

Slow cortical dynamics generate context processing and novelty detection

Highlights

- Multidimensional encoding of stimulus temporal context by neuronal ensembles
- Stimulus statistics and complexity drive cortical areal representation of novelty
- Cortical responses have slow dynamics and historicity, which encode stimulus context
- Recurrent neural networks recapitulate biological findings, generating novelty detection

Authors

Yuriy Shymkiv, Jordan P. Hamm,
Sean Escola, Rafael Yuste

Correspondence

ys2605@columbia.edu

In brief

The brain efficiently processes sensory information by selecting novelty and discarding redundancies. In this work, Shymkiv et al. demonstrate that through a high-dimensional ensemble representation, the slow network dynamics of recurrent cortical circuits serve as a memory that captures stimulus novelty and temporal context.

Report

Slow cortical dynamics generate context processing and novelty detection

Yuriy Shymkiv,^{1,5,*} Jordan P. Hamm,^{1,4} Sean Escola,^{2,3} and Rafael Yuste¹

¹Neurotechnology Center, Department of Biological Sciences, Columbia University, New York, NY, USA

²Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA

³Department of Psychiatry, Columbia University, New York, NY, USA

⁴Present address: Department of Psychiatry and Neuroscience Institute, New York University Grossman School of Medicine and Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY, USA

⁵Lead contact

*Correspondence: ys2605@columbia.edu

<https://doi.org/10.1016/j.neuron.2025.01.011>

SUMMARY

The cortex amplifies responses to novel stimuli while suppressing redundant ones. Novelty detection is necessary to efficiently process sensory information and build predictive models of the environment, and it is also altered in schizophrenia. To investigate the circuit mechanisms underlying novelty detection, we used an auditory “oddball” paradigm and two-photon calcium imaging to measure responses to simple and complex stimuli across mouse auditory cortex. Stimulus statistics and complexity generated specific responses across auditory areas. Neuronal ensembles reliably encoded auditory features and temporal context. Interestingly, stimulus-evoked population responses were particularly long lasting, reflecting stimulus history and affecting future responses. These slow cortical dynamics encoded stimulus temporal context and generated stronger responses to novel stimuli. Recurrent neural network models trained on the oddball task also exhibited slow network dynamics and recapitulated the biological data. We conclude that the slow dynamics of recurrent cortical networks underlie processing and novelty detection.

INTRODUCTION

Effective and timely processing of the sensory environment is critical for living organisms. In mammals, basic stimulus features such as image luminance or sound frequency are processed early in the sensory stream, improving information contrast. Stimulus recognition and the spatiotemporal processing of more abstract features typically appear higher up in the sensory cortex.¹ In addition, both subcortical activity and cortical activity display efficient processing of temporal context, where neuronal responses to repetitive stimuli are suppressed, while responses to unexpected, contextually informative events are amplified.^{2,3}

The importance of temporal context processing is evident in schizophrenia (SZ), a severe chronic mental disorder that affects perception and thought. Contextual processing deficits in SZ patients have traditionally been measured using electroencephalogram (EEG) and the “oddball” paradigm, where oddball (deviant/novel/target) events are interleaved within a sequence of repeated stimuli (standards/redundants).^{2,4–7} The difference between the resulting event-related potential (ERP) waveforms of deviant and redundant events contains a component known as “mismatch negativity” (MMN), which reflects the circuit’s response to sudden changes in the environment (deviance/novelty detection). A reduced MMN component is a well-replicated

biomarker of deficient context processing and cognitive dysfunction in SZ.^{4,6–9}

The source underlying MMN is a population signal, and its circuit basis is of significant interest.^{5,10,11} Beyond its pathological roles, MMN and novelty detection are conceptually related to data compression, normalization, and error prediction, and they are integral to general theories of brain function, such as the Bayesian brain and predictive coding.^{12–15} Recordings from individual neurons in the sensory cortex demonstrate stimulus-specific adaptation (SSA), which is often interpreted as a neuronal correlate of MMN.^{3,16,17} However, SSA generally peaks earlier^{5,18} and is evoked in spatially distinct regions from MMN and deviance detection-related signals.^{18–22} Furthermore, unlike SSA, MMN is broadly tuned,²⁰ can be elicited by inputs that lack repetition altogether,^{5,18} and requires a memory representation.¹⁸ SSA may represent an early bottom-up processing step that precedes the more abstract deviance detection signal that is captured by MMN.^{23,24}

Traditional oddball and MMN studies in humans have been conducted using low-resolution recordings, which provide a coarse measure of activity convolved across thousands of neurons, making the neural circuit mechanisms underlying MMN generation inaccessible. The oddball paradigm has been translated to model organisms,^{3,17} including, more recently, to two-photon calcium imaging experiments in mice,^{22,25–27} thus

enabling recordings from neuronal populations with single-cell resolution. In these mouse experiments, a visual oddball paradigm activated distinct context-specific ensembles of neurons in the primary visual cortex (V1), with particularly large numbers of deviance detection neurons.²⁵

Since auditory hallucinations are among the most salient positive symptoms in SZ, and the majority of human MMN studies have been conducted in the auditory domain, understanding the circuit mechanisms underlying MMN in the auditory cortex (AC) is desirable. Additionally, how temporal context is processed by neural populations across primary and secondary regions, and as a function of stimulus complexity, remains unknown. To address these questions, we conducted two-photon calcium imaging across primary and secondary areas of the AC while awake mice passively listened to auditory oddball paradigms. We found neurons with strong deviance detection responses, for both simple and complex stimuli, in all imaged areas. Discrete neuronal ensembles represented stimulus identity, stimulus structure, and their temporal context. Stimulus responses were particularly long lasting, leading to context-dependent differences in trial-to-trial variability, generating histotoxicity (i.e., a dependency on recent activity). Finally, we modeled the data using artificial recurrent neural networks (RNNs) trained on the oddball task. These RNNs also developed slow dynamics, and their underlying activity recapitulated deviance detection, context- and frequency-specific population states, and the trial-to-trial variability differences observed in *in vivo* data. Thus, complex properties such as SSA and MMN emerged naturally in RNNs, which used slow dynamics to capture and organize stimuli with overlapping features according to their temporal context relevance.

RESULTS

Representation of simple and complex stimuli in AC

We recorded the activity of neuronal populations in the AC of awake mice ($n = 10$) while they were head-fixed, free to move on a circular treadmill, and passively listening to auditory stimuli (Figure 1A). Population calcium activity was measured using the genetically encoded indicator GCaMP6s, transgenically expressed in excitatory neurons²⁸ throughout the AC. Wide-field calcium imaging was performed through a glass cranial window implanted over the left AC, with auditory stimuli presented contralaterally to the right ear. The primary AC (A1), anterior auditory field (AAF; primary region), secondary AC (A2), and ultrasonic field (UF) were identified and subsequently targeted for two-photon experiments^{29–32} (Figures 1B, 1C, and S1). In imaged datasets, calcium activity was demixed³³ and deconvolved into firing rate traces (Figure 1D) and analyzed using trial-averaging and trial-to-trial methods (535- to 367- μm -wide field of view [FOV]; 111 datasets, 214 ± 127 neurons [mean \pm SD]).

Auditory stimuli were presented in a three-block oddball/many-standards paradigm. This three-block approach separates the two MMN components: SSA and deviance detection.^{10,34} Trials were sampled from 10 stimuli with varying features (Figure 1E). For “simple” stimuli, these were amplitude-modulated pure tones of 10 frequencies, logarithmically spaced between 2 and 76.9 kHz (1.5 octaves apart; Figure 1F,

left). In the first oddball block, two of the 10 stimuli were presented in a typical oddball pattern (5.25 ± 0.86 octave separation [mean \pm SD]), with “redundants” played at a 90% probability every second, and “deviants” interleaved with a 10% probability, representing contextually surprising events (Figure 1E, top). The second oddball block was sampled with identities of deviant and redundant stimuli reversed (Figure 1E, middle). Finally, in the control block, all 10 stimuli were randomly presented with a 10% probability every second (Figure 1E, bottom). Across the three blocks, identical stimuli were placed into three distinct temporal contexts: deviant, redundant, and neutral/control.

To measure the frequency tuning of neurons, we analyzed control blocks of simple tones and registered them onto a map of the AC (Figure 1G). Single-neuron tonotopic gradients in A1, A2, and AAF were consistent with wide-field mapping results.³⁵ We found cells tuned to all frequencies in the 2–76.9 kHz range, with a normal distribution centered at 8.10 kHz (± 3.13 octaves [SD], $R^2 = 0.92$; Figures 1H and S2D–S2F). Tone response latencies showed a bimodal distribution, indicative of onset and offset tuning, with underlying neurons consisting of discrete populations^{36–39} (Figures 1I and S2H). Thus, responses propagated through multiple neuronal populations throughout a given trial.

In parallel with the simple tones, we conducted oddball experiments with “complex” auditory frequency gratings,⁴⁰ i.e., wide-band stimuli with amplitude modulation gratings at 10 orientations spaced between 0 and π radians, with orthogonal orientations selected for oddball runs (Figure 1F, right). Tuning to complex stimuli was preferred at orientations 0 and 4/10–7/10 π radians, with the latter representing higher sweep speeds (Figure 1J). The response latency peak was delayed, with a smaller offset peak, compared with simple tones (Figure 1K), likely due to the complex stimulus statistics requiring information to be integrated across both frequencies and time.^{40–42}

Population dynamics are organized in low-dimensional manifolds

Next, we visualized the representations of auditory inputs in the population activity of AC using neuronal manifolds.^{43–46} These can be extracted using linear or nonlinear dimensionality reduction methods, such as principal-component analysis (PCA) and Isomap.⁴⁷ For all control trials and datasets, we created trial-averaged population vectors (Figures S3A–S3C). From these, for the simple stimuli, we estimated the linear dimensionality of input space using the participation ratio⁴⁸ (PR) at 5.54 (83.3% explained variance [EV] with 6 principal components [PCs]; Figures S3D–S3G). Nonlinear embedding with Isomap required only three components for 85% EV, with the population activity shaping into a linear frequency manifold (Figures 1L and S3H). This demonstrated encoding of a low-dimensional input manifold within a high-dimensional population space, likely allowing the network to achieve high-frequency resolution, i.e., the ability to encode nearby stimuli with discrete ensembles.

Applying nonlinear dimensionality reduction to complex stimuli revealed analogous encoding of a low-dimensional, circular manifold in approximately six-dimensional space (PR = 5.44, 86.3% EV with 6 PCs, 85.2% EV with 2 Isomap components; Figures 1M, S3D, and S3I). Interestingly, the grating stimulus space was a circle that folded over itself, revealing neuronal

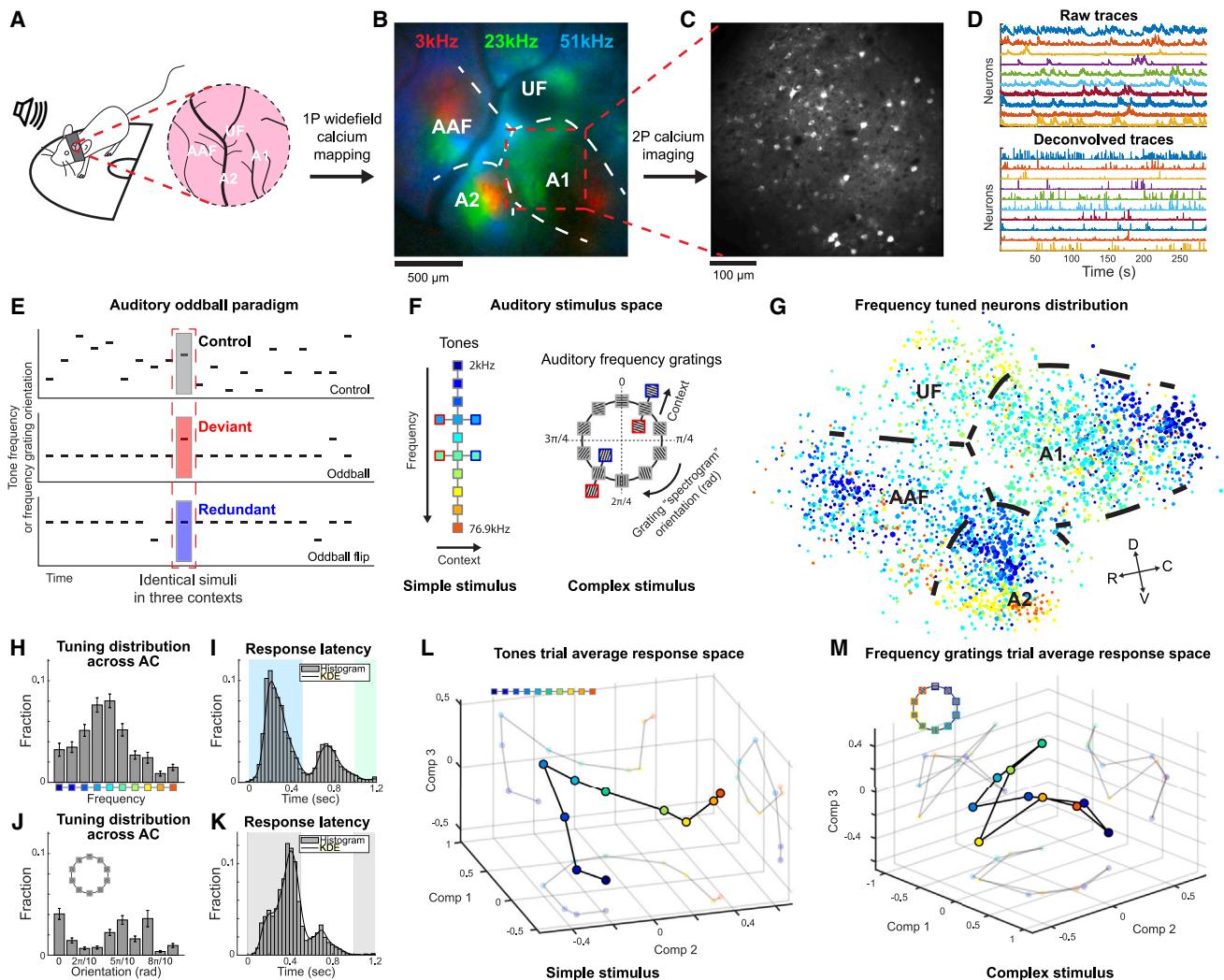


Figure 1. Population representations of simple and complex stimuli in auditory cortex

- (A) Experimental strategy with head-fixed mice on a circular treadmill.
 - (B) Areas of AC mapped using wide-field calcium imaging.
 - (C) Two-photon calcium imaging.
 - (D) Example raw and deconvolved neuronal responses.
 - (E) Stimulus design: three blocks of auditory oddball paradigm—oddball (top), flipped oddball (middle), and control (bottom).
 - (F) Varying stimulus complexity using simple tones (left) and complex auditory frequency gratings (right).
 - (G) Spatial distribution of simple tone-tuned neurons; point color and size represent tuning frequency and magnitude, respectively.
 - (H and I) Distribution cell counts and response latencies for simple tones.
 - (J and K) Same as in (H) and (I), but for complex gratings.
 - (L and M) Low-dimensional embedding of simple tone frequency and complex gratings response spaces.
- Error bars represent SEM across cells.

tuning for both grating sweep speed and orientation, potentially leading to complex adaptation patterns.

Multidimensional encoding of temporal context across auditory areas

We investigated how deviant, control, and redundant contexts were encoded by neurons across the AC. All areas contained neurons that were statistically significantly and in some cases exclusively tuned to context, with these neurons concentrating

in A1 and A2 (Figures 2A and S4A–S4C). Population and trial average calcium activity showed strong MMN-like deviance signals across all imaged areas and stimulus complexities (Figures 2B and 2C). Differences in the population and trial average traces were explained by the counts of context-responsive cells, as well as individual neuron response magnitudes and reliabilities (Figures 2D, 2E, and S4).

For simple tones, A1 and AAF traced out a biphasic response pattern, encoding deviance throughout both the onset and offset

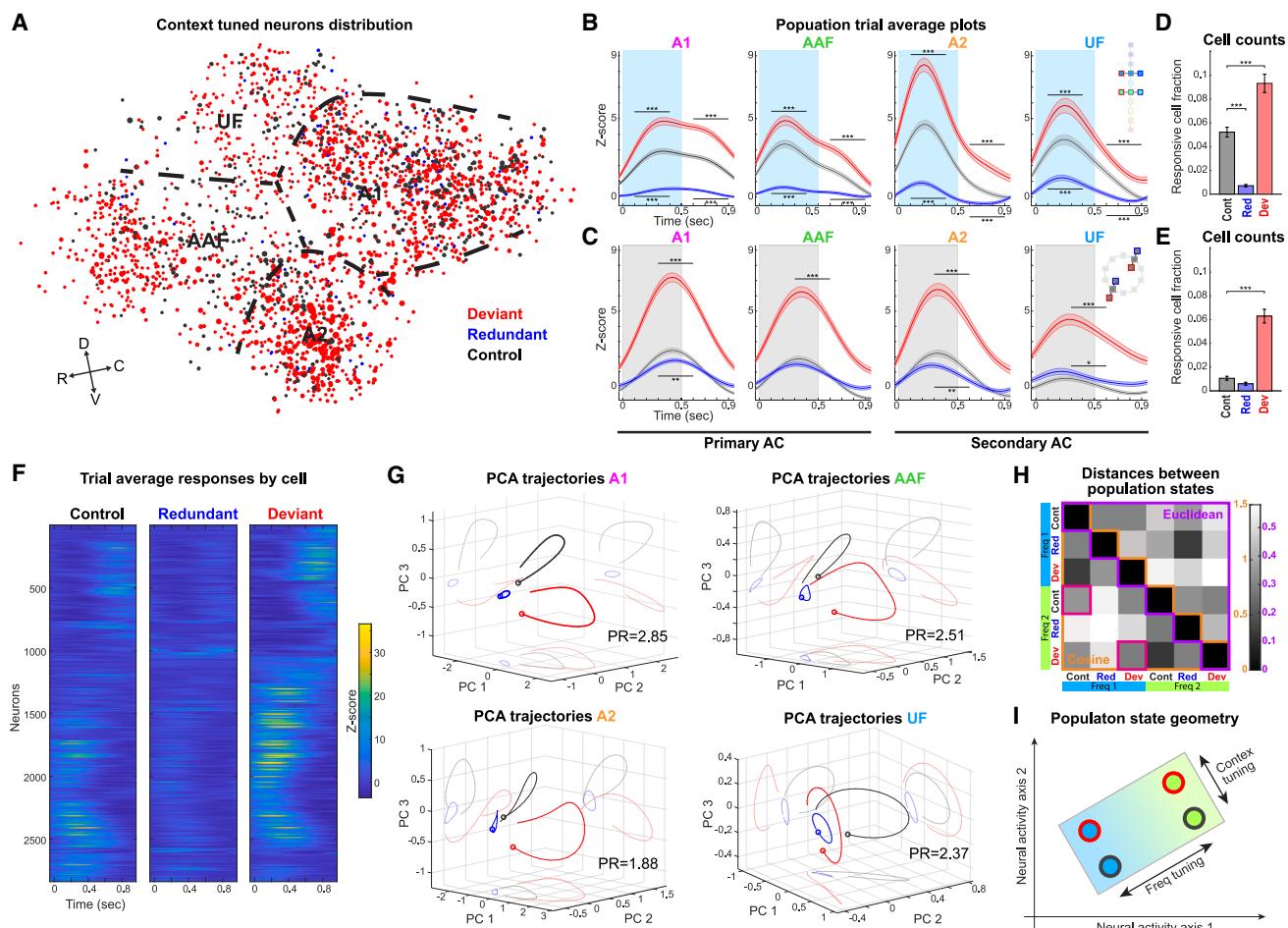


Figure 2. Neuronal ensembles encode temporal context across auditory areas

- (A) Spatial distribution of neurons tuned any of the three contexts; spot color and size represent context and tuning magnitude, respectively.
- (B and C) Population and trial average responses to contexts of simple tones and complex gratings.
- (D and E) Context-responsive cell fractions for simple tones and complex gratings.
- (F) Raster of trial-averaged activity of context-responsive cells, sorted by cosine similarity.
- (G) Dimensionality reduction with PCA of neuronal responses to the three contexts by AC areas; circle corresponds to start of 1-s trial, with 0.5-s stimulus duration.
- (H) Matrix of cosine and Euclidean distances between context-dependent network states; deviant-deviant and control-control distances between different frequencies are highlighted (magenta).
- (I) Graphical depiction of network separation of deviants and controls states of different frequencies.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Error bars reflect SEM across datasets. Extended statistics in Table S1.

periods. Simple tones showed the strongest deviance response in A2 during stimulus onset (Figure S5A), displaying hierarchical processing with increases in abstractness and deviance magnitude relative to A1.⁴⁹ Analysis of the beginnings of oddball runs also revealed that the rate of stimulus adaptation to tones was significantly slower in A1 than in A2⁴⁹ ($\tau_{A1}(1.07 \text{ s}) > \tau_{A2}(0.37 \text{ s})$, $t(14) = 4.16$, $p < 0.001$; Figures S6A–S6C).

Complex frequency gratings showed a single delayed deviance peak in all areas, which was strongest in A1 (Figures 2C and S5C). These results indicate that stimulus statistics and complexity can shape the temporal response profile and determine which cortical areas are engaged in the processing of novelty.^{42,50} Finally, the difference between control and redundant

responses for complex stimuli was not as large as for simple tones, especially in A1 and AAF, suggesting either enhanced adaptation during control trials or a lack of adaptation during redundant trials (Figures 2C and S5D–S5F).

Neuronal ensembles encode stimulus identity and temporal context

To analyze population dynamics underlying novelty detection, we created a trial-averaged raster of neurons responsive to any of the three contexts (Figure 2F). This representation revealed diverse tuning patterns, with some neurons showing exclusive tuning for each context, some displaying joint tuning for control and deviant contexts, and others exhibiting mixed

onset and offset specificity. We then reduced the data dimensionality using PCA and plotted the three largest PCs (Figure 2G). Across all imaged areas, the deviant trial trajectories clearly separated from control trajectories, while the redundant response remained small. The angle between these trajectories represents the difference in corresponding neuronal populations, highlighting the encoding of context by discrete ensembles. The magnitude of the trajectories represents the fractions of responsive cells and the response sizes. The similar shapes of these response manifolds across all conditions suggest that context representation is conserved across AC areas and stimulus complexities (Figures S8E–S8H). We measured the network deviance signal from these low-rank representations using the Euclidean distance between the deviant and redundant trajectories (6-PC dataset, >97% EV; Figures S8A–S8D). In A2, the deviance signal separated the fastest and reached the highest peak, again highlighting A2's importance in context processing of tones. For complex stimuli, the deviance signal profile was again similar across all areas, peaking in A1 (Figures S8I–S8L). This multidimensional representation of the MMN-like signal is potentially more accurate than population and trial averages, as it captures population differences in addition to the magnitude of the response.

We next analyzed the differences between context-specific populations of varying input frequencies by measuring cosine and Euclidean distances between them (Figures 2H, S7G, and S7H). Cosine distance reflects population overlap, whereas Euclidean distance is a joint measure of overlap, responsive cell fraction, and response size. Cosine distances between deviant populations of different frequencies approached orthogonality, similar to distances between corresponding control states (1.03 ± 0.03 and 1.10 ± 0.03 , respectively [mean \pm SEM]). These results demonstrated the full separation of features into discrete neuronal ensembles specific to context and frequency, suggesting the direct involvement of these AC areas in computing novelty detection (Figure 2I).

Trial-to-trial variability encodes stimulus history and temporal context

Averaging data over trials is useful for highlighting stimulus-specific signals, but it can also conceal trial-to-trial dynamics that are important for understanding cortical processing and computations. Here, we investigated how trial-to-trial responses were affected by their local temporal context. First, we computed trial-to-trial data dimensionality for control and deviant trials of the same frequency, as a function of the number of sampled neurons from a dataset (Figure 3A). Deviance response dimensionality quickly saturated, indicating high reliability. In contrast, control dimensionality scaled with the number of added neurons, without reaching saturation. This suggests that control responses are far less reliable and may encode more information than just the current stimulus identity. Analogously, we found a significant increase in trial-to-trial pairwise response similarity for deviant trials relative to control trials (Figure 3B). It is possible that trial-to-trial variability alone could explain the amplified trial average response of deviant traces relative to controls. To address this, we compared the magnitudes of individual trial responses across the three contexts (Figure 3C). Deviant trials

showed significantly larger individual responses relative to controls, suggesting that the combination of response sizes and trial-to-trial variability underlies novelty detection.

To directly measure how stimulus information persists in population activity after a trial ends, we used a bin-wise support vector machine (SVM) to decode the identity of individual trials throughout control runs (Figure 3E). Specifically, for all imaged frames within -1- to 3-s windows around control trials, we trained separate decoders to identify the corresponding stimuli. The advantage of a bin-wise decoder is that it can detect a memory trace, even if the population representation of that stimulus changes throughout the trial. Indeed, stimulus identity was decoded above chance level, starting from stimulus onset and extending into the following trial,^{51,52} with information persisting longest in A1 (Figure 3F).

The increased dimensionality and trial-to-trial variability of control responses suggest that during those runs, neuronal activity encodes more than just the current stimulus identity. One possibility is that a memory trace of the past stimulus persists in the population activity, contributing to the response variability of following trials. To measure the effects of this memory trace, we conducted additional recordings in A1 using control runs with variable interstimulus intervals (ISIs; $n = 5$ mice, 21 datasets, 148.35 ± 105.6 neurons [mean \pm SD]), suspecting that trial-to-trial variability would decrease if network activity was allowed to decay (Figure 3G). Indeed, trial-to-trial variability decreased with increasing ISI, as measured by both cosine similarity and mean pairwise correlations ($\tau_{\text{decay}} = 387 \pm 142$ ms [mean \pm SD]; Figures 3H–3J). Thus, the activity in the AC persists after the stimulus ends, moving through discrete population states, and impacting future responses, in a process we refer to as “slow network dynamics.”

RNN models recapitulate MMN

Our results demonstrate that discrete neuronal ensembles encode seemingly identical stimuli under different temporal contexts, with differences in response variability. We suspected that these findings could be explained by a recurrent network using slow dynamics to encode the temporal context of inputs. To test this, we modeled our data using RNNs trained to perform an oddball task (Figure 4A). We deemed continuous RNNs especially appropriate for this task because they inherently contain a time constant parameter (τ) that represents the speed of the network's response. Tuning this time constant provides a balance between the efficiency of current information discrimination and the duration of the memory trace.

We analyzed the responses of three types of RNNs: (1) trained on oddball tasks, (2) trained on frequency recognition (control) tasks, and (3) untrained. These RNNs contained 25–100 recurrently connected neurons with a ReLU nonlinearity, initialized with irregular activity, and τ set to 0.5 s (Figures S9A–S9E). After training, weights were fixed and RNNs tested with oddball and control inputs, using their neuronal activity for analysis similar to *in vivo* data.

We first analyzed the dimensionality of network activity under both control and oddball inputs (Figures S9F and S9G). Oddball-trained networks showed increased response dimensionality, compared with frequency-trained or untrained networks,

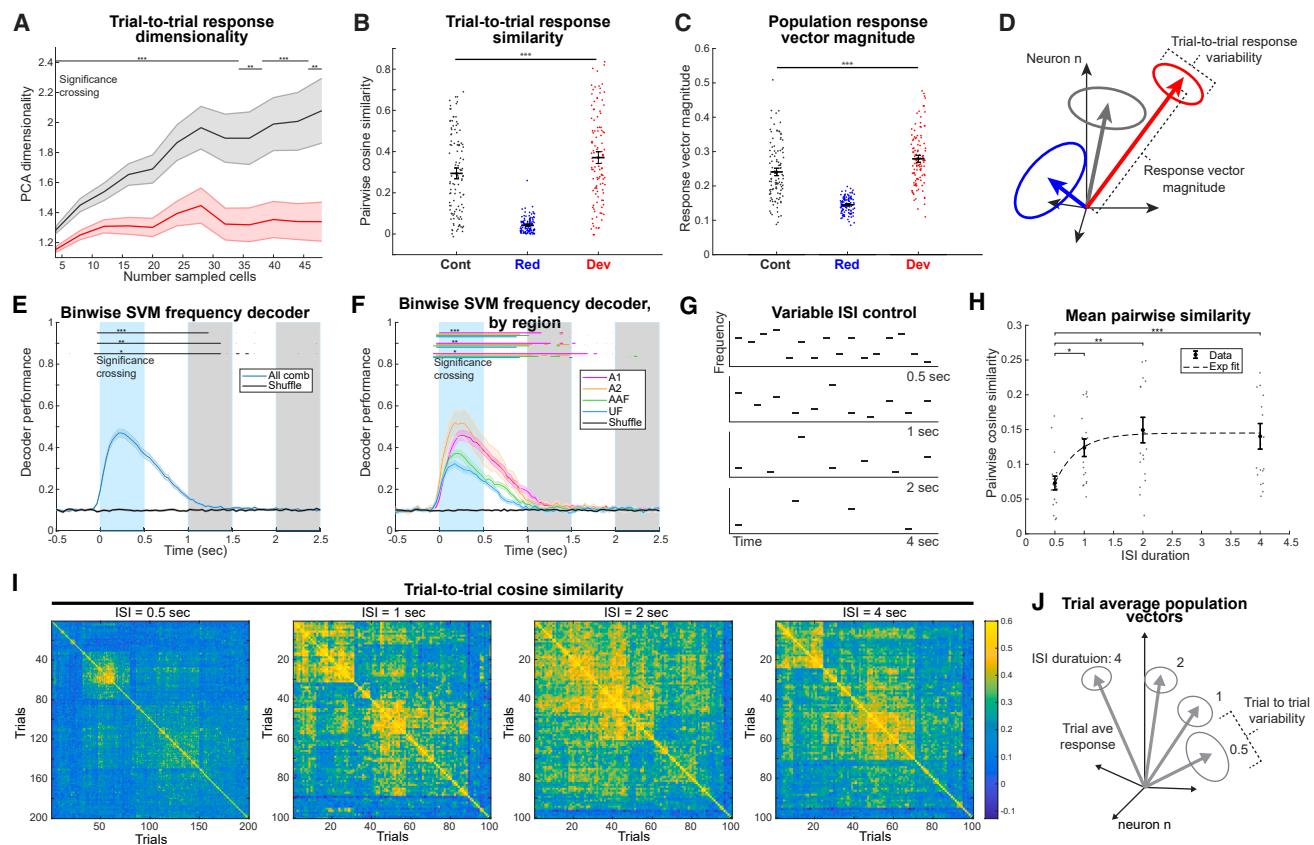


Figure 3. Slow cortical dynamics correlate with trial-to-trial variability and encode stimulus history

- (A) Trial-to-trial response dimensionality as a function of subsampled cells per dataset. Redundant trials were excluded due to low cell number per dataset.
- (B) Pairwise trial-to-trial response similarity for context trials.
- (C) Population vector response magnitudes, normalized by dataset size.
- (D) Graphical depiction of context-dependent response magnitudes and trial-to-trial variability.
- (E) Bin-wise decoding of frequency identity (blue) during the current and into future trials (gray).
- (F) Same as (E), but split by areas.
- (G) Variable ISI control inputs used to measure the network response decay kinetics.
- (H) Mean pairwise trial-to-trial cosine similarity vs. ISI duration, fitted with an exponential function.
- (I) Example sorted trial-to-trial similarity matrices for increasing ISI durations.
- (J) Graphical depiction of response magnitude and trial-to-trial variability dependence on ISI.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Error bars are SEM across stimulus types and datasets. Extended statistics in Table S1.

indicative of a more complex input representation. Indeed, the population and trial average activity of neurons in oddball-trained RNNs exhibited MMN-like responses to oddball inputs, with suppression of redundant responses and amplification of deviant responses, compared with controls (Figure 4B). Additionally, we observed an increase in MMN magnitude with increasing frequency difference, a phenomenon that has been demonstrated in the cortex (Figure S9H).^{2,53,54}

RNN models develop slow dynamics and network historicity

We visualized the input representations of oddball-trained RNNs during oddball and control tasks by reducing activity dimensionality using singular value decomposition (SVD; Figures 4C and S10A). Singular components 1–4 encoded the network transitions during oddball onset and embedded

the redundant frequency space along a linear manifold. Interestingly, RNNs learned to transition slowly, moving from the “quiet period” into frequency-dependent “steady states” over multiple redundant trials, exhibiting slow dynamics. These slow transitions were unique to oddball-trained networks, enabling them to encode memory and process temporal context, whereas frequency-trained and untrained networks quickly returned to baseline after stimuli ended (Figure S10C). Ramping-like activity during redundant stimuli, reminiscent of slow network transitions, has been reported in parvalbumin (PV)-positive⁵⁵ and vasoactive intestinal peptide (VIP)-positive²² interneurons.

Owing to the slow dynamics, redundant frequency identity became the major driving factor in separating network states (Figures 4E and 4F), suggesting its dominant role in driving population activity during oddball. For example, the difference in

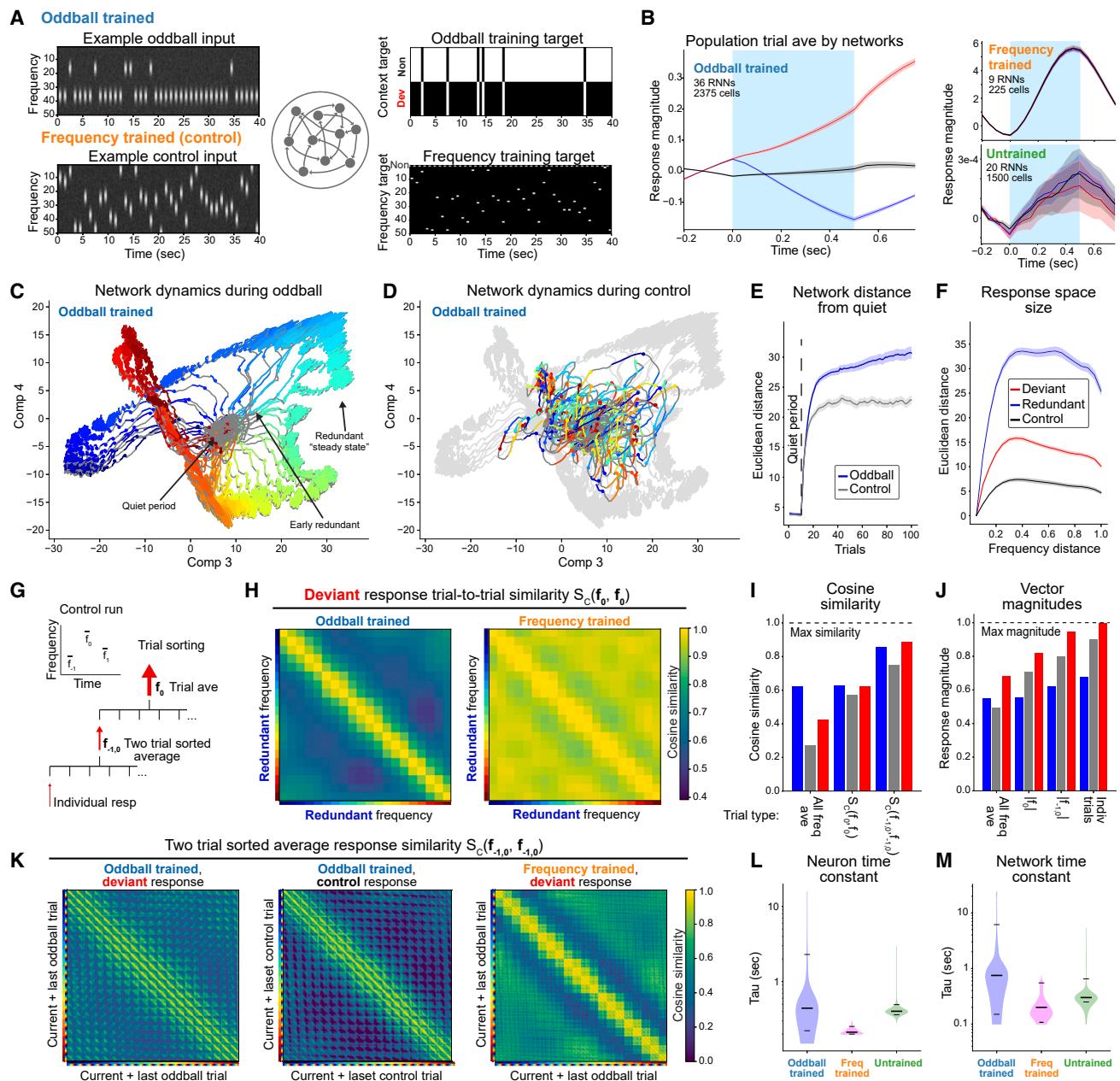


Figure 4. Artificial RNN models encode stimulus context and develop historicity and MMN

- (A) Depiction of network training types, inputs, and targets.
- (B) Population and trial average responses exhibit MMN in oddball-trained RNNs.
- (C) SVD-projected activity of an example oddball-trained network initialized with quiet period (gray) and followed with repeated redundant stimuli for the full frequency range (colors), showing slow transitions into steady states.
- (D) Control responses projected onto oddball singular vectors from (C) show high variability (data from C in gray).
- (E) Mean Euclidean distance networks travel away from the quiet initial state with redundant and control inputs.
- (F) Sizes of frequency space manifolds under each context.
- (G) Graphical description of trial-averaging procedure using current identity (f_0), two-trial identity ($f_{1,0}$), and individual trials.
- (H) Cosine similarity between deviant responses to the same frequency and as a function of preceding redundant frequency, for oddball- (left) and frequency-trained (right) RNNs.
- (I and J) Trial-to-trial cosine similarity and vector magnitudes using different averaging methods.
- (K) Cosine similarity between responses labeled by both current and previous frequency trials ($f_{-1,0}$) for oddball-trained oddball input (left), oddball-trained control input (middle), and frequency-trained oddball input (right) datasets.
- (L and M) Network and neuron time constants for training types under oddball inputs.
- Error bars represent SEM across networks.

redundant frequency induced a cosine angle between populations responding to identical deviant stimuli (Figure 4H).

For control inputs, slow dynamics caused the activity to be unreliable, with the initial states of trials depending on the past stimulus sequence (Figures 4D and S10B). Sorting trials by both the current and previous frequency identities and plotting cosine similarity matrices for oddball and control inputs revealed a memory-dependent structure in oddball-trained RNNs but not in frequency-trained or untrained RNNs (Figure 4K). Furthermore, analyzing trial-to-trial cosine similarity and vector response magnitudes while ignoring prior trial identity led to increased trial-to-trial variability and reduced the trial-averaged response magnitude (Figures 4I and 4J), potentially explaining *in vivo* trial-to-trial differences.

Finally, to directly measure the “slow dynamics” of RNNs, we computed the response decay duration of both the networks and neurons. The network signal was calculated as the Euclidean distance of the population from the initial state. We computed autocorrelations of network and neuronal activities during oddball inputs and measured the time constants of their decay^{56,57} (Figures 4L and 4M). Oddball-trained networks exhibited an increased time constant, likely leveraging this to encode memory and detect novelty.

DISCUSSION

We used two-photon calcium imaging to investigate how novelty detection is represented by populations of AC neurons in awake mice, passively listening to an auditory oddball paradigm with simple- and complex-featured stimuli. All primary and secondary areas exhibited robust context-dependent modulation to both simple and complex stimuli, selectively amplifying responses to deviant events and suppressing redundancies, demonstrating MMN. Underlying the amplified population signal were a greater number of responsive neurons and increased signal magnitudes and reliabilities. Simple tones demonstrated a hierarchical organization of novelty signals⁴⁹ and response abstractness¹ from A1 to A2. Complex stimuli generated a distinct temporal response profile and novelty representation across AC,⁴² likely shaped by the differences in stimulus statistics. Additionally, distinct neuronal populations were engaged throughout the onset and offset periods,^{36–39} indicating that network dynamics traversed multiple population states.

Discrete ensembles selectively encoded stimulus identity and temporal context, separating both the spatial (i.e., auditory frequencies) and temporal features along distinct directions within the high-dimensional neuronal space. Novelty responses, which appeared as linear modulations in the averaged activity, reflected feature tuning in the population space. That is, neuronal responses were not simply augmented or suppressed by context; rather, discrete ensembles were activated throughout each trial, simultaneously encoding stimulus identity, trial structure, and the three temporal contexts.^{25,58} This multidimensional encoding can be interpreted as the spatiotemporal cortical model of inputs under the predictive coding framework, with transitions between context-dependent states driving changes in response precision, while predictions and errors (novelty) are computed through linear combinations of activity.^{5,59,60}

To further investigate the effects of temporal context on population activity, we analyzed single-trial activity during the oddball task. Interestingly, control trial responses showed increased trial-to-trial variability, compared with deviant ones. We suspected that this was due to network historicity, whereby activity from prior stimuli persisted^{51,52} and influenced future responses.⁶¹ Indeed, a linear decoder confirmed that stimulus information persisted across trials. Extending the interval between stimuli in control runs reduced trial history effects and response variability. These findings suggest that stimulus response dynamics in the cortex are slow, propagating through multiple populations and influencing responses to incoming stimuli. This enables the network to store local memories and to perform temporal computations, such as novelty detection.

To test whether recurrent cortical circuits use slow dynamics and historicity to solve temporal tasks, we trained artificial RNNs to perform an oddball task. The RNN models also developed slow dynamics, demonstrating the persistence of stimulus information over time and an increase in trial-to-trial variability during control inputs, recapitulating *in vivo* results. Furthermore, RNNs revealed the spatiotemporal input representation that stores temporal context information in discrete population states. This representation utilized stimulus-specific steady states, induced by sequences of repeated redundants, and demonstrated amplified MMN-like transitions during novel events, representing memory storage and novelty detection.

Our RNN model demonstrates that recurrent network connectivity is sufficient to explain the biological phenomenology and that slow dynamics can serve as a mechanism for performing temporal context computations. In fact, the idea that recurrent connectivity can enable network reverberation, causing sustained responses to stimuli, was first proposed by Lorente de Nò in 1938.⁶² The propagation of activity along a network depends on its synaptic structure, making it plausible that NMDA-receptor blockade can disrupt these slow dynamics.^{63–65} Indeed, past work has suggested that MMN is generated by such synaptic processes.^{5,63} Moreover, it is possible that using their slow dynamics, all cortical areas perform temporal normalization as a canonical computation,¹³ serving as “memory machines” that encode events and process different types of information at varying levels of abstraction.¹

The slow dynamics model can also explain SSA, which is challenging to explain mechanistically at the single-neuron level,¹⁷ requiring stimulus-specific depletion factors.⁶⁶ We show that a recurrent network with slow dynamics naturally generates SSA, without any synaptic or input changes (Figure 4). Stimulus-specific depletion manifests as stimulus-specific redundancy steady states. Furthermore, projecting such activity onto a space orthogonal to the redundant responses would result in the suppression of redundant activity, as observed *in vivo*, reducing the total firing rate and optimizing energy expenditure.

Past work in mouse visual areas suggests that MMN-like responses—and, in particular, deviance detection—are supported by feedback projections to the V1 from higher brain regions, as well as engagement of local inhibitory populations.^{25,67,68} While that mechanism is distinct from the slow recurrent processing we propose here, both mechanisms could work in conjunction, depending on computation timescale, task complexity, and attention

state. On the other hand, different mechanisms and results could stem from differences between auditory and visual areas, as the former may operate on faster timescales than the latter.

Our work thus highlights the role of recurrent circuit connectivity in the processing of novel stimuli as an emergent-level network computation. We propose that the context-specific separations between ensembles and the trial-to-trial variability differences reflect historicity and temporal computation in the cortical network. Moreover, since novelty detection can be interpreted as prediction error, the historicity of the slow dynamics can also provide a mechanistic underpinning for predictive coding frameworks.^{5,59,60} Finally, our results also echo recent descriptions of disrupted temporal responses in patients with SZ and other psychoses,^{57,69} potentially explaining how altered response kinetics can disrupt context processing in patients with SZ, impairing MMN as well as other basic perceptual and cognitive operations.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Yuriy Shymkiv (ys2605@columbia.edu).

Materials availability

This study did not generate unique reagents.

Data and code availability

Calcium imaging data are deposited on Dryad and analysis software is available on GitHub, with locations listed in the [key resources table](#). Any additional information is available upon request from the [lead contact](#).

ACKNOWLEDGMENTS

We thank past and present members of the Yuste lab for help and fruitful discussions and reviewers for helpful comments. This work is supported by the NINDS (RM1NS132981), NEI (R01EY035248), and NIMH (R01MH115900, F31MH122137, and 5T32MH018870).

AUTHOR CONTRIBUTIONS

Y.S., R.Y., and J.P.H. designed the project. Y.S. and J.P.H. conducted experiments. Y.S. analyzed the data. Y.S. and S.E. designed and analyzed the RNN model. Y.S. and R.Y. wrote the article. J.P.H. and S.E. edited the article. R.Y. assembled and directed the team and provided equipment, resources, and funding.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
 - Mice
- [METHOD DETAILS](#)
 - Surgery
 - Widefield imaging
 - Calcium imaging
 - Auditory stimuli generation
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)

- Widefield image analysis
- Image registration
- Preprocessing of two-photon data
- Tuning statistical significance
- Context responsive neurons
- Population and trial average responses
- Extended analysis of context responsive neurons
- Quantifying population dynamics
- Stimulus space analysis and embedding
- Context-dependent trial-to-trial variability
- Extended statistical tests
- RNN model

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.neuron.2025.01.011>.

Received: May 25, 2024

Revised: November 8, 2024

Accepted: January 15, 2025

Published: February 10, 2025

REFERENCES

1. DiCarlo, J.J., Zoccolan, D., and Rust, N.C. (2012). How Does the Brain Solve Visual Object Recognition? *Neuron* 73, 415–434. <https://doi.org/10.1016/j.neuron.2012.01.010>.
2. Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>.
3. Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nat. Neurosci.* 6, 391–398. <https://doi.org/10.1038/nn1032>.
4. Javitt, D.C. (2009). When Doors of Perception Close: Bottom-up Models of Disrupted Cognition in Schizophrenia. *Annu. Rev. Clin. Psychol.* 5, 249–275. <https://doi.org/10.1146/annurev.clinpsy.032408.153502>.
5. Garrido, M.I., Kilner, J.M., Stephan, K.E., and Friston, K.J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* 120, 453–463. <https://doi.org/10.1016/j.clinph.2008.11.029>.
6. Näätänen, R., Shiga, T., Asano, S., and Yabe, H. (2015). Mismatch negativity (MMN) deficiency: A break-through biomarker in predicting psychosis onset. *Int. J. Psychophysiol.* 95, 338–344. <https://doi.org/10.1016/j.ijpsycho.2014.12.012>.
7. Perez, V.B., Tarasenko, M., Miyakoshi, M., Pianka, S.T., Makeig, S.D., Braff, D.L., Swerdlow, N.R., and Light, G.A. (2017). Mismatch Negativity is a Sensitive and Predictive Biomarker of Perceptual Learning During Auditory Cognitive Training in Schizophrenia. *Neuropsychopharmacology* 42, 2206–2213. <https://doi.org/10.1038/npp.2017.25>.
8. Ethridge, L.E., Hamm, J.P., Shapiro, J.R., Summerfelt, A.T., Keedy, S.K., Stevens, M.C., Pearson, G., Tamminga, C.A., Boutros, N.N., Sweeney, J.A., et al. (2012). Neural Activations During Auditory Oddball Processing Discriminating Schizophrenia and Psychotic Bipolar Disorder. *Biol. Psychiatry* 72, 766–774. <https://doi.org/10.1016/j.biopsych.2012.03.034>.
9. Javitt, D.C., and Sweet, R.A. (2015). Auditory dysfunction in schizophrenia: integrating clinical and basic features. *Nat. Rev. Neurosci.* 16, 535–550. <https://doi.org/10.1038/nrn4002>.
10. Ross, J.M., and Hamm, J.P. (2020). Cortical Microcircuit Mechanisms of Mismatch Negativity and Its Underlying Subcomponents. *Front. Neural Circuits* 14, 13. <https://doi.org/10.3389/fncir.2020.00013>.
11. Fitzgerald, K., and Todd, J. (2020). Making Sense of Mismatch Negativity. *Front. Psychiatry* 11, 468. <https://doi.org/10.3389/fpsyg.2020.00468>.

12. Friston, K. (2012). The history of the future of the Bayesian brain. *NeuroImage* 62, 1230–1233. <https://doi.org/10.1016/j.neuroimage.2011.10.004>.
13. Carandini, M., and Heeger, D.J. (2011). Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* 13, 51–62. <https://doi.org/10.1038/nrn3136>.
14. Stefanics, G., Kremláček, J., and Cziger, I. (2014). Visual mismatch negativity: a predictive coding view. *Front. Hum. Neurosci.* 8, 666. <https://doi.org/10.3389/fnhum.2014.00666>.
15. Malmierca, M.S., Anderson, L.A., and Antunes, F.M. (2015). The cortical modulation of stimulus-specific adaptation in the auditory midbrain and thalamus: a potential neuronal correlate for predictive coding. *Front. Syst. Neurosci.* 9, 19. <https://doi.org/10.3389/fnsys.2015.00019>.
16. Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004). Multiple Time Scales of Adaptation in Auditory Cortex Neurons. *J. Neurosci.* 24, 10440–10453. <https://doi.org/10.1523/JNEUROSCI.1905-04.2004>.
17. Nelken, I. (2014). Stimulus-specific adaptation and deviance detection in the auditory system: experiments and models. *Biol. Cybern.* 108, 655–663. <https://doi.org/10.1007/s00422-014-0585-7>.
18. Näätänen, R., Jacobsen, T., and Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology* 42, 25–32. <https://doi.org/10.1111/j.1469-8986.2005.00256.x>.
19. Farley, B.J., Quirk, M.C., Doherty, J.J., and Christian, E.P. (2010). Stimulus-Specific Adaptation in Auditory Cortex Is an NMDA-Independent Process Distinct from the Sensory Novelty Encoded by the Mismatch Negativity. *J. Neurosci.* 30, 16475–16484. <https://doi.org/10.1523/JNEUROSCI.2793-10.2010>.
20. Lakatos, P., O'Connell, M.N., Barczak, A., McGinnis, T., Neymotin, S., Schroeder, C.E., Smiley, J.F., and Javitt, D.C. (2020). The Thalamocortical Circuit of Auditory Mismatch Negativity. *Biol. Psychiatry* 87, 770–780. <https://doi.org/10.1016/j.biopsych.2019.10.029>.
21. Gallimore, C.G., Ricci, D.A., and Hamm, J.P. (2023). Spatiotemporal dynamics across visual cortical laminae support a predictive coding framework for interpreting mismatch responses. *Cereb. Cortex* 33, 9417–9428. <https://doi.org/10.1093/cercor/bhd215>.
22. Bastos, G., Holmes, J.T., Ross, J.M., Rader, A.M., Gallimore, C.G., Wargo, J.A., Peterka, D.S., and Hamm, J.P. (2023). Top-down input modulates visual context processing through an interneuron-specific circuit. *Cell Rep.* 42, 113133. <https://doi.org/10.1016/j.celrep.2023.113133>.
23. Von Der Behrens, W., Bäuerle, P., Kössl, M., and Gaese, B.H. (2009). Correlating Stimulus-Specific Adaptation of Cortical Neurons and Local Field Potentials in the Awake Rat. *J. Neurosci.* 29, 13837–13849. <https://doi.org/10.1523/JNEUROSCI.3475-09.2009>.
24. Grimm, S., Escera, C., Slabu, L., and Costa-Faidella, J. (2011). Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain. *Psychophysiology* 48, 377–384. <https://doi.org/10.1111/j.1469-8986.2010.01073.x>.
25. Hamm, J.P., Shymkiv, Y., Han, S., Yang, W., and Yuste, R. (2021). Cortical ensembles selective for context. *Proc. Natl. Acad. Sci. USA* 118, e2026179118. <https://doi.org/10.1073/pnas.2026179118>.
26. Hamm, J.P., and Yuste, R. (2016). Somatostatin Interneurons Control a Key Component of Mismatch Negativity in Mouse Visual Cortex. *Cell Rep.* 16, 597–604. <https://doi.org/10.1016/j.celrep.2016.06.037>.
27. Zhou, Z.C., Huang, W.A., Yu, Y., Negahbani, E., Stitt, I.M., Alexander, M.L., Hamm, J.P., Kato, H.K., and Fröhlich, F. (2020). Stimulus-specific regulation of visual oddball differentiation in posterior parietal cortex. *Sci. Rep.* 10, 13973. <https://doi.org/10.1038/s41598-020-70448-6>.
28. Daigle, T.L., Madisen, L., Hage, T.A., Valley, M.T., Knoblich, U., Larsen, R.S., Takeno, M.M., Huang, L., Gu, H., Larsen, R., et al. (2018). A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. *Cell* 174, 465–480.e22. <https://doi.org/10.1016/j.cell.2018.06.035>.
29. Stiebler, I., Neulist, R., Fichtel, I., and Ehret, G. (1997). The auditory cortex of the house mouse: left-right differences, tonotopic organization and quantitative analysis of frequency representation. *J. Comp. Physiol. A* 181, 559–571. <https://doi.org/10.1007/s003590050140>.
30. Issa, J.B., Haeffele, B.D., Agarwal, A., Bergles, D.E., Young, E.D., and Yue, D.T. (2014). Multiscale Optical Ca²⁺ Imaging of Tonal Organization in Mouse Auditory Cortex. *Neuron* 83, 944–959. <https://doi.org/10.1016/j.neuron.2014.07.009>.
31. Romero, S., Hight, A.E., Clayton, K.K., Resnik, J., Williamson, R.S., Hancock, K.E., and Polley, D.B. (2020). Cellular and Widefield Imaging of Sound Frequency Organization in Primary and Higher Order Fields of the Mouse Auditory Cortex. *Cereb. Cortex* 30, 1603–1622. <https://doi.org/10.1093/cercor/bhz190>.
32. Wang, C., Jiang, Z.Y., Chai, J.Y., Chen, H.S., Liu, L.X., Dang, T., and Meng, X.M. (2024). Mouse auditory cortex sub-fields receive neuronal projections from MGB subdivisions independently. *Sci. Rep.* 14, 7078. <https://doi.org/10.1038/s41598-024-57815-3>.
33. Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B.L., Koay, S.A., Taxidis, J., Najafi, F., Gauthier, J.L., Zhou, P., et al. (2019). CalmAn an open source tool for scalable calcium imaging data analysis. *eLife* 8, e38173. <https://doi.org/10.7554/eLife.38173>.
34. Harms, L., Michie, P.T., and Näätänen, R. (2016). Criteria for determining whether mismatch responses exist in animal models: Focus on rodents. *Biol. Psychol.* 116, 28–35. <https://doi.org/10.1016/j.biopspsycho.2015.07.006>.
35. Bandyopadhyay, S., Shamma, S.A., and Kanold, P.O. (2010). Dichotomy of functional organization in the mouse auditory cortex. *Nat. Neurosci.* 13, 361–368. <https://doi.org/10.1038/nn.2490>.
36. Recanzone, G.H. (2000). Response profiles of auditory cortical neurons to tones and noise in behaving macaque monkeys. *Hear. Res.* 150, 104–118. [https://doi.org/10.1016/S0378-5955\(00\)00194-5](https://doi.org/10.1016/S0378-5955(00)00194-5).
37. Takahashi, H., Nakao, M., and Kaga, K. (2004). Cortical mapping of auditory-evoked offset responses in rats. *NeuroReport* 15, 1565–1569. <https://doi.org/10.1097/01.wnr.0000134848.63755.5c>.
38. Scholl, B., Gao, X., and Wehr, M. (2010). Nonoverlapping Sets of Synapses Drive On Responses and Off Responses in Auditory Cortex. *Neuron* 65, 412–421. <https://doi.org/10.1016/j.neuron.2010.01.020>.
39. Solyga, M., and Barkat, T.R. (2021). Emergence and function of cortical offset responses in sound termination detection. *eLife* 10, e72240. <https://doi.org/10.7554/eLife.72240>.
40. deCharms, R.C., Blake, D.T., and Merzenich, M.M. (1998). Optimizing Sound Features for Cortical Neurons. *Science* 280, 1439–1443. <https://doi.org/10.1126/science.280.5368.1439>.
41. Nelken, I. (2008). Processing of complex sounds in the auditory system. *Curr. Opin. Neurobiol.* 18, 413–417. <https://doi.org/10.1016/j.conb.2008.08.014>.
42. Walker, K.M.M., Bizley, J.K., King, A.J., and Schnupp, J.W.H. (2011). Multiplexed and Robust Representations of Sound Features in Auditory Cortex. *J. Neurosci.* 31, 14565–14576. <https://doi.org/10.1523/JNEUROSCI.2074-11.2011>.
43. Seung, H.S., and Lee, D.D. (2000). Cognition. The Manifold Ways of Perception. *Science* 290, 2268–2269. <https://doi.org/10.1126/science.290.5500.2268>.
44. Chung, S., and Abbott, L.F. (2021). Neural population geometry: An approach for understanding biological and artificial neural networks. *Curr. Opin. Neurobiol.* 70, 137–144. <https://doi.org/10.1016/j.conb.2021.10.010>.
45. Cunningham, J.P., and Yu, B.M. (2014). Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.* 17, 1500–1509. <https://doi.org/10.1038/nn.3776>.
46. Gallego, J.A., Perich, M.G., Miller, L.E., and Solla, S.A. (2017). Neural Manifolds for the Control of Movement. *Neuron* 94, 978–984. <https://doi.org/10.1016/j.neuron.2017.05.025>.

47. Tenenbaum, J.B., de Silva, V., and Langford, J.C. (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* 290, 2319–2323. <https://doi.org/10.1126/science.290.5500.2319>.
48. Gao, P., Trautmann, E., Yu, B., Santhanam, G., Ryu, S., Shenoy, K., and Ganguli, S. (2017). A theory of multineuronal dimensionality, dynamics and measurement. Preprint at bioRxiv. <https://doi.org/10.1101/214262>.
49. Nieto-Diego, J., and Malmierca, M.S. (2016). Topographic Distribution of Stimulus-Specific Adaptation across Auditory Cortical Fields in the Anesthetized Rat. *PLoS Biol.* 14, e1002397. <https://doi.org/10.1371/journal.pbio.1002397>.
50. Bizley, J.K., Walker, K.M.M., Silverman, B.W., King, A.J., and Schnupp, J.W.H. (2009). Interdependent Encoding of Pitch, Timbre, and Spatial Location in Auditory Cortex. *J. Neurosci.* 29, 2064–2075. <https://doi.org/10.1523/JNEUROSCI.4755-08.2009>.
51. Moshitch, D., Las, L., Ulanovsky, N., Bar-Yosef, O., and Nelken, I. (2006). Responses of Neurons in Primary Auditory Cortex (A1) to Pure Tones in the Halothane-Anesthetized Cat. *J. Neurophysiol.* 95, 3756–3769. <https://doi.org/10.1152/jn.00822.2005>.
52. Campbell, R.A.A., Schulz, A.L., King, A.J., and Schnupp, J.W.H. (2010). Brief Sounds Evoke Prolonged Responses in Anesthetized Ferret Auditory Cortex. *J. Neurophysiol.* 103, 2783–2793. <https://doi.org/10.1152/jn.00730.2009>.
53. Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalogr. Clin. Neurophysiol.* 62, 437–448. [https://doi.org/10.1016/0168-5597\(85\)90054-1](https://doi.org/10.1016/0168-5597(85)90054-1).
54. Tiitinen, H., May, P., Reinikainen, K., and Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature* 372, 90–92. <https://doi.org/10.1038/372090a0>.
55. Natan, R.G., Briguglio, J.J., Mwilaambwe-Tshilobo, L., Jones, S.I., Aizenberg, M., Goldberg, E.M., and Geffen, M.N. (2015). Complementary control of sensory adaptation by two types of cortical interneurons. *eLife* 4, e09868. <https://doi.org/10.7554/eLife.09868>.
56. Murray, J.D., Bernacchia, A., Freedman, D.J., Romo, R., Wallis, J.D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., et al. (2014). A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* 17, 1661–1663. <https://doi.org/10.1038/nn.3862>.
57. Wengler, K., Goldberg, A.T., Chahine, G., and Horga, G. (2020). Distinct hierarchical alterations of intrinsic neural timescales account for different manifestations of psychosis. *eLife* 9, e56151. <https://doi.org/10.7554/eLife.56151>.
58. Audette, N.J., and Schneider, D.M. (2023). Stimulus-Specific Prediction Error Neurons in Mouse Auditory Cortex. *J. Neurosci.* 43, 7119–7129. <https://doi.org/10.1523/JNEUROSCI.0512-23.2023>.
59. Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836. <https://doi.org/10.1098/rstb.2005.1622>.
60. Adams, R.A., Stephan, K.E., Brown, H.R., Frith, C.D., and Friston, K.J. (2013). The Computational Anatomy of Psychosis. *Front. Psychiatry* 4, 47. <https://doi.org/10.3389/fpsyg.2013.00047>.
61. Morcos, A.S., and Harvey, C.D. (2016). History-dependent variability in population dynamics during evidence accumulation in cortex. *Nat. Neurosci.* 19, 1672–1681. <https://doi.org/10.1038/nn.4403>.
62. De N, R.L. (1938). Analysis of the activity of the chains of internuncial neurons. *J. Neurophysiol.* 1, 207–244. <https://doi.org/10.1152/jn.1938.1.3.207>.
63. Javitt, D.C., Steinschneider, M., Schroeder, C.E., and Arezzo, J.C. (1996). Role of cortical N-methyl-D-aspartate receptors in auditory sensory memory and mismatch negativity generation: implications for schizophrenia. *Proc. Natl. Acad. Sci. USA* 93, 11962–11967. <https://doi.org/10.1073/pnas.93.21.11962>.
64. Umbricht, D., Schmid, L., Koller, R., Vollenweider, F.X., Hell, D., and Javitt, D.C. (2000). Ketamine-Induced Deficits in Auditory and Visual Context-Dependent Processing in Healthy Volunteers: Implications for Models of Cognitive Deficits in Schizophrenia. *Arch. Gen. Psychiatry* 57, 1139–1147. <https://doi.org/10.1001/archpsyc.57.12.1139>.
65. Schmidt, A., Diaconescu, A.O., Komterer, M., Friston, K.J., Stephan, K.E., and Vollenweider, F.X. (2013). Modeling Ketamine Effects on Synaptic Plasticity During the Mismatch Negativity. *Cereb. Cortex* 23, 2394–2406. <https://doi.org/10.1093/cercor/bhs238>.
66. Yarden, T.S., and Nelken, I. (2017). Stimulus-specific adaptation in a recurrent network model of primary auditory cortex. *PLoS Comput. Biol.* 13, e1005437. <https://doi.org/10.1371/journal.pcbi.1005437>.
67. Furutachi, S., Franklin, A.D., Aldea, A.M., Mrsic-Flogel, T.D., and Hofer, S.B. (2024). Cooperative thalamocortical circuit mechanism for sensory prediction errors. *Nature* 633, 398–406. <https://doi.org/10.1038/s41586-024-07851-w>.
68. Jordan, R., and Keller, G.B. (2020). Opposing Influence of Top-down and Bottom-up Input on Excitatory Layer 2/3 Neurons in Mouse Primary Visual Cortex. *Neuron* 108, 1194–1206.e5. <https://doi.org/10.1016/j.neuron.2020.09.024>.
69. Hamm, J.P., Gilmore, C.S., Picchetti, N.A.M., Sponheim, S.R., and Clementz, B.A. (2011). Abnormalities of Neuronal Oscillations and Temporal Integration to Low- and High-Frequency Auditory Stimulation in Schizophrenia. *Biol. Psychiatry* 69, 989–996. <https://doi.org/10.1016/j.biopsych.2010.11.021>.
70. Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L.F., Dalgleish, H., Carandini, M., and Harris, K.D. (2017). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. Preprint at bioRxiv, 061507. <https://doi.org/10.1101/061507>.
71. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. Preprint at arXiv. <https://doi.org/10.48550/ARXIV.1912.01703>.
72. Trivedi, U. (2017). Isomap. MATLAB Central File Exchange. https://www.mathworks.com/matlabcentral/fileexchange/62449-isomap-d-n_fcn-n_size-options.
73. Shymkiv, Y., and Yuste, R. (2023). Aberration-free holographic microscope for simultaneous imaging and stimulation of neuronal populations. *Opt. Express* 31, 33461–33474. <https://doi.org/10.1364/OE.498051>.
74. Malone, B.J., Scott, B.H., and Semple, M.N. (2007). Dynamic Amplitude Coding in the Auditory Cortex of Awake Rhesus Macaques. *J. Neurophysiol.* 98, 1451–1474. <https://doi.org/10.1152/jn.01203.2006>.
75. Picton, T.W., John, M.S., Dimitrijevic, A., and Purcell, D. (2003). Human auditory steady-state responses: Respuestas auditivas de estado estable en humanos. *Int. J. Audiol.* 42, 177–219. <https://doi.org/10.3109/14992020309101316>.
76. Grent-t-Jong, T., Brickwedde, M., Metzner, C., and Uhlhaas, P.J. (2023). 40-Hz Auditory Steady-State Responses in Schizophrenia: Toward a Mechanistic Biomarker for Circuit Dysfunctions and Early Detection and Diagnosis. *Biol. Psychiatry* 94, 550–560. <https://doi.org/10.1016/j.biopsych.2023.03.026>.
77. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>.
78. Linkert, M., Rueden, C.T., Allan, C., Burel, J.-M., Moore, W., Patterson, A., Loranger, B., Moore, J., Neves, C., MacDonald, D., et al. (2010). Metadata matters: access to image data in the real world. *J. Cell Biol.* 189, 777–782. <https://doi.org/10.1083/jcb.201004104>.
79. Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B* 57, 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Ai162D Mice	Jax	RRID:IMSR_JAX:031562
Slc17a7-IRES2-Cre Mice	Jax	RRID:IMSR_JAX:023527
Software and Algorithms		
CalmAn	Giovannucci ³³	RRID:SCR_021533; https://github.com/flatironinstitute/CalmAn
Suite2P	Pachitariu ⁷⁰	RRID:SCR_016434; https://github.com/MouseLand/suite2p
Pytorch	Paszke ⁷¹	RRID:SCR_018536; https://pytorch.org/
Isomap	Trivedi ⁷²	https://www.mathworks.com/matlabcentral/fileexchange/62449-isomap-d-n_fcn-n_size-options
Caiman_sorter	This paper	https://github.com/shymkivy/caiman_sorter
SLM_GUI	Shymkiv ⁷³	https://github.com/shymkivy/SLM_control_GUI
AC_2p_analysis	This paper	https://github.com/shymkivy/AC_2p_analysis
Motion_corr_Ys	This paper	https://github.com/shymkivy/motion_corr_Ys
Register_2p_to_wf	This paper	https://github.com/shymkivy/register_2p_to_wf
AC_mapping_analysis	This paper	https://github.com/shymkivy/AC_mapping_analysis
RNN	This paper	https://github.com/shymkivy/RNN
Deposited Data		
Calcium Imaging Data	This paper	Dryad database: https://doi.org/10.5061/dryad.xsj3tx9q6

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mice

All experiments were approved by Columbia University Institutional Animal Care and Use Committee, and conducted in accordance with the National Institutes of Health guidelines. We used transgenic TIGRE2.0 mice,²⁸ both male and female, with GCaMP6s under Vglut1 promoter (Ai162D x Slc17a7-IRES2-Cre[Vglut1]), aged 8–17 weeks ($n = 10$). We did not analyze the effect of sex on our results since it was included of the original study design and limitations in mouse number per gender group (i.e. there were more females than males).

Since C57 mice are known to start losing high-frequency hearing, for a fraction of the mice we plotted the number of identified tuned neurons as a function of the age of the experiment (6/10 mice for which data were available; Figures S2I and S2J). We observed a small correlation of decline in number of tuned neurons across all frequencies combined, which did not seem to be important for our experiments. Furthermore, to ensure that mice perceived sounds during oddball, we used mapping results to guide the selection of oddball frequencies, only choosing sounds that elicited strong widefield responses.

METHOD DETAILS

Surgery

All surgeries and imaging were performed over the left auditory cortex (AC). Mice were anesthetized with isoflurane (1–2%) administered through a vaporizer in oxygen. Craniotomy was centered over the (AC), 2.5 mm caudal and 5 mm lateral from Bregma. The location was first identified with a stereotactic injection, with the mouse head in an upright orientation, and marked with a skin marker. The 3–4 mm cranial window ended up over the squamosal bone, posteromedial to the jugal bone. Then, the mouse was rotated ~60° to continue the surgery. After this, we mounted a titanium head plate over the left hemisphere with a circular opening around the AC. The headplate was fixed with dental cement and allowed to dry. Using a dental drill, we carved out the outline of the cranial window in the skull, slightly smaller than the coverslip size, and carefully removed it with forceps. The open brain surface was infused with sterile saline solution throughout the surgery. After removing the skull and washing out any blood with saline, the dura was removed during every surgery, significantly improving the imaging resolution. Before mounting the coverslip, 1.5% low melting temperature agarose was poured onto the brain. A coverslip (0.17 mm thickness, 3–4 mm diameter) was applied over agarose and pressed onto the bone surrounding craniotomy. We made sure that the coverslip sat tightly on the brain surface, preventing pockets for the brain to bulge out

into. Tight mounting is also likely important for preventing excessive motion artifacts. Finally, we fixed the coverslip to the skull with dental cement. We let the mice recover for at least a couple of hours before performing widefield mapping of the AC areas.

Widefield imaging

Both widefield and two-photon imaging were performed using a Bruker Ultima In Vivo two-photon system. We built a custom stage with adjustable height and angle of the head bar holding the head plate, to maneuver the angle of the mouse cranial window with respect to the objective. The head-fixed mouse was positioned on a circular treadmill. The rotation speed of the treadmill was recorded with an LED/photodarlington pair (Honeywell S&C HOA1877-003).

Widefield one-photon calcium imaging was performed with widefield LED illumination (3.3 mm FOV) and camera-recorded calcium signals in response to pure tone auditory inputs. The LED light source (470 nm, ~1 mW/mm²) was coupled to the Bruker microscope for excitation. Data collection was carried out with 4x objective (0.1 NA) and Hamamatsu ORCA flash4.0 CMOS camera with a large chip (13.3 mm x 13.3 mm), resulting in an approximately 5 mm FOV. We used a Bruker orbital adapter to bend the light path ~60° clockwise in the roll axis. To ensure that the optical window was aligned perpendicularly to the light path, we used a high NA 4x (0.28 NA) objective with thin depth of field to align it by eye until the edges around the cranial window were simultaneously in focus. After aligning the cranial window angle, we switched to the low NA 4x (0.1 NA) objective and focused on 100–200 μm below the surface of the brain. Imaging was performed at 30 Hz with 512x512 pixel acquisition, which was further binned down by a factor of two along both axes during analysis. Voltage signals from the stimulus software were acquired using the PrairieView (Bruker) voltage recording module. Frame times were downloaded from the HCimage software controlling the camera.

Calcium imaging

As previously described,²⁵ before imaging, mice underwent head plate mounting and chronic cranial window surgeries over the AC. Then, we used Bruker Ultima In Vivo two-photon system with a two-photon Ti:sapphire (MaTi DeepSee) at 940nm for imaging. The laser was gated and power modulated with a Pockels cell (Conoptics 350-105, with 302 RM driver).

The mouse was set up on a custom stage with a circular treadmill, similar to widefield experiments. Two-photon point scanning was performed in 256x256 pixel format, at approximately 60 Hz and averaged online over two frames to yield 30 Hz output. Voltage recordings of electrical signals describing parameters from the stimuli and locomotion were recorded with PrairieView software. As described for widefield imaging, we aligned the optical path to the cranial window and switched to the 20x (0.95 NA) Olympus water immersion objective. We used ultrasound gel (Aquasonic Clear) diluted to 25–50% in water instead of using water for immersing the objective.

For trial-to-trial datasets for **Figures 3G–3J**, we used the same Bruker system, but now customized with a spatial light modulator (SLM) to do holographic multiplane imaging via remote focusing.⁷³ A 60 Hz scan was effectively converted to collecting planes 5 of calcium imaging data at 9.35 Hz (with 25 μm spacing between planes), with a 25x (1.05 NA) objective.

Auditory stimuli generation

All auditory stimuli were generated with custom MATLAB code available online (github.com/shymkivy/stim_scripts) to be used in conjunction with the TDT RZ6 Multi I/O Processor, MF1 multi-field magnetic speakers in the closed-field configuration, custom RPvdsEx circuits at 200 kHz. For all the studies we used sinusoidal amplitude-modulated (SAM) frequency tones as simple stimuli, or auditory frequency gratings as complex stimuli. SAM tones were generated with a TDT RPvdsEx circuit controlled by a MATLAB script to trigger stimulus onset, offset and carrier frequency identity. For complex frequency gratings, auditory waveforms were created in MATLAB and played using another RPvdsEx.

Simple SAM tones for both widefield mapping and oddball were created using 10 carrier frequencies (500 ms duration), starting from 2 kHz and increased in 1.5 octave steps up to 76.9 kHz (2, 3, 4.5, 6.75, 10.1, 15.2, 22.8, 34.2, 51.3, and 76.9 kHz). We then added 40 Hz sinusoidal amplitude modulation (multiplied by squared and normalized 40 Hz sine wave) to potentially increase sound salience⁷⁴ and enhance robustness of cortical responses.^{75,76} SAM tones were generated with TDT RPvdsEx circuit and controlled by a MATLAB script, where we specified stimulus start, stimulus end, carrier frequency and amplitude.

Auditory frequency grating waveforms (0.5 s duration) were generated in MATLAB and the sound was played through the RZ6 amplifier (**Figure S1B**). We synthesized the sounds by first generating an array of tones for all frequencies between 2–90 kHz with 0.0052 octave resolution. Then, we generated amplitude modulation waveform square profiles for each of 10 orientations (0–2 π) and multiplied them onto the waveform matrix. These sounds were summed and the total power was adjusted to control the volume output. Additionally, 5 ms cosine on- and off-ramps were added to the outer ends of the frequency gratings. The code for controlling the sequence of sounds or frequency gratings using the RZ6 amplifier is available in MATLAB. For frequency gratings, the paradigm structure was exactly the same, but instead, we chose from 10 gratings (corresponding to ± 3.5, 7.9, 14.6 and 33.0 octaves/s speeds, for gratings 2–5 and 7–10).

Widefield mapping was done with SAM tones, consisting of 400 trials, each with a 0.5 s stimulus duration and a 2 s ISI. On each trial stimuli were randomly sampled from 10 SAM tones. And played by controlling the RZ6 amplifier with custom MATLAB script (audio_ACmapping.m). In addition, we performed test mapping runs with pure tones to verify that the results were the same as with SAM tones.

For the oddball paradigm, we used SAM tones as “simple” stimuli and auditory frequency gratings as “complex” stimuli. For both simple and complex paradigms, we started from 10 stimuli representing the “stimulus spaces” (Figure 1F), consisting of SAM tones and auditory frequency gratings respectively. Aside from stimulus identity, the structures of the simple and complex oddball paradigms were identical. The oddball paradigm was divided into three blocks (Figure 1E), a control block, oddball block, and flipped oddball block. In all three blocks, stimuli were presented with a 0.5 s duration, a 0.5 s ISI period plus a random jitter of 0–50 ms (appended at end of each ISI). The control block was played first, and consisted of 400 trials, with one of the 10 stimuli randomly sampled on each trial, and expected stimulus probability of 10% (same as deviant stimuli). For the two oddball blocks, we selected two frequencies with sufficient separation (5.25 ± 0.86 octaves separation [mean \pm SD]), and used 600 trials (to ensure sufficient number of deviants). First of selected frequencies was assigned as redundant and second as deviant. Initial 20 trials were set as redundant, in order to prime the network for oddball events. On each following trial, we performed a coin flip, where there is 90% probability that redundant is played, and 10% probability that deviant was played. To prevent extended periods of redundants, and increase the number of deviants in oddball runs, we set the probability of deviant stimuli to increase after 20, 21, 22, and 23 continuous redundant trials, to 20%, 25%, 50% and 100% respectively. For the flipped oddball block, the identities of stimuli used for redundant and deviant were reversed. We paused for 60 s between each of the three blocks. Overall, the three oddball blocks placed seemingly identical stimuli in three different contexts, neutral control, deviant, and redundant.

For the variable ISI experiments, we conducted four distinct control block runs with simple tones from the paradigm described above. For the 0.5 s ISI, we used 2000 trials, while for the remaining 1 s, 2 s, and 4 s runs we used 1000 trials.

In the beginning of the study, the noise in the room was measured with a decibel meter at ~ 61 dB (4–30 kHz working range), and the average volume of sounds was set to 5 dB above that, by scaling the voltage in the RPvdsEx circuit. To ensure that sounds across full frequency range were indeed played in our setup, we used a factory-calibrated Piezoelectronics 378C01 microphone (10–100,000 Hz working range), to record pure frequency tones generated by the MF1 speaker under experimental conditions (closed-field and 4 V modulation amplitude; Figure S1A). Additionally, we did not perform individual tone volume calibration to equalize the volume across frequencies, so there was MF1 microphone-dependent volume variation across frequencies. For frequencies above 65 kHz, sound volume produced by MF1 microphone suffers (<https://www.tdt.com/component/mf1-multi-field-magnetic-speakers/>), and those frequencies were likely harder to perceive. Despite the frequency-dependent volume differences, as shown in our study, we observed significant responses to the vast majority of stimuli, mostly in the middle range of stimuli for tones (Figures S2D–S2G). Throughout the study we guided the selection of tones used in the oddball task with mapping results, selecting only those frequencies that elicited large widefield calcium responses. As a result, the majority of frequencies used were in the middle range (tones 2–8; between 3–34 kHz, with 5.25 ± 0.86 octaves separation [mean \pm SD]). For complex frequency gratings, since the sounds are wideband, we suspect stimuli recognition was equivalent for all gratings.

QUANTIFICATION AND STATISTICAL ANALYSIS

Widefield image analysis

The pipeline to analyze widefield data is available online (AC_mapping_analysis). Data were collected with HCimage software in raw format (.cxd typically lost fewer frames than.tif). Frame times were extracted from HCimage. Initially, the 30 Hz, 512x512 movie was binned down to 256x256 using averaging (for faster processing) and converted to TIFF format using ImageJ⁷⁷ and bioformat plugin.⁷⁸ First we created an average frame and cropped the movie to a smaller rectangular active region for faster processing. Then, using MATLAB, we performed 3D interpolation to fill in missing frames (usually lost to computer freezing, often in bursts, on average <5 frames missing per recording). Trials with multiple missing frames were discarded. We then used stimuli onset times to sort the data into a matrix of size KxMxN (K = number of pixels; M = number of trials per stim; N = number of frames per trial). 300 ms before stimulus onset was used as the baseline and response was tracked for 1.5 s after stimulus ended. For each stimulus, activity was averaged across all trials. To create a spatial image of responses to stimuli (Figures S1C and S1D), we averaged frames during the stimulus period (50 to 450 ms) and subtracted the frame average during baseline (-300 to 0 ms), resulting in a single 2D image for each frequency.

Image registration

To register the individual mouse widefield maps onto each other we used a custom MATLAB GUI (register_2p_to_wf). For each dataset we manually marked the peak response locations for each frequency, and for all four AC areas, and extracted those coordinates from each widefield image. Next, we used those coordinates to register the field of images across all mice with rigid affine transformation and estimated a transformation matrix.

An analogous process was performed for registering two-photon fields of view onto the widefield image, using common structural features, such as blood vessels, as markers instead of functional response locations. Before or after each two-photon experiment we captured an image of the surface of the brain at the same (x, y) locations. In the same GUI, we used the surface frames and widefield images to manually label anatomical structures that were present in both images (typically small vessels, or ends and turns of vessels). The coordinates were then used to perform a nonrigid affine transformation and compute the transformation matrix. The two transformation matrices allowed us to register the locations of single cells onto a common map across all datasets (Figures 1G and 2A).

We manually defined AC regions (A1, A2, AAF, and UF) using tonotopic gradients and prior literature.^{29–32} Finally, we used coordinates from single cells to assign them to the corresponding four auditory regions. The gradients appeared well-defined on average, but locally broke down into a more random pattern.³⁵ Examining local frequency ranges showed that A2 was enriched with low-frequency-responsive neurons, while UF was expectedly²⁹ devoid of them (Figures S2F and S2G). In addition, the number of tuned cells for both simple tones and complex frequency gratings was significantly higher in A2 and A1 compared to AAF and UF (Figures S2C and S2M), suggesting that those areas were more engaged.

Preprocessing of two-photon data

Two-photon movies were acquired in.tif format using PrairieView. First, movies were motion-corrected with a custom suite2P⁷⁰-based script (motion_corr_Y5). Then, we used CNMF³³ for calcium signal demixing. Important parameters to consider were neuron half-size (“gsig” set to 5–6, for zoom 1–1.2) and, for deconvolution, the “fudge factor” was set to 0.99 (lower values caused spikes overfitting to noise). CNMF was performed in patches and refined in one cycle.

Visualization of CNMF outputs and determination of cell quality thresholds were done with a custom GUI (github.com/shymkiv/caiman_sorter). We used the same signal-to-noise ratio (SNR) threshold of 4 and “r_value” of 0.6 (CalmAn output, which represents the fraction of total signal of region of interest (ROI) pixels that the demixed signal represents) for all datasets. CalmAn CNN predictions did not perform well with the 256x256 pixel datasets collected with 20x magnification and 0.95 NA objective, and consequently, they were not used. After the neuron traces were extracted for all cells, we converted the calcium trace to a firing rate proxy using a smoothed (100 ms SD Gaussian kernel smoothing), rectified, first derivative function. Finally, all firing rate traces were self-standardized, by dividing the activity by its maximal response magnitude, resulting in all activity ranging between a baseline of 0 (derivative-based deconvolution process removed the baseline) and maximal firing rate of 1.

Tuning statistical significance

Code for all analyses is available online (AC_2p_analysis). To estimate the statistical significance of tuning for the frequency and context trials we first computed trial-averaged responses for all the trials, including the 10 distinct control frequencies (or gratings for complex frequency gratings), deviant context for two of the frequencies, and redundants for the two frequencies (redundants 2–7 were pooled into same trial type). Each stimulus type typically contained at least 40 trials. To estimate the trial average response used the “peak detection” method, where for -0.5 to 2 s around each trial start time we estimated peak magnitude and location (peak magnitude was averaged over 100 ms around the peak location). To estimate the statistical significance of those responses, we used bootstrapping to generate null distributions of peak response magnitudes, by randomly resampling from 10 control and the two deviant trials (redundant trials were omitted due to the response being significantly smaller, to reduce false positives), and applying peak detection in the same process. This process was repeated 5000 times, and all the peak magnitudes were used to generate a null response distribution. We computed a threshold corresponding to 99.7% in the null distribution and all cell responses crossing that threshold were deemed significant (responsive cells). To calculate z-scores, we standardized the data by subtracting the null distribution mean and dividing it by the threshold for significance (we did not use standard deviation because null distribution was not normal).

In parallel to peak detection, we also tried to define stimulus response windows to estimate tuning (onset window [100–400 ms] and offset window [600–900 ms]), and averaged the response within those windows. We preferred the peak detection method because it did not assume the response time, and was able to simultaneously detect onset- and offset-responsive neurons.

Context responsive neurons

For all neurons, we computed z-scores for tuning to the three context trials (deviant, redundant and control), each for the two frequencies under regular and flipped oddball runs. Some neurons showed exclusive tuning, while others were jointly tuned for more than one trial type. Neurons were defined as deviant-, control-, and redundant-tuned neurons when they had the highest z-scores for the corresponding contexts (Figure 2A). In the population and trial average plots (Figures 2B and 2C), the process appeared as linear modulation, where deviant and redundant responses seemed to be augmented or suppressed relative to control. At the population level, the process resembled feature tuning, where each context trial responded with a different population (Figure S7). Deviance and control populations partly overlapped, while the redundant population was orthogonal to the others. Despite the overlap between the deviant and control populations, there were cells with exclusive tuning that were not responsive to any other context.

Population and trial average responses

To create the population and trial average plots (Figures 2B and 2C), we selected all cells significantly tuned to any of the three contexts for the corresponding frequency (context-responsive cells) from each dataset. The data were combined across both oddball frequencies and across all datasets resulting in a 3D matrix NxMxT (N = number of cells; M = three contexts, T = number of trial frames, 0–950 ms of trial). Then, for each context trial, we plotted the average response with the standard error of the mean (over cells). Global MMN-like calcium activity corresponds to the difference between deviant and redundant traces (Figures S5D–S5F). To compute the statistical significance of deviance and redundancy signals compared to control, we averaged them over onset or offset response windows for simple tones, and applied a paired t-test (onset window 100–400 ms; offset window 600–900 ms). For complex frequency gratings, since the responses peaked in the middle of the trial, we averaged over middle response window

(300–600 ms). These results demonstrated the presence of a robust MMN in the mouse AC for both simple and complex stimuli, elicited with an auditory oddball paradigm, similar to the one used in human SZ studies.

For complex frequency gratings, we found that control and redundant traces in population and trial average plots are not significantly different in magnitude (Figure 2C). This is interesting and suggests that either adaptation during redundant trials is not as strong as during controls, or there is significant adaptation during control trials, or both. The lack of adaptation can potentially be explained by considering the frequency grating stimuli statistics, which are wideband stimuli where all frequencies are played together and modulated with different temporal profiles (Figure S1B). Unlike tones, the frequency grating features are constantly changing throughout stimulus presentation, and information from different features needs to be integrated over time and across frequencies. Therefore, it is plausible that constant frequency tone over 500 ms may elicit stronger adaptation than a dynamic frequency grating. Stronger adaptation during frequency grating control runs is also plausible. We showed that frequency grating-tuned neurons show not only orientation tuning but also sweep speed tuning (Figure 1M). Under these circumstances each grating does not only get adapted by itself, but also by the grating of similar sweep speeds. This would induce cross-stimulus adaptation and may have increased the adaptation-like effect during frequency grating control runs.

Extended analysis of context responsive neurons

To examine in greater detail what underlies context-dependent differences in population response curves, we analyzed cell counts and response magnitudes of neurons responsive to any of the three contexts. Across all imaged areas, for simple tones, we found almost twice the number of deviance detection neurons compared to their control counterparts, and more than ten times the number of deviance detection neurons compared to redundant detection neurons (Figure 2D). Analysis by area showed the most deviance detection neurons in A2, followed by A1, UF, and AAF (Figure S4D). Plotting the empirical cumulative distribution function (ECDF) of response z-scores showed that deviance detection neurons showed significantly larger responses compared to control or redundancy tuned neurons, suggesting that deviance-tuned neurons are not only most abundant, but also most reliable (Figures S4E and S4F). Consistent with previous data, A2 exhibited the best tuning for simple tone stimuli (Figure S4E). The z-scores of redundancy neuron responses were, on average, much lower than either deviant or control, largely due to low reliability.

Analyzing the cell counts and magnitudes for frequency grating contexts showed results analogous to those for tones (Figures 2E and S4G–S4J). The fractions of deviance detection neurons were also significantly higher than those of redundancy- or control-tuned neurons, with no significant area-specific differences. The response magnitude ECDF plots once again showed that deviance detection neurons on average have the highest z-scores and reliability (Figure S4K), with A1 being most engaged in deviance detection of frequency gratings (Figure S4L).

Quantifying population dynamics

We used cosine and Euclidean distances between population states to analyze dynamic changes throughout the trial and between trial types in our data. Cosine distance (or similarity, which equals 1 minus the distance) is a measure of population overlap between given states. On the other hand, Euclidean distance is a joint measure of population overlap, responsive neuron counts, and response size. One way to understand Euclidean distance is by examining the distance of a single response vector relative to zero. The Euclidean distance will simply be the vector magnitude, which, for three neurons, is $V(n_1, n_2, n_3) = \sqrt{(n_1^2 + n_2^2 + n_3^2)}$. From this definition, we see that the magnitude increases with number of neurons and scales with their response sizes. The distance between two response vectors is the magnitude of the difference vector between the two individual vectors $V_{diff} = V_1 - V_2$, which consequently depends on both neuron numbers and response magnitudes of individual vectors. Furthermore, if populations responding to two stimuli are overlapping, meaning the same neurons are being activated, the angle between the two vectors will be small (same idea as when taking dot products). Alternatively, if the populations are nonoverlapping, the angle should be orthogonal, and it will contribute to an increase in the Euclidean distance between the two responses. Thus, Euclidean distance between two populations is a complex measure that depends on cell number, response sizes, and population overlap. Because of similar reasoning we should not compare Euclidean distances between datasets that have different numbers of neurons. To address this, we used “normalized Euclidean distance” as a metric when comparing datasets with different numbers of neurons (Figures 2H, S7F, S7H, and S7N), where the neuron number-dependent differences are rescaled and the responsive cell fractions become more important.

To visualize the population activity during the full oddball paradigm, we first plotted a neuronal raster for an example dataset, displaying the activity of all neurons as a function of time (Figure S3A). Neurons were sorted according to their cosine similarity, while trials were sorted according to trial types and chronologically within each type. These sorted raster plots revealed ensembles of neurons that were reliably coactivated during respective frequency and context trials, representing the auditory stimulus space of inputs in the AC population response. During the control stimulus neuronal groups responded significantly and reliably to a single or local range of frequencies, representing their frequency tuning, with varying reliability. Response patterns during the redundant stimuli were far sparser, with smaller groups of neurons evoked and lower trial-to-trial reliability, consistent with the cell counts and response reliability data (Figures S3A, S4E, S4F, S4K, and S4L). The responses of neurons to deviant stimuli appeared similar to those of the control period for the same frequency but were notably more reliable.

To compute trial average multidimensional plots (Figures 2G, 2H, and S8E–S8H) we gathered the trial-averaged trace data and performed dimensionality reduction using PCA. For each dataset, we selected cells responding to any of the three contexts. We

concatenated the data from all datasets along the cell axis, resulting in a dataset with dimensions $N \times M \times T$ (N = number of cells; M = three contexts, T = number of trial frames). We reshaped the data into 2D matrix of dimensions $N \times (M^*T)$ and applied PCA. The top three principal components (PCs) were used to generate the 3D plots.

To quantify these PCA representations, we computed the dimensionality of the trial average context response manifold (Figures 2G and 2H). With only three PCs, we were able to encode 88.7% of trial-averaged data (participation ratio⁴⁸ — PR = 2.85, 2.51, 1.88, and 2.37 for A1, AAF, A2, and UF respectively), indicating that the manifold for tone context responses is less than three-dimensional. Next, we measured the Euclidean distances between each context type throughout the trial, from low-rank trial-averaged dataset projected onto 6 PCs (97.9% total variance; Figures S8A–S8D). The distances did not increase inclusion of additional components, as the trial-averaged data is virtually noise free. The distances between deviant and redundant trajectories represent the network MMN component, and showed that in A2 the deviant trials separate from redundant trials faster and reach a higher peak relative to all other regions, again highlighting the importance of A2 in context processing of tones (Figures S8A and S8M). In addition, during the offset period of trial, A1 had a larger separation between deviant and redundant or control and redundant trajectories, supporting the idea that A1 is highly engaged in offset deviance detection (Figures S8A–S8C and S8N). Finally, analyzing the distances between deviant and control trial indicated that A2 is most sensitive to processing neutral versus deviant contexts of tones (Figure S8C).

Stimulus space analysis and embedding

To visualize the full stimulus space of population activity during the control and oddball inputs, we analyzed trial-averaged population vectors, displaying the activity of all neurons as a function of time (Figures S3A–S3G). We used peak response magnitudes for all responsive cells, discarding the response latency in these analyses. These vectors were initially concatenated across all datasets and analyzed using PCA to extract the structure of a low-dimensional linear manifold (Figures S3E–S3G). The first 2 PCs separated middle-range frequencies, where the highest numbers of tuned neurons were found. PCs 3 and 4 separated low-end frequencies, while PCs 5 and 6 separated high-end frequencies. These results indicate that a relatively high number of PCs are needed to capture the trial-averaged data in linear frequency space (6 PCs were required to explain 83.3% of the total variance, out of 10 total dimensions; PR = 5.54; Figure S3D). This likely reflects the fact that frequency inputs were well-separated and mostly encoded by unique groups of neurons.

To embed the population vector data into Isomap⁷² space, similar in concept to PCA, we can embed either the different population vectors corresponding to trial types (Figures 1L and 1M) or the different signal vectors representing cell tuning similarities (Figures S3H and S3I). For the former case, we computed the cosine distances between population vectors for all input trials across all datasets, resulting in an $M \times M$ matrix where M is the number of trials used. This matrix was input into Isomap MATLAB code, that outputs embedding with a specified number of components. Similarly, for the signal vectors, we computed cosine distances between tuning of all cell pairs to corresponding trial types, resulting in an $N \times N$ matrix, where N is the number of neurons used. The distance matrix was then input into the Isomap code.

To visualize the range of different types of functionally tuned cells (as opposed to trials), we also applied Isomap embedding of the neuron signal vectors (trial-averaged response vectors), pooled across all datasets. A similar manifold shape appeared, representing the input space with all neurons scattered around it. The distances between neurons in the signal space correspond to the difference in their tuning profiles, indicating that there is a large individual neuron tuning variability, which, as a whole, encompasses the same frequency space manifold as trial average response space.

Context-dependent trial-to-trial variability

We believe that using trial-to-trial variability is a powerful method to measure the network stimulus decay duration and time constant indirectly (Figures 3G–3J). The direct method for measuring the time constant involves using a binwise decoder on population activity to detect how long stimulus memory persists (Figures 3E and 3F), but those results were not as clear, probably because our sample of neurons was not complete. To visualize the slow network responses, in sensory cortex that lacks pyramidal neurons with sustained activity, we would need to record from a complete sample of neurons capturing the full response and apply PCA. The individual neuronal responses together should represent a sequence of activations, while in the high-dimensional space it should appear as a slow trajectory moving through multiple populations.

Our hypothesis that a network needs slow dynamics to encode local stimulus history and perform novelty detection. The lack of novelty detection should correlate with faster recurrent time constants and reduced trial-to-trial variability in disease models (discussion section “Network deficits in schizophrenia”). From this perspective, trial-to-trial variability assay probes the mechanism that ultimately leads to novelty detection. It is possible that the trial-to-trial variability measure could be used as an alternative to oddball in some cases, where multiple neurons (or pixels in fMRI) are available for analysis.

Extended statistical tests

Kernel density estimation (KDE; Figures 1I and 1K) was used as a nonparametric smoothing method with a Gaussian kernel to estimate the continuous probability distribution.

A one-way ANOVA was conducted to compare distributions across groups using the F-statistic. If a significant result was found it was followed up by Fisher’s Least Significant Difference (LSD) Method and a two-tailed T-statistic with Benjamini-Hochberg⁷⁹ false discovery rate (FDR) adjustments to compute the significance of the difference in the means.

To compare the significance of cosine and Euclidean distances of a state compared to shuffled data (Figures 2H, S7G, and S7H) we used a paired t-test and followed by FDR correction. Specifically, for each dataset we used a total of 6 trials (3 contexts for 2 different oddball frequencies) and created 100 shuffled versions, where trial identity was permuted for each cell. In this case, the population states were randomized. For each shuffled vector, we computed a distance between each of 15 trial pairs, and computed the mean distance across shuffles. For each dataset we computed paired t-test for real distance and mean of shuffled dataset, to compute p-value. FDR correction was done for the 15 different pairs of trials.

RNN model

We used PyTorch library⁷¹ and Python to train recurrent neural networks (RNNs) and analyze the neurons in a manner analogous to how we analyzed *in-vivo* calcium imaging data (training and analysis code available at <https://github.com/shymkiv/RNN>). Initially we trained networks with 25 neurons, and then extended to 50, 75, 100, 150, and 250. The loss saturated after 100 neurons (Figures S9A and S9B), so for all multi-network analysis we pooled networks between 25–100 neurons. Analyzing networks with more neurons than required for the task typically led to part of the neurons not participating, washing out the population average activity. RNNs were created with input (W^{in} , 50xT), recurrent (W^{rec} , NxN), and output (W^{out} , Nx2 or Nx51) layers, where all three layers included a bias layer, and were jointly updated on each training iteration.

$$\frac{dr}{dt} = -r + f(W^{rec}r + b^{rec} + W^{in}u + b^{in})$$

$$\text{Update : } r_t = \left(1 - \frac{dt}{\tau}\right)r_{t-1} + \frac{dt}{\tau}f(W^{rec}r_{t-1} + b^{rec} + W^{in}u_t + b^{in})$$

$$\text{Output : } z = g(W^{out}r + b^{out})$$

Weights for all three layers were initialized with a normal distribution, with the standard deviation adjusted by the gain factor $g = 1$, $normal(0, \sqrt{g^2/N})$, where N is the number of neurons. The network timestep (dt) was set to 0.05 s, and time constant (τ) was set to 0.5 s. Furthermore, we used Rectified Linear Unit (ReLU) nonlinearity and crossentropy loss. The ReLU function is representative of biological neurons, with a linear response above zero and set to zero below zero. Oddball inputs were generated to consist of 20 trials, with 0.5 s stimulus duration and 0.5 s ISI, combining to 400 time steps (T) per training sample. Training sample iterations were performed with batches of 100 samples, and updated using the AdamW optimizer. For each sample neuronal activity was initialized with a uniform distribution between -1 and 1. The learning rate was kept constant throughout all iterations at 1e-3, and training typically reached a performance plateau after 1e6 to 2e6 iterations.

Auditory inputs to the network were created to represent a spectrogram of size 50xT, modeling 50 different frequencies separated logarithmically (Figure 4A). In order to create overlap between frequencies and model a linear input space, each individual frequency was spatially convolved with a Gaussian kernel with a standard deviation of 3 frequency bins. For oddball RNN inputs during training, for each sample we generated new inputs by randomly sampling 2 out of the 50 frequencies to be deviant and redundant, representing all possible combinations. The oddball trace was generated similarly to *in-vivo* experiments, with the redundant frequency played every 1 second at 90% probability, and interleaved with the deviant frequency at 10% probability. The oddball training target was same length as the input, T, and at each time step, it was set to 1 or 0, indicating whether the deviant stimulus was on or not. With this structure both the redundant stimuli and interstimulus intervals were labeled as non-deviant. We also tried to label specifically the redundants, as third input type, but this led to the activity being dominated by unwanted 1 Hz cycle representing stimuli turning on and off. Cross-entropy loss outputs were weighted with deviant = 0.95 and non-deviant = 0.05 of the loss, corresponding to the fractions of those bins in the target. For control inputs during frequency training we randomly sampled from all 50 frequencies. For frequency-trained networks the target was of size 51xT, requiring the identification all of the 50 frequencies and offset periods.

After training, all three networks—oddball-trained, frequency-trained, and untrained—were tested with both oddball and control inputs. For analysis, we treated the neuronal activity from RNN as neurons in our data, with a few key distinctions. First, since these were not spiking networks, the activity did not represent action potentials or firing rates; instead, it should be interpreted as synaptic currents or membrane voltage.

During inputs, unlike in real neurons, the absolute magnitude of neuronal activity is important, and baseline subtraction should not be applied. The network largely separates states based on the absolute magnitude of the activity, and key structure is encoded in these differences. Due to this, in population dynamics analysis PCA should be avoided because of the mean subtraction step, and SVD should be used instead.

In population and trial average plots (MMN), we saw the network respond with large changes in activity to deviant trials, while those changes were suppressed during redundant trials (Figure 4B). We interpret these activity changes by considering that an oddball-trained RNN with slow dynamics can reach a “steady state” over multiple redundant inputs. This becomes a relatively low-energy state, and any further redundant inputs cause the network to move only a small amount. Whenever a different frequency trial is

presented (deviant), the new low-energy steady state is now far away, and the network starts to make large steps again towards the new target, before saturating with repeated inputs. Frequency-trained networks, on the other hand, do not develop slow dynamics, because their goal is current frequency discrimination in control-like tasks, causing them to reach steady states quickly, throughout the 500 ms of stimulus presentation, and fall back to baseline during the 500 ms ISI period. The time constants of frequency-trained networks are reduced compared to untrained networks, in order to optimize current frequency recognition (Figures 4L and 4M). Untrained RNN dynamics do not reflect inputs (Figure S9C), because the input layer has not been scaled with training to increase the signal of the data relative to network noise.