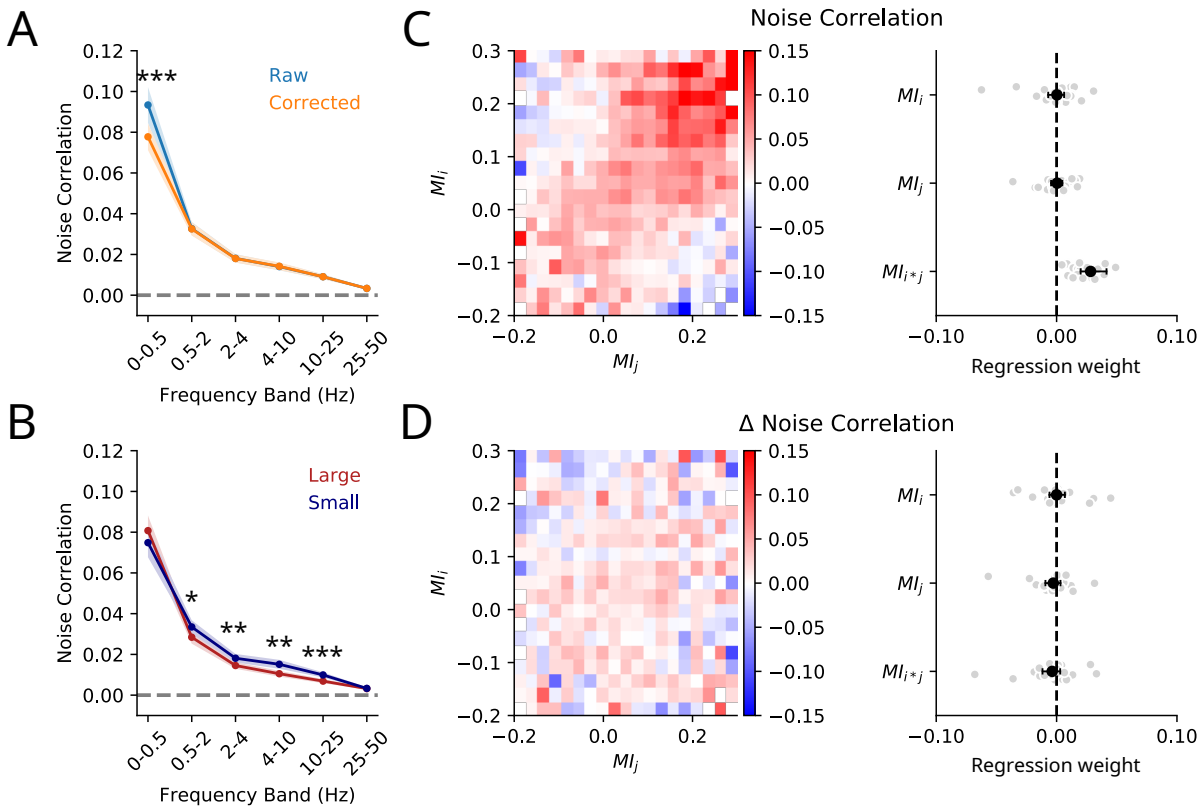

Figure 6. First- and second order arousal effects operate on distinct timescales and neural populations. A. Noise correlations computed after filtering residual spike counts (raw and pupil-corrected) into six non-overlapping temporal frequency bands. To minimize bias from one recording session, noise correlation was first averaged across neuron pairs within session before measuring the significance of the difference between conditions (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, Wilcoxon signed-rank test, $n = 25$ recording sessions). Shading indicates standard error across recording sessions. **B.** Pupil-corrected noise correlations measured after splitting data by median pupil size. Statistics as in A. **C.** Left: Pairwise noise correlations (spikes binned at 4Hz as in Figure 2), grouped into 20 evenly spaced bins as a function of modulation index (MI) for each pair of units (i, j). Right: Regression coefficients for predicting overall noise correlation from MI of individual neurons (MI_i and MI_j) and their interaction (MI_{i*j}). Dots represent single recording sessions; error bars represent 95% confidence intervals over parameter values. **D.** Same as C but for change in noise correlation between large and small pupil.

Arousal-dependent reduction of correlated variability is low dimensional

Finally, we wondered how closely arousal-dependent changes in noise correlation predicted changes in stimulus discriminability. We hypothesized that if arousal-dependent changes are restricted to a limited subspace of the population activity, then their impact on discriminability will depend on how well that subspace aligns with the discrimination axis. To test this, we measured a primary axis of modulation, defined as the projection of population activity onto

the axis where noise correlation changed most between large and small pupil (Figure 7A-D). We then reasoned that the alignment of the sensory discrimination axis ($\Delta\boldsymbol{\mu}$) and modulation axis should predict the magnitude of arousal-dependent changes in discriminability uniquely attributable to second-order effects (Figure 5C). Indeed, we found that this was the case. For pairs of stimuli where the cosine similarity between $\Delta\boldsymbol{\mu}$ and modulation axis was high, we observed that second-order effects were larger (Figure 7G, $p = 0.021, W = 3, n = 9$, Wilcoxon signed-rank test). Furthermore, for most sessions, we found that the modulation axis was low-dimensional (Figure 7D); in 9/11 recording sessions, we found only one significant dimension along which noise correlation changed between large and small pupil ($p < 0.05$, permutation test, Methods). This dimension was consistently similar to the first principal component of the pooled noise data (7E, F), suggesting that arousal modulates correlated variability along high variance dimensions in the data.

Based on these results, we conclude that arousal-dependent changes in correlated variability are restricted to a relatively low-dimensional subspace and that these changes contribute to improved sound discriminability when the discrimination and noise correlation modulation axes are aligned.

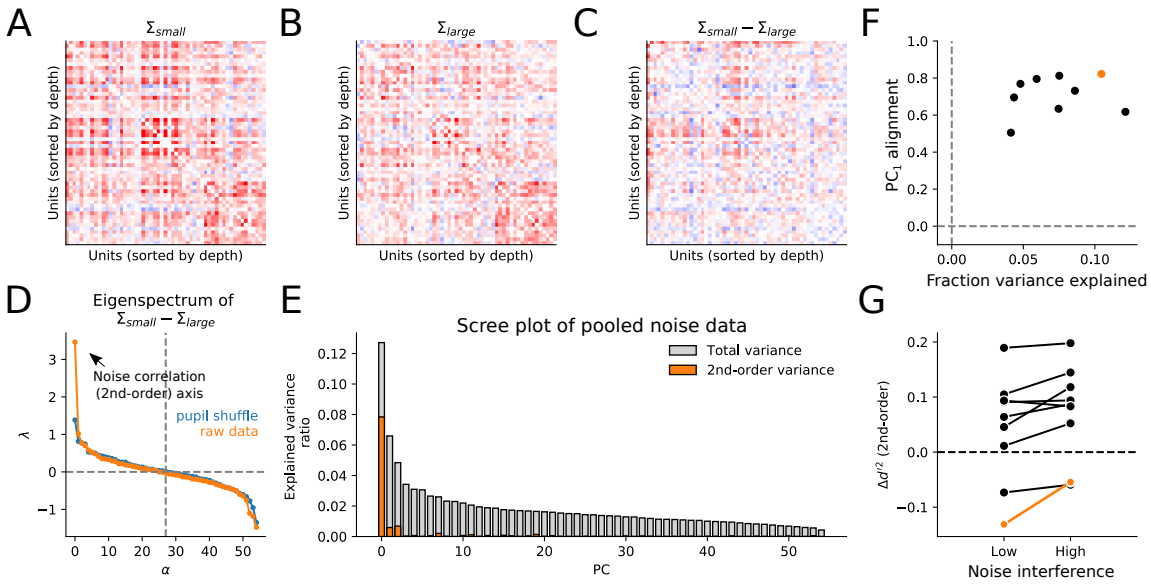

Figure 7. Alignment of pupil-related changes in noise correlation and stimulus discriminability axes. A. Covariance matrix of z-scored spike counts for small-pupil states during one recording session. **B.** Same as in A for large pupil state. **C.** Difference between covariance matrices in panels A and B. Colormap scale shared between panels A-C. **D.** Eigenspectrum of the difference matrix in C (orange). The first eigenvalue corresponds to the noise correlation modulation axis. The blue line shows the noise floor for the eigenspectrum, computed with pupil size shuffled prior to classifying large/small pupil trials. Only sites with at least one significant eigenvalue were included subsequently in panels F/G (9/11 sites, $p < 0.05$, permutation test). **E.** Scree plot shows ratio of total variance explained along each principal component of the pooled noise data (grey) for the example site in A-D. The fraction of variance along each PC that can be accounted for by the modulation axis is overlaid in orange *e.g.* If the modulation axis were perfectly aligned with the first PC, the first bar would be colored entirely orange. **F.** Scatter plot compares fraction of total variance in the data explained by the modulation axis (x-axis) versus the ratio of modulation axis variance to PC_1 variance (y-axis) *i.e.* The ratio between the height of the orange and grey bar for PC_1 in panel E. Each point represents one site. Example site in A-E indicated in orange. **G.** Changes in $\Delta d'^2$ uniquely attributable to second-order effects of arousal, grouped by cosine similarity between the noise correlation and discrimination axes per site (low vs. high noise interference: median split). Stimulus pairs with high cosine similarity showed significantly greater discrimination improvements due to 2nd-order changes ($\mu_{high} = 0.074 \pm 0.032, \mu_{low} = 0.043 \pm 0.028, p = 0.02, n = 9$, Wilcoxon signed-rank test).

Discussion

Previous studies of how internal behavioral state influences sensory coding by neural populations have primarily focused on small, low-dimensional stimulus sets. It is not clear if results from this work can be extended to coding of more diverse, natural stimuli. Theoretical studies have shown that the impact of population statistics, such as correlated variability, on coding accuracy can vary dramatically depending on how the population is tuned to the

sensory stimulus.[39] Consistent with these predictions, we found that the influence of arousal on discriminability of natural sounds by neural populations in primary auditory cortex (A1) varied substantially with sensory tuning. This result identified a strong dependence of discriminability on the geometric relationship between sound evoked activity and two distinct, low dimensional modulatory signals within the neural population.

Effects of shared intrinsic variability on discriminability are stimulus-dependent

Correlated, intrinsic variability within neural populations is ubiquitous in cortex. However, even before this phenomenon was observed experimentally, there was a substantial theoretical effort to understand how such activity might affect coding by neural populations.[23–30] This early work established that correlated variability can interfere substantially with the brain's ability to accurately discriminate sensory stimuli. Therefore, experimental characterization of this phenomenon is critical to fully understand neural population codes.

Although evidence for intrinsic correlation is widespread, experimental studies have provided conflicting evidence as to whether or not it does in fact interfere with population coding.[31–37,48] There are at least two reasons why effects of correlated variability might vary across studies. First, the dimensions containing interfering noise could be very low variance. In this case, detecting these correlations would require recording large amounts of data, both over many neurons and over many trials, a methodology that has only recently become feasible.[31–33] A second possibility is that the impact of correlated noise depends on the tuning of the neurons in the population that is being read out, and its relationship with the noise space.[34,48] Thus, discrepancies in previous work might be explained by differences in sampling of neural populations or in the stimuli tested. Because the alignment of sensory signal and intrinsic noise may depend on the stimuli that are presented, it is important to characterize coding accuracy across diverse sets of stimuli. Our results showed that, indeed, effects of correlated variability are highly dependent on the stimuli tested.

Questions about stimulus-dependent changes in discriminability are difficult to address in a single study, as there is a trade-off between the number of stimuli that can be presented and the number of times that each can be repeated during a single recording session. In contrast to recent work,[31–33] we presented a diverse set of stimuli over a relatively small number of repeats and neurons. Thus, our ability to measure low-variance dimensions in our population

data was limited and we were unable to draw strict conclusions about the presence (or absence) of information limiting noise. However, developing a novel dimensionality reduction approach (*TDR*) allowed us to reliably estimate the interaction between the highest-variance noise dimension and the sensory discrimination axis across a large portion of the sensory response space, something that has not yet been studied in detail.

Separate circuitry drives first- and second-order neural variability

Recent studies of selective attention have suggested that fluctuations in second-order, correlated variability result from the coherent modulation of first-order, evoked response rates of many neurons by an intrinsic behavioral state variable.[22,37,41] For example, in a recent study by Denfield et al., macaques were trained on a cued change-detection task in which the stability of spatial selective attention was manipulated between behavioral blocks.[22] Because the gain of evoked responses is known to be modulated by attention,[49] the authors proposed that the magnitude of correlated variability should be highest when attention itself was most variable. Indeed, when animals were forced to switch attention between multiple locations within a behavioral block, noise correlations were strongest.[22] This is presumably because the spotlight of attention must alternate between receptive field locations to maximize reward. These results are consistent with work by Rabinowitz et al. who demonstrated that the variability of coherent first-order gain modulation by attention is reduced during attend-in relative to attend-out conditions.[41] Collectively, these results offer a parsimonious explanation for why pairwise noise correlations are reduced during traditional cued change-detection tasks, where attention is focused stably on a single spatial location.[19]

Unlike the case of visual selective attention, we found that arousal-dependent first- and second-order effects in auditory cortex could be fully dissociated, and thus did not arise from a common intrinsic signal. First-order modulations were slow, on the order of many seconds. We suggest these may be analogous to previously described slow-drift signals in cortex.[41,50] Second-order correlated variability, however, was relatively fast, operating on timescales of tens to hundreds of milliseconds. We could not predict second-order arousal effects from first-order effects; it was not more likely for two cells showing large first-order modulation to have their correlation be modulated by arousal. Given these findings, we conclude that, unlike the case of visual selective attention, first- and second-order effects of arousal are mediated by distinct neural circuits in A1. These effects could be described with low-dimensional modulatory

signals, each operating on a different A1 population and on a different timescale.

Although isolating the circuitry underlying these effects is not feasible here, we propose a few non mutually exclusive possibilities that may mediate first- and second-order effects. One option is that the effects are inherited in a bottom-up fashion. First-order response modulation by arousal has been observed in multiple stages of the auditory pathway, even before auditory cortex.[51,52] Further, stochastic variability in the ascending auditory pathway may cause noise correlation in neurons that share ascending input. However, this seems unlikely as this type of noise has been reported to be very weak in cortex.[31–33] Alternatively, neuromodulatory signals originating in subcortical regions may be involved. Slow changes in pupil-indexed arousal have been tied to cholinergic and adrenergic signaling in auditory cortex[53] and stimulation of the basal forebrain is known to enhance evoked response rates and reduce correlated variability in V1.[54] A third possibility is that the effects arise due to intracortical communication. Inhibitory interneuron populations in A1 are known to receive input from motor cortex.[55] Because motor activity itself may coincide with increases in arousal,[21] it is possible that increased inhibition via this intracortical pathway is responsible for decorrelation during high arousal states. Simultaneous monitoring of motor activity and pupil size would be needed to test this hypothesis more rigorously.

State-dependent coding in auditory cortex

It is increasingly clear that neural activity in primary sensory regions of the brain is significantly modulated by non-sensory variables, including arousal.[16,18,21,38,56] Arousal here refers to spontaneous changes in alertness as measured by pupil diameter, even in the absence of a behavioral task.[56] Similar to previous work, we find that increased arousal is associated with enhanced excitability and reduced correlated variability in A1.[16,18,40] These effects enhance the neural signal to noise ratio in V1[21] and improve population coding accuracy of tonal stimuli in A1.[16]

Building on previous work, we explored the effects of arousal more completely by measuring the changes in population coding accuracy across a large space of natural sounds. Consistent with previous work, pupil-indexed increases in arousal led to improved discriminability between sounds on average. However, the relative magnitude of this improvement varied substantially across stimuli. Improvements were largest when intrinsic correlated variability

interfered with sensory discrimination. Thus, a general suppression of correlated variability and increase in evoked response rates during high arousal states may gate behavioral states that require improved perceptual discrimination, like selective attention.

These results highlight the importance of a systematic exploration of arousal effects across the sensory response space. In future work, parametric stimuli might be used to more systematically probe sound feature representations across behaviorally relevant stimuli. For example, one existing study of auditory processing began to address this question in anesthetized animals.[17] In this work, Kobak et al. measured A1 population coding of sounds that varied in inter-aural level (ILD) and absolute binaural level (ABL). By inducing different states of cortical activation with urethane anaesthesia, the authors demonstrated that in the active (desynchronized) state, noise and signal subspaces shift to become approximately orthogonal, thereby facilitating accurate encoding across both ILD and ABL. Extending this approach to spectro-temporally variable, naturalistic stimuli that vary along behaviorally relevant dimensions[57] will be critical for understanding state-dependent population coding of natural sounds.

Materials and Methods

Surgical procedure:

All procedures were performed in accordance with the Oregon Health and Science University Institutional Animal Care and Use Committee (IACUC) and conform to standards of the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC). All animal care and procedures were similar to those described previously.[8,18,58,59] Five young adult male ferrets were acquired from an animal supplier (Marshall Farms). Head-post implantation surgeries were then performed in order to permit head-fixation during neurophysiology recordings. Stainless steel head-posts were fixed to the animal using Charisma composite or bone cement (Palacos), which bonded to the skull and also to stainless steel screws that were inserted into the skull. After a two-week recovery period, animals were habituated to a head-fixed posture and auditory stimulation. At this point, a small (1 - 2 mm) craniotomy was opened above primary auditory cortex for neurophysiology recordings.

Acoustic stimuli:

Digital acoustic signals were transformed to analog (National Instruments), amplified (Crown), and delivered through a free-field speaker (Manger) placed 1 m from the animal's head and 30°contralateral to the the hemisphere in which neural activity was being recorded (Fig. 1). Stimulation was controlled using custom MATLAB software (`https://bitbucket.org/lbhb/baphy`) and all experiments took place inside a custom double-walled sound-isolating chamber (Professional Model, Gretch-Kenn).

In the present study, we measured neural responses to natural sound stimulation. Some of this data has been previously published.[18,60] The full sound library can be accessed at `https://bitbucket.org/lbhb/baphy`). A non-exhaustive list of sound stimuli included in the natural stimulus sets (set 1: 93 sounds, 3 sec/sample, set 2: 306 sounds, 4 sec/sample with 1 sec ISI or 1 sec/sample with 0.5 sec ISI in a subset of experiments) were conspecific and heterospecific vocalizations, speech, music, and environmental sounds which were developed to characterize natural sound statistics.[43] All sound stimuli were presented at 65 dB SPL. During every experimental session, a subset of sound samples were repeated at least ten times, while the remainder of sound samples were played only once. In order to study the trial-to-trial variability in neural responses, only sounds repeated at least ten times were included in this manuscript. The order in which stimuli were presented was generated pseudo-randomly. Following this structure, stimuli were played continuously until all sound samples in the library had been presented. This meant that experiments lasted approximately 40 minutes.

Neurophysiology:

After opening a small craniotomy (1 - 2 mm) over auditory cortex, 1 - 4 tungsten micro-electrodes (FHC, 1-5 MΩ) were inserted to characterize the tuning and response latency of the region of cortex. Primary auditory cortex (A1) was identified by characteristic short latency responses, frequency selectivity, and tonotopic gradients of single or multiunit activity across multiple penetrations.[61,62] Once A1 was located, subsequent penetrations were made with a single (64-channel) or dual shank (128-channel) silicon probe[42] to allow recording from multiple units simultaneously. On each consecutive recording day, we shifted the location of the electrode penetration slightly to access fresh cortical tissue, expanding the craniotomy if necessary. Data were digitized at 30 KHz, amplified, and band-pass filtered between 0.3 and 3000 Hz (Intan Technologies, Open Ephys) and saved to disk for further analysis.

Electrode contacts on our laminar probes were spaced 20 μm horizontally and 25 μm vertically, collectively spanning 1.05 mm of cortex. Data recorded were spike sorted offline using Kilosort[63] or Kilosort2, the latter of which includes modifications to Kilosort including correction for electrode drift during recording sessions (code available: `https://github.com/MouseLand/Kilosort2`). Spike sorting results were manually curated in phy (`https://github.com/cortex-lab/phy`). For all sorted and curated spike clusters, a contamination percentage was computed by measuring the cluster isolation in feature space. All sorted units with contamination percentage less than or equal to 5 percent were classified as single-unit activity. All other units manually curated and saved in phy but not meeting this strict isolation criterion were labeled as multi-unit activity.

Pupillometry:

During neurophysiological recordings, video of the ipsilateral pupil (relative to the recording hemisphere) was collected using an open source camera (Adafruit TTL Serial Camera) fitted with a lens (M12 Lenses PT-2514BMP 25.0 mm) whose focal length allowed placement of camera 10 cm from the ferret's eye. Contrast was increased using infrared illumination. Ambient light levels were fixed for each experiment at roughly 1500 lux to provide maximum dynamic range of pupil size.[18] Pupil size was measured offline by fitting an ellipse to each video frame using using a custom machine learning algorithm written in Python (version 3.6.8). The minor axis of the fit ellipse was extracted and saved for further analysis with neurophysiological data. Blinks were detected and excluded as in[18] and pupil data was shifted by 750 ms relative to spike times in order to account for the lagged relationship between changes in pupil size and neural activity in auditory cortex and to allow for comparison with previous research.[40]

Because arousal is an uncontrolled variable in our experiments, it is likely that the same arousal states were not sampled from day to day, between different animals, or between different stimulus pairs within the same experimental session. As an attempt to control for this variability, we created a normalized metric of pupil variance for each stimulus, and then computed the mean of this variance metric for each recording session. For each stimulus pair, trials were split in half based on median pupil size across all trials. Next, the difference of mean large pupil size and mean small pupil size was normalized by the standard deviation of pupil across the entire experiment. Therefore, for a stimulus pair that sampled a large range of pupil states (relative to pupil during that particular recording session) this number was high. Finally,

we averaged this metric across all stimulus pairs at a recording site and compared it with the magnitude of pupil-dependent changes in neural discriminability for each recording session (Figure S6).

Pupil tracking algorithm:

We developed an automatic method for pupil size detection using a deep learning approach. Our model architecture was based on DenseNet201,[64] which is available through Keras (`https://keras.io/`). In order to transform the output of the model to pupil ellipse predictions, we added a single global pooling layer and a final prediction layer in which five pupil ellipse parameters (x-position, y-position, minor axis, major axis, and rotation) were fit to each video frame. In order to initialize model weights, the model was pre-trained on ImageNet,[65] then fine-tuned using roughly 500 previously analyzed, nonconsecutive frames from video of the pupil of multiple different ferrets (data from[18]). Qualitatively, after this first round of training the model performed well, even on novel video frames of pupil from new animals. However, in cases where the pupil video quality was poor, or differed substantially from the video frames in our training data set, we noticed failures in the model predictions. To further improve the model, we employed an active learning procedure. For each new analyzed video, pupil ellipse fits were analyzed qualitatively by experimenters. If the fit quality was deemed poor, predictions for these frames were manually corrected and added to the training data set. The model was then retrained and the analysis rerun. The network became robust to varying levels of video quality and performed consistently without the need for user intervention. The code for this analysis is available at `https://github.com/LBHB/nems_db`.

Pupil-dependent GLM:

In order to characterize the dependence of first-order response statistics on pupil-indexed arousal, we built a state-dependent generalized linear model. For each recorded unit, i, the input to this model was defined as the peri-stimulus time histogram (PSTH) response averaged over all stimulus repetitions ($r_0(t)$). The predicted firing rate was calculated by scaling the PSTH by a pupil-dependent multiplicative and additive factor to model pupil-dependent changes in gain and baseline firing rate over time (Eqn. 1). To account for any potential saturation of pupil-dependent modulation, pupil was first passed through a static sigmoid nonlinearity, F (Figure 2). The parameters of this nonlinearity, as well as the gain (β_0) and baseline (β_1) coefficients were fit independently for each cell using 10-fold jackknifed cross validation using

the Neural Encoding Model System (NEMS, `https://github.com/LBHB/NEMS`).

$$\hat{r}_i(t) = \beta_0\Big(1 + F\big(p(t)\big)r_{0,i}(t)\Big) + \beta_1 F\big(p(t)\big) \tag{1}$$

Because of the static nonlinearity applied to pupil, simply extracting gain (β_0) and baseline (β_1) parameters did not provide an easily interpretable way to evaluate the prevalence of gain vs. baseline effects in our model. Therefore, we instead computed the unique modulation of firing rate due to each, respectively, as a function of pupil condition:

$$MI_{\beta_0} = \beta_0\Big(< F(p_{large}(t)) > - < F(p_{small}(t)) > \Big) \tag{2}$$

$$MI_{\beta_1} = \beta_1\Big(< F(p_{large}(t)) > - < F(p_{small}(t)) > \Big) \tag{3}$$

Where <> indicates the mean over time and *large* / *small* refer to median splits of pupil size across the experimental session.

To summarize overall pupil-dependent modulation without differentiating between baseline and gain, we measured an overall pupil modulation index (MI). MI was defined by as the mean sound-evoked response when pupil was large minus the mean response when pupil was small, normalized by the sum of these two quantities. Large and small trials were defined based on a median split of pupil size across the entire recording session.

$$MI = \frac{\bar{r}_{large} - \bar{r}_{small}}{\bar{r}_{large} + \bar{r}_{small}} \tag{4}$$

Noise correlation:

Pairwise noise correlations were measured by grouping spike counts into 250 ms bins, extracting only evoked periods (epochs when sound stimuli were playing), and computing Pearson's correlation between all pairwise combinations of z-scored spike counts. Z-scores were calculated for each stimulus independently, as in Eqn. 5, where $r(t)$ is the single trial response, r_0 is the trial averaged response, and σ is the standard deviation of spike counts across repetitions. Therefore, the z-scored spike counts $Z(t)$ of each neuron i for each stimulus s had mean zero and standard deviation one.

$$Z_i(t) = \frac{r_i(t) - r_{0,i}(s)}{\sigma_i(s)} \tag{5}$$

Noise correlations were also measured after first band-pass filtering spike counts. In order to achieve high frequency resolution, data were binned at 100 Hz (10 ms bins) for this analysis. To band-pass filter the data, we first computed the Fast Fourier transform of each unit's spike counts. Next, in the frequency domain we used a Tukey window to extract the frequencies of interest. Finally, we inverted the modified signal back into the time domain, z-scored the data, and computed Pearson's correlation as above. For pupil-corrected noise correlations, we performed the same procedure after first performing the pupil-correction (see below).

Pupil-correction:

To remove changes in firing rate that were associated with pupil, we subtracted the pupil-dependent GLM prediction ($\hat{r}_i(t)$) from each unit's response ($r_i(t)$). Next, we added back the average stimulus response, or PSTH ($r_0(t)$), in order to preserve the sensory information encoded by each unit. Thus, the pupil-corrected response ($r_{corrected}$) was calculated in two steps, as follows:

$$r_{residual,i}(t) = r_i(t) - \hat{r}_i(t) \tag{6}$$

$$r_{corrected,i} = r_{residual,i}(t) + r_{0,i}(t) \tag{7}$$

Pairwise stimulus discrimination:

Natural sound samples were broken into non-overlapping 250 ms segments, similar to the procedure followed by Pachitariu et al., 2015.[11] In order to obtain cross-validated estimates of stimulus discriminability (Figure S1), we required that each stimulus be repeated at least 20 times. This was the case for 11 / 25 recording sites. These 11 recording sites were collected from four different animals. For the remaining sites (which contained only 10 repetitions per stimulus), we could not perform cross-validation. Therefore, for all analyses included in the main text in which we report neural discriminability (d'^2) estimates (Figures 3, 4, 5, and 7) only the high-repetition count recording sessions were included. However, we repeated key analyses using all 25 recording sessions without performing cross-validation and found that

results were largely consistent with the more rigorous approach (Figure S7).

For each pair of stimulus segments we extracted the N neuron X k trial response matrices, A and B. Because the number of recorded neurons was greater than the number of stimulus repetitions, we performed Targeted Dimensionality Reducton (TDR) to preserve only significant dimensions of the population response (see following section and Figures S2/S3). This allowed us to accurately estimate the population statistics. We quantified encoding accuracy in this reduced-dimensionality space by measuring neural stimulus discriminability, d'^2, the discrete analog of Fisher information:[28,30–33,39]

$$d'^2 = \Delta\boldsymbol{\mu}^T \Sigma^{-1} \Delta\boldsymbol{\mu} = \Delta\boldsymbol{\mu}^T \mathbf{w}_{opt} \tag{8}$$

where $\Delta\boldsymbol{\mu}$ represents the vector connecting the mean ensemble responses to stimulus A and stimulus B, $\Sigma = \frac{1}{2}(\Sigma_A + \Sigma_B)$ represents the mean noise-covariance matrix, and $\mathbf{w}_{opt}$ is the optimal decoding axis, *i.e.* the vector orthogonal to the optimal linear discrimination hyperplane in state-space. In practice, for our high repetition count data set we estimated $\mathbf{w}_{opt}$ using 50-percent of trials (training data) then projected the held out 50-percent of trials (test data) onto this vector and measured discriminability. For a detailed schematic of this procedure, see Figure S1. Pupil-dependent measurements of d'^2 followed an identical procedure, but before measuring discriminability, the test data was first split in half based on median pupil size.

Throughout this work, we describe stimulus discriminability as a function of the relationship between the signal and noise subspace. To quantify this relationship, we defined the following two axes which we found to be critical for understanding encoding accuracy:

Discrimination axis magnitude ($|\Delta\boldsymbol{\mu}|$): For a given pair of stimuli A and B, we defined the discrimination axis (*i.e.* the axis containing sensory information) to be the vector $\Delta\boldsymbol{\mu}$. When two stimuli drove very different population activity, the magnitude of this vector, $|\Delta\boldsymbol{\mu}|$, was large and discriminability was high.

Noise interference ($|cos(\theta_{\Delta\boldsymbol{\mu},\mathbf{e}_1})|$): The structure of trial to trial variability in responses can vary depending on the stimulus.[36,66,67] Therefore, within the reduced dimensionality TDR space (see below) for each pair of stimuli, we defined the correlated variability axis, $\mathbf{e}_1$, as the first eigenvector of the average covariance matrix, $\Sigma = \frac{1}{2}(\Sigma_A + \Sigma_B)$. We then asked how this axis interacted with the signal by measuring the cosine similarity between $\Delta\boldsymbol{\mu}$ and $\mathbf{e}_1$. Thus, when the two were perfectly aligned our measure of noise interference was equal to one and when they were perfectly orthogonal, interference was zero.

To quantify the dependence of d'^2 and $\Delta d'^2$ on these two axes, we used the following linear regression models (Eqns. 9, 10). All variables were z-scored prior to model fits.

$$d'^2 = \beta_0 + \beta_1 |cos(\theta_{\Delta\boldsymbol{\mu},\mathbf{e}_1})| + \beta_2 |\Delta\boldsymbol{\mu}| + \beta_3 |cos(\theta_{\Delta\boldsymbol{\mu},\mathbf{e}_1})||\Delta\boldsymbol{\mu}| \tag{9}$$

$$\Delta d'^2 = \beta_0 + \beta_1 |cos(\theta_{\Delta\boldsymbol{\mu},\mathbf{e}_1})| + \beta_2 |\Delta\boldsymbol{\mu}| + \beta_3 |cos(\theta_{\Delta\boldsymbol{\mu},\mathbf{e}_1})||\Delta\boldsymbol{\mu}| \tag{10}$$

Targeted Dimensionality Reduction:

During our experiments we typically recorded from between 20 and 60 units simultaneously at a single recording site (on one 64-channel probe shank). However, at most, we repeated each individual stimulus only 24 times. Therefore, because the number of neurons was generally greater than the number of trials, estimation of the covariance matrix for a particular stimulus, A (Σ_A), was unreliable (Figure S3).[31] This poses a challenge for estimating neural discriminability as this process depends on accurate estimation of Σ. One technique that has been proposed to help deal with such limitations is dimensionality reduction.[31]

Because we were interested specifically in studying neural discriminability, we developed a Targeted Dimensionality Reduction (*TDR*) approach whereby we reduced our *N*-dimensional space, on a per-stimulus pair basis, to just two dimensions containing the most critical information for estimating stimulus discriminability: The sensory discrimination axis, $\Delta\boldsymbol{\mu}$, and the global noise axis, which we defined as the first principal component of the pooled noise data over all stimuli and neurons (raw spike counts after subtracting the mean stimulus response from each neuron).[17] Thus, the noise axis was fixed between stimulus pairs but $\Delta\boldsymbol{\mu}$ was not. We defined $TDR_1 = \Delta\boldsymbol{\mu}$ and TDR_2 as the axis orthogonal to $\Delta\boldsymbol{\mu}$ and in the plane spanned by $\Delta\boldsymbol{\mu}$ and the noise axis. We illustrate this process graphically (Figure S2). Analysis of neural discriminability using *TDR* in simulated data sets (Figure S3) demonstrated the efficacy of this method both in preventing over-fitting and in facilitating accurate estimation of discriminability. Furthermore, though it is not a focus of the current work, we demonstrate that although this method only explicitly preserves the first noise principal component, it is still capable of detecting information limiting correlations[30] that recent work has suggested might be restricted to relatively low variance dimensions of the data.[31,33]

Pupil-dependent changes in stimulus discriminability

d'^2, measured across pupil states, could vary greatly across the sensory response. Therefore, in order to measure pupil-dependent changes in coding accuracy, we used a normalized metric, $\Delta d'^2$ that allowed us to directly compare relative changes in discriminability for different stimulus pairs and recording sites. To measure this, we computed a pupil-dependent modulation index of d'^2. For each stimulus pair, $\Delta d'^2$ was defined as the d'^2 measured during large pupil trials minus d'^2 for small pupil trials, normalized by the sum of these two quantities.

$$\Delta d'^2 = \frac{d'^2_{large} - d'^2_{small}}{d'^2_{large} + d'^2_{small}} \tag{11}$$

Response simulations:

To dissociate arousal-dependent modulation of first-order statistics, such as response gain, from second-order effects, such as correlated variability, we followed a procedure similar to that in Cohen et al., 2009.[19] Simulated data sets were generated independently for both large and small pupil conditions. In the first-order simulation, first-order statistics were allowed to vary between large and small pupil while second-order statistics were fixed. In the first+second-order simulation, both first- and second-order statistics changed between large and small pupil.

First-order simulation: Pupil-dependent first-order statistics (mean and variance) were measured using the raw data. Next, we built covariance matrices, $\Sigma_{1,large}$ and $\Sigma_{1,small}$, where the off-diagonal elements in each pupil condition were fixed to their measured values over all trials and the diagonal elements of each matrix were set to the measured variance in each respective pupil condition. We simulated first-order population responses, R_1 by drawing trials from a multivariate Gaussian distributions as follows:

$$R_{1,large} = \mathcal{N}(\mu_{large}, \Sigma_{1,large}), \qquad R_{1,small} = \mathcal{N}(\mu_{small}, \Sigma_{1,small}) \tag{12}$$

First + Second-order simulation: Pupil-dependent first-order statistics (mean and variance) *and* second-order statistics (off-diagonal elements of the covariance matrix) were measured using the raw data. We used these values to define covariance matrices $\Sigma_{2,large}$ and $\Sigma_{2,small}$ where the off-diagonal elements were set based on their covariance values in each respective pupil condition and the diagonal elements were set to the measured single-unit variances. We simulated first+second-order population responses, R_{1+2} as above by drawing trials from a

multivariate Gaussian distributions. All simulations were performed on a per-stimulus basis.

Identification of noise correlation modulation axis

We identified the axis in state space along which noise correlations were most modulated by pupil. First, we measured large and small pupil noise covariance matrices. Noise covariance matrices were defined as the covariance matrix of the z-scored spiking activity (z-scored on a per-stimulus basis). We then set the diagonal of this matrix equal to zero, because for this analysis we were interested only in capturing shared patterns of variability. Finally, we computed the difference covariance matrix (small pupil matrix minus large pupil matrix) and performed an eigen-decomposition. The largest eigenvalue corresponded to the axis along which noise correlations changed most between the two pupil states. Significance of this axis was evaluated with a permutation test.

Statistical analyses

Our data followed a nested structure; multiple cells were recorded from the same animal and many different stimuli were presented during each experimental session. Therefore, it is possible our results could be biased by differences between animals and/or experimental recording session. To account for this in our statistical tests, we took one of the following two approaches: (1) Averaged metrics across cells (or pairs of cells) and sound stimuli within a recording session before performing statistical tests or (2) Performed statistical tests using hierarchical bootstrapping.[68] Although each approach reduces statistical power relative to treating each individual measurement as wholly independent, they provide conservative estimates of p-values and reduce the chance of detecting false positives.[68] The only exception to this approach was in the estimates of pupil-dependent effects on evoked firing rates in single neurons shown in Figure 1.

For estimating significance of pupil-effects in single neurons, we performed a jackknifed t-test for each individual neuron. Data were split into 20 non-overlapping estimation / validation sets that together tiled the entire experimental session. For each validation set, a prediction correlation was computed for both the full pupil-dependent GLM and the pupil-shuffled GLM. Cells where the mean pupil-dependent model performance was 2 standard errors greater than the mean shuffled model performance ($p \leq 0.05$) were considered to have significant pupil effects.

For all statistical tests measuring large vs. small pupil effects where we first averaged results within recording session, we performed a two-tailed Wilcoxon signed-rank test. For each test, we report the test statistic, W, the p-value, and the exact n number of recording sessions used to perform the test. In cases where we performed a hierarchical bootstrap, we report the direct bootstrap probability of the null hypothesis.[68] In both cases, we also provide the mean and standard error of the number of measurements per recording session

To quantify patterns in d'^2 and $\Delta d'^2$ heatmaps, we performed standard linear regression on a per-recording site basis using the statsmodel package for Python (`https://www.statsmodels.org/stable/index.html`). To determine the significance of regression coefficients at the group level, we performed a Mann-Whitney U test to determine if the parameter distribution over recording sites was significantly different than zero. For each test, we report the test statistic, U, the p-value, and the exact n number of recording sessions used to perform the test.

To measure the significance of the correlation between pupil-dependent changes in stimulus discriminability and pupil variance per recording session, we performed a permutation test. We randomly shuffled our distribution of mean $\Delta d'^2$ across recording sites and computed the correlation of these re-sampled values with pupil variance. This was repeated 1000 times in order to calculate a p-value, which is reported in the text along with the true measured correlation coefficient.

Finally, to determine if the measured noise correlation modulation axis for a given recording session was significant, we performed a permutation test. For each recording session, pupil size was shuffled in time before classifying trials as large or small pupil. After shuffling, spike covariance matrices were measured for large and small pupil trials and an eigen-decomposition was performed on the difference between these matrices. This was repeated 20 times and the mean / standard error of the resulting eigenvalues was stored. Significant dimensions were those in which the actual measured eigenvalue was 2 standard errors greater than the mean shuffled value.

Glossary

- *Discrimination axis* ($\Delta\boldsymbol{\mu}$): Vector connecting the mean population response to stimulus A and stimulus B.
- *Pooled noise data*: All single trial data after subtracting off mean stimulus responses from each neuron.

- *Global noise axis*: First principal component of pooled noise data.
- *Correlated variability axis* ($\mathbf{e}_1$): Stimulus-pair specific estimate of noise axis. First principal component of noise data pooled across the two stimuli within *TDR* space.
- *Optimal decoding axis* ($\mathbf{w}_{opt}$): Axis orthogonal to the optimal linear discrimination boundary in *TDR* space.
- *Discrimination axis magnitude* ($|\Delta\boldsymbol{\mu}|$): Vector magnitude of the discrimination axis.
- *Noise interference* ($|cos(\theta_{\mathbf{e}_1,\Delta\boldsymbol{\mu}})|$): Cosine similarity between the correlated variability axis and discrimination axis.

Data Availability

The datasets analyzed in this study are available from the corresponding author upon reasonable request.

Author contributions

CRH and SVD designed experiments. CRH performed experiments, analyzed data, and developed data analysis and pupil tracking software. DS assisted with data collection. ZPS assisted with software development for pupil tracking and assisted with data collection. CRH and SVD wrote the manuscript. All authors edited the manuscript.

Acknowledgments

This work was supported by a National Science Foundation Graduate Research Fellowship (NSF GRFP, GVPRS0015A2) (CRH), the National Institute of Health (NIH, R01 DC0495) (SVD), Achievement Rewards for College Scientists (ARCS) Portland chapter (CRH), and by the Tartar Trust at Oregon Health and Science University (CRH).

References

1. A. Ryan and J. Miller. Effects of behavioral performance on single-unit firing patterns in inferior colliculus of the rhesus monkey. *Journal of Neurophysiology*, 40(4):943–956, July 1977.

2. Jonathan Fritz, Shihab Shamma, Mounya Elhilali, and David Klein. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience*, 6(11):1216–1223, November 2003. Number: 11 Publisher: Nature Publishing Group.

3. Jonathan Fritz, Mounya Elhilali, and Shihab Shamma. Active listening: task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. *Hearing Research*, 206(1-2):159–176, August 2005.

4. Gonzalo H. Otazu, Lung-Hao Tai, Yang Yang, and Anthony M. Zador. Engaging in an auditory task suppresses responses in auditory cortex. *Nature Neuroscience*, 12(5):646–654, May 2009.

5. Chen-Chung Lee and John C. Middlebrooks. Auditory cortex spatial sensitivity sharpens during task performance. *Nature Neuroscience*, 14(1):108–114, January 2011.

6. Daniel P. Knudsen and Timothy Q. Gentner. Active recognition enhances the representation of behaviorally relevant information in single auditory forebrain neurons. *Journal of Neurophysiology*, 109(7):1690–1703, April 2013.

7. Pingbo Yin, Jonathan B. Fritz, and Shihab A. Shamma. Rapid Spectrotemporal Plasticity in Primary Auditory Cortex during Behavior. *The Journal of Neuroscience*, 34(12):4396–4408, March 2014.

8. S. J. Slee and S. V. David. Rapid Task-Related Plasticity of Spectrotemporal Receptive Fields in the Auditory Midbrain. *Journal of Neuroscience*, 35(38):13090–13102, September 2015.

9. J. D. Downer, M. Niwa, and M. L. Sutter. Task Engagement Selectively Modulates Neural Correlations in Primary Auditory Cortex. *Journal of Neuroscience*, 35(19):7565–7574, May 2015.

10. Kishore V Kuchibhotla, Jonathan V Gill, Grace W Lindsay, Eleni S Papadoyannis, Rachel E Field, Tom A Hindmarsh Sten, Kenneth D Miller, and Robert C Froemke. Parallel processing by cortical inhibition enables context-dependent behavior. *Nature Neuroscience*, 20(1):62–71, January 2017.

11. M. Pachitariu, D. R. Lyamzin, M. Sahani, and N. A. Lesica. State-Dependent Population Coding in Primary Auditory Cortex. *Journal of Neuroscience*, 35(5):2058–2073, February 2015.

12. Joshua D. Downer, Brittany Rapone, Jessica Verhein, Kevin N. O'Connor, and Mitchell L. Sutter. Feature-Selective Attention Adaptively Shifts Noise Correlations in Primary Auditory Cortex. *The Journal of Neuroscience*, 37(21):5378–5392, May 2017.

13. Jermyn Z See, Craig A Atencio, Vikaas S Sohal, and Christoph E Schreiner. Coordinated neuronal ensembles in primary auditory cortical columns. *eLife*, 7:e35587, June 2018.

14. Michele N Insanally, Ioana Carcea, Rachel E Field, Chris C Rodgers, Brian DePasquale, Kanaka Rajan, Michael R DeWeese, Badr F Albanna, and Robert C Froemke. Spike-timing-dependent ensemble encoding by non-classically responsive cortical neurons. *eLife*, 8:e42409, January 2019. Publisher: eLife Sciences Publications, Ltd.

15. Sophie Bagur, Martin Averseng, Diego Elgueda, Stephen David, Jonathan Fritz, Pingbo Yin, Shihab Shamma, Yves Boubenec, and Srdjan Ostojic. Go/No-Go task engagement enhances population representation of target stimuli in primary auditory cortex. *Nature Communications*, 9(1), December 2018.

16. Pei-Ann Lin, Samuel K. Asinof, Nicholas J. Edwards, and Jeffry S. Isaacson. Arousal regulates frequency tuning in primary auditory cortex. *Proceedings of the National Academy of Sciences*, November 2019.

17. Dmitry Kobak, Jose L Pardo-Vazquez, Mafalda Valente, Christian K Machens, and Alfonso Renart. State-dependent geometry of population activity in rat auditory cortex. *eLife*, 8:e44526, April 2019.

18. Zachary P. Schwartz, Brad N. Buran, and Stephen V. David. Pupil-associated states modulate excitability but not stimulus selectivity in primary auditory cortex. *Journal of Neurophysiology*, 123(1):191–208, November 2019. Publisher: American Physiological Society.

19. Marlene R Cohen and John H R Maunsell. Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*, 12(12):1594–1600, December 2009.

20. Jude F. Mitchell, Kristy A. Sundberg, and John H. Reynolds. Spatial Attention Decorrelates Intrinsic Activity Fluctuations in Macaque Area V4. *Neuron*, 63(6):879–888, September 2009.

21. Martin Vinck, Renata Batista-Brito, Ulf Knoblich, and Jessica A. Cardin. Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. *Neuron*, 86(3):740–754, May 2015.

22. George H. Denfield, Alexander S. Ecker, Tori J. Shinn, Matthias Bethge, and Andreas S. Tolias. Attentional fluctuations induce shared variability in macaque primary visual cortex. *Nature Communications*, 9(1), December 2018.

23. P. L. E. van Kan, R. P. Scobey, and A. J. Gabor. Response covariance in cat visual cortex. *Experimental Brain Research*, 60(3):559–563, November 1985.

24. K. H. Britten, M. N. Shadlen, W. T. Newsome, and J. A. Movshon. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, 12(12):4745–4765, December 1992.

25. Ehud Zohary, Michael N. Shadlen, and William T. Newsome. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, 370(6485):140–143, July 1994.

26. M. N. Shadlen, K. H. Britten, W. T. Newsome, and J. A. Movshon. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *Journal of Neuroscience*, 16(4):1486–1510, February 1996.

27. A. J. Parker and W. T. Newsome. SENSE AND THE SINGLE NEURON: Probing the Physiology of Perception. *Annual Review of Neuroscience*, 21(1):227–277, March 1998.

28. L. F. Abbott and Peter Dayan. The Effect of Correlated Variability on the Accuracy of a Population Code. *Neural Computation*, 11(1):91–101, January 1999.

29. Bruno B. Averbeck, Peter E. Latham, and Alexandre Pouget. Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, 7(5):358–366, May 2006.

30. Rubén Moreno-Bote, Jeffrey Beck, Ingmar Kanitscheider, Xaq Pitkow, Peter Latham, and Alexandre Pouget. Information-limiting correlations. *Nature Neuroscience*, 17:1410, September 2014.

31. Oleg I. Rumyantsev, Jérôme A. Lecoq, Oscar Hernandez, Yanping Zhang, Joan Savall, Radosław Chrapkiewicz, Jane Li, Hongkui Zeng, Surya Ganguli, and Mark J. Schnitzer. Fundamental bounds on the fidelity of sensory cortical coding. *Nature*, March 2020.

32. Ramon Bartolo, Richard C. Saunders, Andrew R. Mitz, and Bruno B. Averbeck. Information limiting correlations in large neural populations. *The Journal of Neuroscience*, pages 2072–19, January 2020.

33. MohammadMehdi Kafashan, Anna Jaffe, Selmaan N. Chettih, Ramon Nogueira, Iñigo Arandia-Romero, Christopher D. Harvey, Rubén Moreno-Bote, and Jan Drugowitsch. Scaling of information in large neural populations reveals signatures of information-limiting correlations. *bioRxiv*, page 2020.01.10.902171, January 2020.

34. A. S. Ecker, P. Berens, A. S. Tolias, and M. Bethge. The Effect of Noise Correlations in Populations of Diversely Tuned Neurons. *Journal of Neuroscience*, 31(40):14272–14283, October 2011.

35. Felix Franke, Michele Fiscella, Maksim Sevelev, Botond Roska, Andreas Hierlemann, and Rava Azeredo da Silveira. Structures of Neural Correlation and How They Favor Coding. *Neuron*, 89(2):409–422, January 2016.

36. Joel Zylberberg, Jon Cafaro, Maxwell H. Turner, Eric Shea-Brown, and Fred Rieke. Direction-Selective Circuits Shape Noise to Ensure a Precise Population Code. *Neuron*, 89(2):369–383, January 2016.

37. Alexander S. Ecker, George H. Denfield, Matthias Bethge, and Andreas S. Tolias. On the Structure of Neuronal Population Activity under Fluctuations in Attentional State. *The Journal of Neuroscience*, 36(5):1775–1789, February 2016.

38. Carsen Stringer, Marius Pachitariu, Nicholas Steinmetz, Charu Bai Reddy, Matteo Carandini, and Kenneth D. Harris. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437):eaav7893, April 2019.

39. Bruno B. Averbeck and Daeyeol Lee. Effects of noise correlations on information encoding and decoding. *Journal of Neurophysiology*, 95(6):3633–3644, June 2006.

40. Matthew J. McGinley, Stephen V. David, and David A. McCormick. Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron*, 87(1):179–192, July 2015.

41. Neil C. Rabinowitz, Robbe L. Goris, Marlene Cohen, and Eero P. Simoncelli. Attention stabilizes the shared gain of V4 populations. *eLife*, 4:e08998, November 2015.

42. Jiangang Du, Timothy J. Blanche, Reid R. Harrison, Henry A. Lester, and Sotiris C. Masmanidis. Multiplexed, High Density Electrophysiology with Nanofabricated Neural Probes. *PLOS ONE*, 6(10):e26204, October 2011.

43. Josh H. McDermott, Michael Schemitsch, and Eero P. Simoncelli. Summary statistics in auditory perception. *Nature Neuroscience*, 16(4):493–498, April 2013. Number: 4 Publisher: Nature Publishing Group.

44. Tatiana A. Engel, Nicholas A. Steinmetz, Marc A. Gieselmann, Alexander Thiele, Tirin Moore, and Kwabena Boahen. Selective modulation of cortical state during spatial attention. *Science*, 354(6316):1140–1144, December 2016.

45. A. M. Ni, D. A. Ruff, J. J. Alberts, J. Symmonds, and M. R. Cohen. Learning and attention reveal a general relationship between population activity and behavior. *Science*, 359(6374):463–465, January 2018.

46. Nima Mesgarani and Edward F. Chang. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*, 485(7397):233–236, May 2012. Number: 7397 Publisher: Nature Publishing Group.

47. Robbe L T Goris, J Anthony Movshon, and Eero P Simoncelli. Partitioning neuronal variability. *Nature Neuroscience*, 17(6):858–865, June 2014.

48. Carsen Stringer, Michalis Michaelos, and Marius Pachitariu. High precision coding in visual cortex. *bioRxiv*, page 679324, November 2019.

49. Robert Desimone and John Duncan. Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, 18(1):193–222, March 1995. Publisher: Annual Reviews.

50. Benjamin R. Cowley, Adam C. Snyder, Katerina Acar, Ryan C. Williamson, Byron M. Yu, and Matthew A. Smith. Slow Drift of Neural Activity as a Signature of Impulsivity in Macaque Visual and Prefrontal Cortex. *Neuron*, 0(0), August 2020. Publisher: Elsevier.

51. Daniela Saderi, Zachary P. Schwartz, Charlie R. Heller, Jacob R. Pennington, and Stephen V. David. Dissociation of task engagement and arousal effects in auditory cortex and midbrain. *bioRxiv*, page 2020.06.16.155432, June 2020. Publisher: Cold Spring Harbor Laboratory Section: New Results.

52. Gioia De Franceschi and Tania Rinaldi Barkat. Task-induced modulations of neuronal activity along the auditory pathway. preprint, Neuroscience, July 2020.

53. Jacob Reimer, Matthew J. McGinley, Yang Liu, Charles Rodenkirch, Qi Wang, David A. McCormick, and Andreas S. Tolias. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications*, 7:13289, 2016.

54. Michael Goard and Yang Dan. Basal forebrain activation enhances cortical coding of natural scenes. *Nature Neuroscience*, 12(11):1444–1449, November 2009.

55. David M. Schneider, Anders Nelson, and Richard Mooney. A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature*, 513(7517):189–194, August 2014.

56. Matthew J. McGinley, Martin Vinck, Jacob Reimer, Renata Batista-Brito, Edward Zagha, Cathryn R. Cadwell, Andreas S. Tolias, Jessica A. Cardin, and David A. McCormick. Waking State: Rapid Variations Modulate Neural and Behavioral Responses. *Neuron*, 87(6):1143–1161, September 2015.

57. Shi Tong Liu, Pilar Montes-Lourido, Xiaoqin Wang, and Srivatsun Sadagopan. Optimal features for auditory categorization. *Nature Communications*, 10(1):1302, March 2019. Number: 1 Publisher: Nature Publishing Group.

58. S. V. David, J. B. Fritz, and S. A. Shamma. Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proceedings of the National Academy of Sciences*, 109(6):2144–2149, February 2012.

59. Daniela Saderi, Brad N. Buran, and Stephen V. David. Streaming of Repeated Noise in Primary and Secondary Fields of Auditory Cortex. *Journal of Neuroscience*, 40(19):3783–3798, May 2020. Publisher: Society for Neuroscience Section: Research Articles.

60. Jacob Pennington and Stephen David. Complementary effects of adaptation and gain control on sound encoding in primary auditory cortex. preprint, Neuroscience, January 2020.

61. Jennifer K. Bizley, Fernando R. Nodal, Israel Nelken, and Andrew J. King. Functional organization of ferret auditory cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 15(10):1637–1653, October 2005.

62. Diego Elgueda, Daniel Duque, Susanne Radtke-Schuller, Pingbo Yin, Stephen V. David, Shihab A. Shamma, and Jonathan B. Fritz. State-dependent encoding of sound and behavioral meaning in a tertiary region of the ferret auditory cortex. *Nature Neuroscience*, 22(3):447–459, 2019.

63. Marius Pachitariu, Nicholas Steinmetz, Shabnam Kadir, Matteo Carandini, and Harris Kenneth D. Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. *bioRxiv*, page 061481, June 2016.

64. Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Densely Connected Convolutional Networks. *arXiv:1608.06993 [cs]*, January 2018. arXiv: 1608.06993.

65. Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, June 2009. ISSN: 1063-6919.

66. Adam Kohn and Matthew A. Smith. Stimulus Dependence of Neuronal Correlation in Primary Visual Cortex of the Macaque. *Journal of Neuroscience*, 25(14):3661–3673, April 2005.

67. Douglas A. Ruff and Marlene R. Cohen. Stimulus Dependence of Correlated Variability across Cortical Areas. *Journal of Neuroscience*, 36(28):7546–7556, July 2016.

68. Varun Saravanan, Gordon J. Berman, and Samuel J. Sober. Application of the hierarchical bootstrap to multi-level data in neuroscience. *arXiv:2007.07797 [q-bio]*, July 2020. arXiv: 2007.07797.