PV neurons improve cortical complex scene analysis by enhancing timing-based coding

Short title: PV neurons enhance cortical coding

Jian Carlo Nocon[1,2,3,4], Howard J. Gritton[5,6], Nicholas M. James[1,2,3,4], Xue Han[1,2,3,4], Kamal Sen[1,2,3,4,*]

[1]Neurophotonics Center, Boston University, Boston, Massachusetts, United States of America

[2]Center for Systems Neuroscience, Boston University, Boston, Massachusetts, United States of America

[3]Hearing Research Center, Boston University, Boston, Massachusetts, United States of America

[4]Department of Biomedical Engineering, Boston University, Boston, Massachusetts, United States of America

[5]Department of Comparative Biosciences, University of Illinois, Urbana, Illinois, United States of America

[6]Department of Bioengineering, University of Illinois, Urbana, Illinois, United States of America

*Corresponding author

Email: kamalsen@bu.edu (KS)

Author contributions

J.C.N. and H.J.G. performed all experiments. J.C.N. collected and analyzed the data. N.M.J. provided technical support and code for data analysis. X.H. and K.S obtained funding and supervised the study. J.C.N, H.J.G, X.H., and K.S. wrote the manuscript and contributed to the interpretation of the results.

Abstract

Cortical representations underlying a wide range of cognitive abilities, which employ both rate and spike timing-based coding, emerge from underlying cortical circuits with a tremendous diversity of cell types. However, cell-type specific contributions to cortical coding are not well-understood. Here, we investigate the role of parvalbumin (PV) neurons in cortical complex scene analysis. Many complex scenes contain sensory stimuli, e.g., natural sounds, images, odors or vibrations, which are highly dynamic in time, competing with stimuli at other locations in space. PV neurons are thought to play a fundamental role in sculpting cortical temporal dynamics; yet their specific role in encoding complex scenes via timing-based codes, and the robustness of such temporal representations to spatial competition, have not been investigated. Here, we address these questions in auditory cortex using a cocktail party-like paradigm; integrating electrophysiology, optogenetic manipulations, and a family of novel spike-distance metrics, to dissect the contributions of PV neurons towards rate and timing-based coding. We find that PV neurons improve cortical discrimination of dynamic naturalistic sounds in a cocktail party-like setting by enhancing rapid temporal modulations in rate and spike timing reproducibility. Moreover, this temporal representation is maintained in the face of competing stimuli at other spatial locations, providing a robust code for complex scene analysis. These findings provide novel insights into the specific contributions of PV neurons in cortical coding of complex scenes.

Introduction

The cerebral cortex is critical to perception, attention, decision-making, memory, and motor output. A fundamental question of how cortical circuits achieve these functions remains a central unsolved problem in systems and cognitive neuroscience. One line of questioning towards solving this problem has been to investigate how the cortical code is organized; specifically, to assess whether cortical coding is based on rate or spike timing[1]. Previous studies have demonstrated both rate and spike timing-based coding in cortex[2-4]. However, a mechanistic understanding of how cortical circuits implement such cortical codes is still missing.

A second line of questioning towards addressing this central problem has been to investigate the structure of cortical circuits using a combination of anatomy, physiology, and new technologies, e.g., optogenetics, to interrogate key cortical circuits and neuron types[5, 6] This concerted approach has allowed systems neuroscience to identify biophysical properties of specific cell types, e.g., parvalbumin-expressing (PV), somatostatin-expressing (SOM) and vasoactive intestinal peptide-expressing (VIP) neurons)[7]. However, the specific contributions of these diverse cell types to the cortical code remain poorly understood.

A potentially powerful strategy for unraveling cell-type specific contributions to cortical coding is to probe cortical circuits in the context of naturalistic, behaviorally relevant problems that require a significant cortical contribution. An important example of such a problem is complex scene analysis, e.g., recognizing objects in a scene cluttered with multiple objects. The brain displays an astonishing ability to perform complex scene analysis in everyday settings, an impressive feat yet to be matched by state-of-the-art machines. How cortical circuits and cell-types contribute to this powerful computational ability remains unclear.

70 PV neurons are a prominent group of inhibitory neurons in cortex. Several studies have
71 investigated the role played by PV neurons in the generation of oscillations[8] and spike
72 synchronization[9]. However, the influence of PV neurons influence on cortical coding at the
73 single neuron level is unknown. PV neurons may play a specific role in complex scene analysis.
74 Complex natural scenes are characterized by stimuli that are highly dynamic in time, e.g.,
75 communication sounds, faces, odors and vibrations, with competing stimuli at other locations in
76 space. The biophysical properties of PV neurons are well suited for rapid temporal processing[7],
77 suggesting they may play a critical role in encoding dynamic natural stimuli. Moreover, such a
78 temporal representation must be robust in the face of competing stimuli at other spatial locations.
79 These observations motivate several open questions: Do PV neurons play a critical role in timing-
80 based codes for dynamic naturalistic stimuli in complex scenes? Are such temporal representations
81 robust to competing stimuli at other locations in space? Here, we address these questions in
82 auditory cortex using a combination of high-density electrophysiology, optogenetic tools for
83 silencing of PV neurons, and a family of novel spike distance metrics[10, 11] to dissect specific
84 contributions of PV neurons to the cortical code.

85 The auditory cortex is well suited to investigate these issues. It is thought to play a key role
86 in solving the "cocktail party problem"[12, 13], one of the most impressive examples of complex
87 scene analysis. For example, at a crowded cocktail party, we can hear what a colleague is saying
88 among other speakers, music, and background noise. Auditory cortical neurons show both
89 feature[13] and spatial selectivity[14, 15] that may contribute to solving the cocktail party problem.
90 Previous studies have demonstrated that auditory cortical neurons can employ both rate and spike
91 timing-based codes[3, 4]; and provided insight into the roles of inhibitory neurons in shaping
92 frequency tuning[16-19], frequency discrimination[20, 21], adaptation[22], sparseness[23], and

gap encoding[24]. However, the contributions of inhibitory cell types to complex scene analysis via rate vs. timing-based codes is not currently known. Here, we integrate a cocktail party-like paradigm[14] with optogenetic silencing of PV neurons to investigate the specific contribution of PV neurons to rate vs. timing-based coding in mouse primary auditory cortex (A1). We find that PV neurons enhance rapid temporal modulations in firing rate and spike timing-based coding in A1; improving cortical discrimination of dynamic naturalistic sounds in a cocktail party-like environment. Moreover, this temporal representation is upheld when challenged by competing sounds at other spatial locations, providing a robust foundation for complex scene analysis.

Results

Investigating cortical coding in mouse A1 using a cocktail party-like paradigm.

To better understand cortical coding of complex scenes in a mouse model amenable to circuit interrogation using genetic tools, we adopted a cocktail party-like experimental paradigm[14] while recording from the auditory cortex of PV-Arch transgenic mice (see Methods, 'Recording sessions and data acquisition'). The recording set-up consisted of 4 speakers arranged around the mouse at 4 locations on the azimuthal plane: directly in front (0°), two contralateral (45° and 90°) and 1 ipsilateral (-90°). Target stimuli consisted of white noise modulated by human speech envelopes extracted from a speech corpus[25]. We utilized two target waveforms (target 1 and target 2) and a masker consisting of unmodulated white noise. Mice were exposed to target alone trials and all possible target-masker combinations from the 4 locations (Figs 1A-C).

Mouse A1 neurons show spatial configuration sensitivity between competing auditory stimuli.

We assessed cortical coding using neural discriminability and a variety of other quantitative response measures. Neural discriminability between the two targets (% correct) was computed both without the masker ("Clean" condition, Fig 1A); and with the masker, for all possible combinations of target-masker locations ("Masked" condition, Figs 1B-C). Values of 100% and 50% represent perfect discrimination performance and chance performance, respectively. We refer to the matrix of performance values from all speaker configurations as "spatial grids", and positions of high performance (>70%) as "hotspots". These hotspots represent locations of enhanced discriminability between the two targets using a spike distance-based classifier to determine how well target identity can be predicted given the spike train from that site based on dissimilarities in spike timing and instantaneous rate[14] (Fig 1D: see Methods, 'Neural discriminability using SPIKE-distance').

Fig 1A illustrates spike trains from an example site that shows high discriminability under both target-only conditions (Figs 1A and 1D: black); and for a specific spatial configuration in the presence of a competing noise masker (Figs 1B and 1D: red). In the masked condition, discriminability depends strongly on the spatial configuration of the target and masker, indicating that the response at this site is spatial configuration sensitive. (Figs 1C and 1D, red versus green).

We next quantified the proportion of neurons that displayed hotspots. We used automated detection and sorting algorithms to identify 135 units from n=9 animals used in this study[26, 27]. Consistent with previous studies from auditory cortex, we found 69 of these 135 units, or ~51% of

136 units, were significantly sound modulated in response to the auditory stimulus. Further analysis
137 revealed that 40 of these 69 units showed significant hotspots (≥70%) at one or more spatial
138 configurations tested or ~45% of all auditory responsive units. Using this criterion, we found a
139 total of 94 spatial hotspots in the clean condition and 50 spatial hotspots in the masked condition
140 from the 135 units recorded. This finding suggests that many individual sites recorded in A1 of
141 the mouse showed relatively high discrimination performance for specific spatial locations or
142 target-masker configurations, and that competing maskers generally lower discrimination
143 performance.

144

145 Hotspots are present throughout the cortical column.

146 We next assessed hotspot locations relative to depth in the cortical column. Because
147 electrodes were chronically implanted to target recording sites across all cortical layers, we could
148 use current source density analysis to estimate laminar depth using the LFPs recorded from eight
149 electrode contacts within the same shank as described previously[28] (S1 Fig). Analysis revealed
150 we had strong multi-unit activity representation across supragranular, granular, and infragranular
151 regions across the four recording shanks (S1B-C Fig). We further compared the number of sites
152 with hotspots across the cortical column and the absolute level of discrimination performance from
153 those sites (Figs 1E-F, S2 and S3A Fig). There was no evidence that performance was better under
154 clean conditions across layers, but we did find that in the presence of a competing masker, layer
155 2/3 neurons showed significantly better performance than layer 6 neurons (Figs 1F and S2 Fig).
156 We also analyzed the cumulative distribution functions of performance by layer (S3B Fig). Using
157 this alternative measure, we found that L2/3 and L4 were significantly different from one another

158 under masked conditions. Combined, these analyses suggest that supragranular layers show
159 hotspots with higher discrimination performance compared to granular or infragranular layers.
160

161 Silencing of PV neurons reduces discrimination performance at 162 hotspots.

163 To investigate whether PV interneurons play a critical role in auditory discrimination
164 performance, we compared responses at hotspots, with and without optogenetic silencing of PV
165 neurons in A1. As seen in the individual examples in Fig 2A, and described in detail below, we
166 found that PV silencing generally increased the baseline firing rate, thereby reducing the
167 modulation depth during stimuli. To better understand how optogenetic silencing impacted
168 performance at hotspots, we compared performance across sites in the presence or absence of
169 optogenetic silencing. We found that hotspot performance decreased in both clean and masked
170 conditions (Fig 2B, S2 Fig). We next analyzed the data set using a spike distance-based classifier
171 to compare spike timing, which also revealed the effects of PV silencing on spike timing-based
172 discrimination under both clean and masked conditions (Fig 2C).

173 We next wanted to examine if discrimination performance was correlated with other
174 response measures, i.e., the driven firing rate (FR) and temporal sparseness of activity of units in
175 the presence or absence of PV silencing. We found that the driven FR increased significantly in
176 both the clean and masked condition (S4A Fig). Temporal sparseness is a measure of the
177 distribution of a unit's response over time, and sparse coding is thought to increase the efficiency
178 of encoding by minimizing the number of spikes necessary to convey information about the
179 stimulus[29]. We found there was a significant reduction in temporal sparseness for both the

cleaned and masked conditions (S4B Fig). These results indicate that PV silencing strongly affects both firing rate and temporal sparseness during the stimulus. We found no significant relationship between driven FR and performance in either the optogenetic ($r = 0.01480$, $p = 0.8603$) or the non-optogenetic ($r = 0.004741$, $p = 0.9550$) conditions. Therefore, although optogenetic silencing elevates firing rate and reduces performance, there was no direct evidence of a relationship between firing rate and performance. When we considered how sparseness was correlated with performance, we did find a weakly positive relationship between sparseness and performance under both control ($r = 0.1730$, $p = 0.03810$) and laser ($r = 0.1772$, $p = 0.03363$) conditions. These findings indicate that sparseness is strongly modified by PV silencing and is weakly correlated with performance.

Reduced discrimination performance driven by changes in rapid firing rate modulation and spike timing.

To determine the extent to which changes in rapid firing rate modulation, spike timing, and average firing rate changes that occur during silencing might affect performance, we calculated different performance metrics using the same hotspots as shown in Fig 2. Specifically, we used inter-spike interval (ISI)-distance, rate-independent (RI)-SPIKE-distance, or spike count, as the basis for discriminability between spike trains. ISI-distance calculates the distance between two spike trains based on the average dissimilarity in instantaneous firing rate, while RI-SPIKE-distance measures spike timing dissimilarity between two trains without being sensitive to firing rate differences[10]. Spike count distance is the absolute difference in spike count between trains, effectively measuring differences in total firing rate. We found that both ISI-distance-based (Fig

202 3A) and RI-SPIKE-distance-based (Fig 3B) performances showed significant decreases with
203 optogenetic silencing. Spike count distance-based performance, however, did not show significant
204 changes in performance during the optogenetic condition (Fig 3C). These findings suggest that
205 discrimination based on both rapid rate modulation and spike timing-based coding are profoundly
206 influenced by optogenetic silencing of PV neurons, but discrimination based on average firing rate
207 or total spike count is not affected by PV silencing.

208 To investigate the contribution of firing rate modulation more closely, we quantified the
209 effects of optogenetic silencing on the floor firing rate (S5A Fig) as well as the peak firing rate
210 (S5B Fig) during the auditory stimulus. Previous work has demonstrated that decreases in
211 spontaneous activity can change sound representations to promote sound perception during
212 behavior[30]. We found that under both clean and masked conditions, the firing rate of individual
213 hotspots increased during optogenetic silencing. In particular, we noted how the change in the
214 firing rate under optogenetic conditions enhanced the floor rate significantly above the near zero
215 rate seen in non-optogenetic trials. Therefore, we calculated the modulation index[31] (S5C Fig;
216 see Methods, 'Modulation index'), in the presence or absence of optogenetic silencing. We found
217 that the modulation index was significantly reduced under optogenetic conditions, suggesting PV
218 neurons improve discrimination by enhancing the modulation depth of responses.

219 Overall, these findings demonstrate that the observed decrease in discrimination
220 performance upon PV silencing is due to effects on rapid temporal modulations in firing rate and
221 spike timing, and not due to changes on global rate measures, i.e., average rate or total spike count
222 (Fig 3D).

223

Silencing of PV neurons reduces discrimination performance by increasing rate modulation similarity between auditory targets and reducing spike timing reproducibility within trials.

We further evaluated how PV silencing impacts specific components of discrimination performance. High neural discriminability is dependent both on differences in responses between targets, and reproducibility of responses within a target. To determine how optogenetic silencing independently affects these two factors, we calculated distances either between targets, or within the same target using the spike distance measures. We first compared spike trains between targets using the ISI-distance, which quantifies discriminability from instantaneous rate modulations. We found that optogenetic silencing of PV neural activity decreased discrimination primarily by decreasing the distance between targets (Figs 4A-B and S6A Fig). This effect was observed for both clean and masked trials (Fig 4Bii), suggesting that PV neurons improve cortical representation of auditory stimuli by enhancing the dissimilarity in rapid modulations in firing rate between targets.

We next considered how optogenetic silencing of PV neurons affected spike timing-based coding beyond measures of modulations in rate. To do so, we calculated the RI-SPIKE-distance between spike trains[10]. We found that PV silencing decreased discriminability primarily by increasing the distance within targets across trials (Figs 4A, 4C, and S6B Fig). This finding suggests that PV neurons improve cortical representations of auditory stimuli by enhancing spike timing reproducibility across trials, and that silencing PV neurons decreases this reproducibility. These results, in combination with the between-target measures, suggest that PV silencing

245 decreases performance by altering both dissimilarity between targets and within trials, and that the
246 effects of PV silencing are more profound in the presence of competing maskers.

247 Taken together, our results demonstrate that PV neurons improve cortical discrimination
248 of dynamic naturalistic stimuli in complex scenes by enhancing rapid temporal modulations in
249 firing rate and spike timing reproducibility. Moreover, this temporal representation is maintained
250 in the presence of competing stimuli at other spatial locations, providing a robust foundation for
251 cortical complex scene analysis.

252

253 Discussion

254 Diversity of cortical cell types and the cortical code.

255 One of the most striking features of the cerebral cortex is the tremendous diversity of its
256 cell types[32]. Understanding the computational role of such diversity in cortical coding is central
257 to systems neuroscience. Addressing this central question requires understanding cell type-specific
258 contributions to the cortical code at both the single neuron and the population level. A small
259 number of previous studies have demonstrated a role of specific cell types in cortical population
260 coding; specifically, the generation of oscillations[33, 34], and synchrony across cortical layers
261 and areas[35, 36]. However, cell type-specific contributions to the cortical code at the single and
262 multi-unit level, a fundamental aspect of the cortical code, remain poorly understood. In this study,
263 we addressed this fundamental gap by investigating the role of PV neurons in cortical coding of a
264 complex naturalistic scene, i.e., a cocktail party-like setting, in mouse A1.

265

Cortical discrimination & PV neurons; rate, spike timing and temporal codes.

We assessed cortical coding using neural discrimination and other quantitative measures. There is a rich history of quantitative work on cortical discrimination[37, 38]. This study extends this body of work in two significant ways. First, we assessed the contribution of a specific cell type, i.e., PV neurons, to cortical discrimination using optogenetic silencing of PV neurons. Second, we dissected the contributions of rate modulation, spike timing and spike count towards cortical discrimination, using a family of spike distance metrics designed to quantify these specific components. This combined approach revealed that optogenetic silencing of PV neurons generally degraded discrimination performance in both clean and masked conditions by decreasing rapid rate modulations between targets and spike timing reproducibility within targets. Overall, our findings suggest that PV neurons enhance timing-based coding via rapid modulations in firing rate and spike timing-based coding.

An influential review on neural coding defined the precise notion of a "temporal code" as one that contains information in spike timing beyond rate modulations[39]. Temporal codes have been challenging to identify because the contributions of rate vs. spike timing are often difficult to decouple. Our results based on spike distance measures which quantify both rate-dependent and rate-independent components of coding suggest that PV neurons also contribute to temporal coding in cortical discrimination. This computational portrait of PV neurons fits well with their established electrophysiological specializations, which support fast, efficient, and temporally precise responses[7].

A general cortical representation for the cocktail party problem across species.

From a comparative standpoint, we found that the key features present at the cortical level within A1 of the mammalian mouse was consistent with previous findings in songbirds. Specifically, a previous study by Maddox et al. found hotspots at particular spatial configurations of target and masker on the spatial grids of cortical level neurons in songbirds[14]. Songbirds and mice have different frequency ranges of hearing and therefore the cues used for spatial processing, e.g., interaural time difference (ITD), and interaural level difference (ILD) are frequency-dependent, and the peripheral representations of these cues are likely to be different across species with different frequency ranges of hearing. This suggests the emergence of general cortical representations for solving the cocktail party problem despite different peripheral representations of acoustic cues across species. In a cocktail party-like setting, multiple sound sources usually arise from different spatial locations. The position of hotspots at specific spatial configurations for single neurons, and the diversity of hotspot positions across the population, may enable cortical representation of targets for different target-masker configurations (S7 Fig).

Within channel vs. cross-channel inhibition.

Cortical inhibitory neurons can mediate feedforward, recurrent and di-synaptic feedback inhibition in cortical circuits, illustrated in a schematic model in Fig 5A. Previous modeling studies have demonstrated that feed-forward within-channel inhibition can improve discrimination performance[40]; whereas lateral inhibition across different channels can lead to the formation of

309 hotspots[41]. Our results suggest that PV neurons mediate within-channel inhibition,
310 corresponding to I neurons in the schematic model (Fig 5A). This is consistent with our finding
311 that although silencing PV neurons reduced discrimination performance, it did not completely
312 eliminate the presence of hotspots on the spatial grids. Moreover, many hotspots with the highest
313 discrimination performance before PV silencing were the locations with the highest level of
314 performance in the PV silencing condition (Fig 2Bii). This suggests that inhibition of PV neurons
315 does not eliminate the pattern of spatial configuration sensitivity in the spatial grids, but instead
316 more globally decreases encoding accuracy. Based on these observations, we hypothesize that a
317 separate cell-type (X neurons in Fig 5) mediate cross-channel lateral inhibition, resulting in the
318 generation of hotspots and the specific pattern of spatial configuration sensitivity on spatial grids.
319 A candidate cell type that correspond to X cells are somatostatin-positive (SOM) neurons, which
320 have been implicated in feedback lateral inhibition[42, 43] and surround suppression[44]. In this
321 context, it is also interesting to note that a small number of sites showed improved discrimination
322 upon optogenetic silencing of PV neurons (see Target 0˚, Masker 0˚ location in Fig 2B: 64%
323 compared to 69% under laser condition for example), although these represented a small fraction
324 of total hotpots (22 of 116 in the clean condition and 16 of 66 in the masked condition). Such
325 secondary hotspots may arise via complex network interactions which unmask other inhibitory
326 connections upon silencing of PV neurons (Fig 5B), a plausible scenario given the complex
327 network connectivity between excitatory pyramidal, PV and SOM neurons[45].

328

329 # Materials and methods

330 ## Subjects

331 All procedures involving animals were approved by the Boston University Institutional
332 Animal Care and Use Committee (IACUC). A total of 9 mice were used in this study. Original
333 breeding pairs of parvalbumin-Cre (PV-Cre: B6;129P2-Pvalbtm1(cre)Arbr/J), and Ai40 mice
334 (Arch: B6.Cg-*Gt(ROSA)26Sor*$^{tm40.1(CAG-aop3/EGFP)Hze}$/J) mice were obtained from Jackson
335 Laboratory, (Maine) and all breeding was done in house. Subjects consisted of both male and
336 female PV-Arch offspring 8-12 weeks old on the day of recording.

337

338 ## Surgery

339 Mice were surgically implanted with a head-plate as described previously[28, 46]. Briefly,
340 under isoflurane anesthesia, stereotaxic surgery was performed install a headplate, electrode, and
341 optical fiber. The custom head-plate was mounted anterior to bregma allowing access to AC
342 caudally. The headplate was anchored to the skull with 3 stainless steel screws and dental cement.
343 A fourth screw was connected to a metal pin and placed in the skull above the contralateral
344 cerebellum to serve as the reference. A craniotomy was made above the right auditory cortex (AP
345 -2.3 to -3.6, ML + 4.0 to +4.5, DV). Using a stereotaxic arm, a 32-contact linear probe
346 (Neuronexus, Ann Arbor, MI; model: A 4x8-5mm-100-400-177-CM32) with 100µm spacing
347 between electrode contacts and 400µm spacing between shanks, was positioned into AC,
348 perpendicular to the cortical surface. Because of the curvature of the AC cortical surface, not all
349 four shanks could be placed at precisely the same depth during each experiment. Probes were
350 advanced until all electrode contacts were within the cortical tissue and shanks were positioned
351 along the rostro-caudal axis of the AC (Fig 6A-C). An optical fiber, 200µm in diameter, was placed

352 medially to the 4 shanks and positioned between the two innermost shanks terminating at the
353 cortical surface (Fig 6A). After implantation, mice were allowed to recover for 4-7 days before
354 undergoing habituation to being head-fixed as described below.

355

356 Habituation

357 Following surgery and complete recovery, mice were first handled for several days before
358 being head-fixed to the recording apparatus. Mice were gradually exposed to longer restraint
359 periods at the same time of day as subsequent recording sessions. Each animal received at least 6
360 habituation sessions prior to the first recording day. Under head-fixed conditions, mice were
361 loosely wrapped with a piece of lab tissue (Kimwipes: Kimberly-Clark, Irving, TX) to encourage
362 reduced movement. At the conclusion of habituation, between three and five days of recording
363 sessions were done per animal and had the same duration as habituation sessions (Fig 6Di).

364

365 Recording sessions and data acquisition

366 All recordings were made with a Tucker Davis Technologies RZ2 recording system in an
367 electrically-shielded sound attenuation chamber. Broadband neural signals at 24414.0625 Hz were
368 recorded for each of the 32 channels. Local field potentials (LFPs) were band-pass filtered between
369 1 and 300 Hz, notch-filtered at 60 Hz, and digitized at 3051.8 Hz.

370 Recording sessions consisted of both non-optogenetic and optogenetic trials in random
371 order. Inter-trial interval was 5 seconds, with 3 seconds of stimulus playback followed by 2
372 seconds of silence (Fig 6Dii). Mice were exposed to target-alone ("clean") trials and target-masker

373 ("masked") combinations (Fig 6E). 10 trials were given per target identity for all possible
374 combinations of target location, masker location (including "clean" trials), and optogenetic
375 silencing of PV-expressing neurons. Thus, animals received a total of 800 trials per ~60 minute
376 recording session.

377

378 **Auditory stimuli**

379 All auditory stimuli were generated in Matlab and auditory stimuli consisted of either
380 target, masker or combination stimuli played from the four ES-1 electrostatic speakers. Target
381 stimuli consisted of white noise modulated in time by human speech envelopes extracted from
382 recordings of Harvard sentences. Masker stimuli consisted of unmodulated white noise. Before
383 speaker calibration, all stimuli were generated with the same RMS value, and sampling frequency
384 was 195312 Hz to capture the frequency range of hearing in mice. Stimuli were loaded onto a
385 custom RPvdsEx circuit on an RZ6 Multi I/O processor, which was connected to two PM2R
386 multiplexers that controlled the location of target and masker stimuli during trials.

387 During recordings, the stimuli were presented 18 cm from the mouse's head using four ES-
388 1 electrostatic speakers, driven by two ED-1 speaker drivers. The four speakers were arranged
389 around the mouse at four locations on the azimuthal plane: directly in front (0°), two contralateral
390 (45° and 90°) and 1 ipsilateral (-90°) (Fig 6F). Before recording sessions, stimulus intensity was
391 calibrated using a conditioning amplifier and microphone (Brüel and Kjaer, amplifier: model 2690,
392 and microphone model: 4939-A-011). For 7 of the 9 mice, all stimuli were at a measured 75dB
393 intensity at the mouse's head. For the remaining 2 mice, stimulus intensity was set to 70dB.
394 Stimulus playback lasted 3 seconds with a 1ms cosine ramp at onset and offset.

Optogenetic stimulation

Laser light for optogenetic stimulation of nucleus basalis was delivered through a multimode optically-shielded 200um fiber (Thorlabs, Newton NJ; model: BFH48-200), coupled to a 532nmDPSS laser (Shanghai Laser Ltd., Shanghai, China; model: BL532T3-200FC). Laser power was calibrated to 10mW at the fiber tip using a light meter calibrated for 532nm wavelength (PM100D, Thor Labs, Newton,926 NJ). The intensity was determined based on optogenetic cortical PV silencing studies using Archaerhodopsin from the literature[20, 47]. During optogenetic trials, the laser was turned on 50ms before stimulus onset and co-terminated with auditory stimuli (Fig 6Dii). Square light pulses lasting 3.05 s were delivered via TTL trigger from the RZ2 recording system to the laser diode controller (ADR-1805). Optogenetic trials were randomized throughout the recording session such that animals received all stimulus/masker pairs from each location with and without laser.

Data analysis

Spike extraction

Kilosort2 (https://github.com/MouseLand/Kilosort) was used to detect multi-units within the recordings[26]. Before spike detection and sorting, the broadband signal was band-passed between 300 and 5000 Hz using a 3rd-order Butterworth filter. Kilosort results were then loaded onto Phy2 (https://github.com/cortex-lab/phy), a graphical user interface for examining spike cluster information[27]. Clusters with artifact-like spike waveforms were deemed as noise, while

the remaining clusters were deemed as multi-units. The spikes toolbox (https://github.com/cortex-lab/spikes) was then used to import the multi-unit information from Phy to Matlab and extract spike waveforms from the high-passed signal[26]. To remove any remaining artifacts, all spikes with waveforms above an absolute threshold of 1500 mV were discarded.

Multi-units were deemed auditory-responsive if the mean z-score during non-optogenetic was larger than an absolute value of +1 between 10 and 30 ms after stimulus onset or during the entire 3 seconds of stimulus playback, based on the average response up to 1 second before stimulus onset. Multi-unit peri-stimulus time histograms (PSTHs) were calculated using 20 ms bins for plotting purposes.

Current source density analysis

Current source density (CSD) analysis estimates the second spatial derivative of LFP signals to determine the relative current across the cortical laminar depth. CSDs were calculated using LFPs, as described previously[28]. LFPs from control masked trials were used, as the rise time was more similar between target identities than clean stimuli. LFPs were low-passed filtered at 150 Hz before being down-sampled by a factor of 8 to 381 Hz. LFPs recorded from each contact were averaged across all trials prior to CSD calculation.

Before CSD estimation, channels that did not show an evoked response were interpolated using neighboring sites on the same shank. After this, LFPs were spatial smoothed across the eight channels in each shank:

$$\phi(z) = \frac{\phi(z + \Delta z) + 2\phi(z) + \phi(z - \Delta z)}{4}$$

437 where z is the depth perpendicular to the cortical surface, Δz is the electrode spacing, and Φ is the
438 potential. CSD was then estimated as:

439 $$CSD(z) = -\frac{\phi(z+\Delta z) - 2\phi(z) + \phi(z-\Delta z)}{\Delta z^2}$$

440 For display purposes, CSDs were linearly interpolated and plotted as pseudo-color images.

441 To determine the granular layer in each shank, CSD sink onset times were calculated. Sink
442 onset for each channel was defined as the time when the CSD goes below -3 times the standard
443 deviation of pre-stimulus activity. If more than one channel was found to have the earliest sink
444 onset, the channel whose neighbors had the earliest onsets was deemed the granular layer, or layer
445 4. This layer identification was used to anatomically classify the layer locations of channels in
446 each animal[48].

447

448 **Neural discriminability performance using SPIKE-distance**

449 Neural discrimination performance was calculated using a template-matching approach
450 similar to our previous study[14]. Briefly, spike trains were classified to one of the two target
451 stimuli based on whose template, one from each stimulus, yielded a smaller spike distance. For
452 each target-masker configuration, 10 iterations of template matching were done. Each iteration
453 consisted of 2000 template matches and left out one of the 10 spike trains in each target identity,
454 and the final value of neural discriminability performance is the average across all 10 iterations.
455 SPIKE-distance[10] calculates the dissimilarity between two spike trains based on differences in
456 spike timing and instantaneous firing rate without additional parameters. For one spike train in a
457 pair, the instantaneous spike timing difference at time t is:

$$S_1(t) = \frac{\Delta t_P^{(1)}(t) x_F^{(1)} + \Delta t_F^{(1)}(t) x_P^{(1)}}{x_{ISI}^{(1)}(t)}, t_P^{(1)} \leq t \leq t_F^{(1)}$$

where Δt_P represents the distance between the preceding spike from train 1 ($t_P^{(1)}$) and the nearest spike from train 2, Δt_F represents the distance between the following spike from train 1 ($t_F^{(1)}$) and the nearest spike from train 2, x_F is the absolute difference between t and $t_F^{(1)}$, and x_P is the absolute difference between t and $t_P^{(1)}$. To calculate $S_2(t)$, the spike timing difference from the view of the other train, all spike times and ISIs are replaced with the relevant values in train 2. The pairwise instantaneous difference between the two trains is calculated as:

$$S''(t) = \frac{S_1(t) + S_2(t)}{2\langle x_{ISI}^1(t), x_{ISI}^2(t)\rangle}$$

Finally, $S_1(t)$ and $S_2(t)$ are locally weighted by their instantaneous spike rates to account for differences in firing rate:

$$S(t) = \frac{S_1(t) x_{ISI}^2(t) + S_2(t) x_{ISI}^1(t)}{2\langle x_{ISI}^1(t), x_{ISI}^2(t)\rangle^2}$$

For a train of length T, the distance is the integral of the dissimilarity profile across the entire response interval, with a minimum value of 0 for identical spike trains:

$$D_S = \frac{1}{T}\int_0^T S(t)dt$$

cSPIKE (http://wwwold.fi.isc.cnr.it/users/thomas.kreuz/Source-Code/cSPIKE.html) was used to calculate all spike train distances between all possible spike train pairs for all spatial grid configurations[10].

Spatial grid configurations with performance above 70% during control trials were deemed "hotspots" and used for comparisons between control and optogenetic conditions. Configurations with at least 3 trials without spiking were excluded from performance calculations. To determine how rapid firing rate modulation, spike timing, and average firing rate contribute to discriminability, we used different distance measures as inputs to the template matching approach. For all hotspots, performances using ISI-distance, RI-SPIKE-distance, and spike count distance, the absolute difference in spike count between trains, were also calculated and compared to SPIKE-distance-based values.

ISI-distance

To determine how optogenetic silencing affects rapid temporal modulations in firing rate, ISI-distances were calculated. The ISI-distance calculates the dissimilarity between two spike trains based on differences in instantaneous rate synchrony. For a given time point:

$$I(t) = \frac{\left| x_{ISI}^{(1)}(t) - x_{ISI}^{(2)}(t) \right|}{\max\left(x_{ISI}^{(1)}(t), x_{ISI}^{(2)}(t)\right)}$$

This profile is then integrated along the spike train length to give a distance value, with values of 0 obtained for either identical spike trains or pairs with the same constant firing rate and a global phase shift difference.

RI-SPIKE-distance

To determine how optogenetic silencing affects reproducibility in spike timing between responses to the same stimulus, RI-SPIKE-distances between spike trains were calculated. The RI-SPIKE-distance is rate-independent, as it does not take differences in local firing rate between the two spike trains into account. From SPIKE-distance calculations, the final step of weighing $S_1(t)$ and $S_2(t)$ by their instantaneous spike rates is skipped, yielding:

$$S_{1,2}^{RI}(t) = \frac{S_1(t) + S_2(t)}{2\langle x_{ISI}^1(t), x_{ISI}^2(t)\rangle}$$

Like the other measures, the dissimilarity profile is integrated to give a distance value, with values of 0 obtained for two identical spike trains.

Layer and performance analysis

To determine if cortex layer had a significant effect on neural discriminability, one-way ANOVA was used to compare performance across layers during non-optogenetic trials. This analysis was done separately for clean and masked configurations. If the ANOVA yielded a significant effect from layers, post-hoc multiple comparisons were made between mean performances at different layers. Finally, Kolmogorov-Smirnov tests were used to determine if the cumulative distribution functions of performance were significantly different between layers. Bonferroni's correction was applied to account for multiple comparisons, and tests that yielded p-values below 0.0025 ($\alpha < 0.05/20$) were deemed significant.

Comparisons between control and optogenetic conditions

To determine the effects of optogenetic silencing on neural responses, paired t-tests between control and optogenetic conditions were done for neural discriminability and firing rate during stimulus playback. Like the ANOVA for performance and layers, these comparisons were separately done for clean and masked trials.

Sparseness

To determine how the sensitivity of neural responses changes with optogenetic silencing, temporal sparseness was analyzed. Given a PSTH with N bins and individual bins r_i:

$$S = \frac{N - \frac{(\sum_i^N r_i)^2}{(\sum_i^N r_i^2)}}{N - 1}$$

Sparseness is bound between 0 and 1, with 0 indicating a dense code and 1 indicating a sparse code. Bins of size 10 ms were used, and the driven responses of both target stimuli were concatenated in the time axis before calculations.

Floor and peak firing rate estimations and modulation index

To estimate the minimum firing rate in each response, the 20ms PSTH bin with the fewest spikes was used, and all trial snippets within that bin were resampled 100 times to obtain a 95% confidence interval for the minimum firing rate. This confidence interval was then used to find all PSTH bins that showed minimal spiking. A final estimate of minimum firing rate was obtained using 100 iterations of jackknife resampling on all trial snippets within these PSTH bins. This process was repeated for estimating peak firing rate.

Finally, to determine how optogenetic silencing affects the depth of firing rate modulation, modulation index[31] was calculated using the estimates for floor and peak firing rates:

$$MI = \frac{FR_{peak} - FR_{floor}}{FR_{peak} + FR_{floor}}$$

Acknowledgements

We would like to thank Alberto Cruz-Martin, Michael Economo, and Conor Houghton for comments and suggestions on the manuscript.

References

1. deCharms RC, Zador A. Neural representation and the cortical code. Annu Rev Neurosci. 2000;23:613-47. Epub 2000/06/09. doi: 10.1146/annurev.neuro.23.1.613. PubMed PMID: 10845077.

2. Zuo Y, Safaai H, Notaro G, Mazzoni A, Panzeri S, Diamond ME. Complementary contributions of spike timing and spike rate to perceptual decisions in rat S1 and S2 cortex. Curr Biol. 2015;25(3):357-63. Epub 2015/01/27. doi: 10.1016/j.cub.2014.11.065. PubMed PMID: 25619766.

3. Yao JD, Sanes DH. Temporal Encoding is Required for Categorization, But Not Discrimination. Cereb Cortex. 2021;31(6):2886-97. Epub 2021/01/12. doi: 10.1093/cercor/bhaa396. PubMed PMID: 33429423; PubMed Central PMCID: PMCPMC8107795.

554 4. Lu T, Liang L, Wang X. Temporal and rate representations of time-varying signals in the
555 auditory cortex of awake primates. Nat Neurosci. 2001;4(11):1131-8. PubMed PMID: 11593234.
556 5. Lohse M, Bajo VM, King AJ, Willmore BDB. Neural circuits underlying auditory contrast
557 gain control and their perceptual implications. Nat Commun. 2020;11(1):324. Epub 2020/01/18.
558 doi: 10.1038/s41467-019-14163-5. PubMed PMID: 31949136; PubMed Central PMCID:
559 PMCPMC6965083.
560 6. Guo W, Hight AE, Chen JX, Klapoetke NC, Hancock KE, Shinn-Cunningham BG, et al.
561 Hearing the light: neural and perceptual encoding of optogenetic stimulation in the central auditory
562 pathway. Sci Rep. 2015;5:10319. Epub 2015/05/23. doi: 10.1038/srep10319. PubMed PMID:
563 26000557; PubMed Central PMCID: PMCPMC4441320.
564 7. Tremblay R, Lee S, Rudy B. GABAergic Interneurons in the Neocortex: From Cellular
565 Properties to Circuits. Neuron. 2016;91(2):260-92. Epub 2016/08/02. doi:
566 10.1016/j.neuron.2016.06.033. PubMed PMID: 27477017; PubMed Central PMCID:
567 PMCPMC4980915.
568 8. Antonoudiou P, Tan YL, Kontou G, Upton AL, Mann EO. Parvalbumin and Somatostatin
569 Interneurons Contribute to the Generation of Hippocampal Gamma Oscillations. J Neurosci.
570 2020;40(40):7668-87. Epub 2020/08/30. doi: 10.1523/JNEUROSCI.0261-20.2020. PubMed
571 PMID: 32859716; PubMed Central PMCID: PMCPMC7531548.
572 9. Jang HJ, Chung H, Rowland JM, Richards BA, Kohl MM, Kwag J. Distinct roles of
573 parvalbumin and somatostatin interneurons in gating the synchronization of spike times in the
574 neocortex. Sci Adv. 2020;6(17):eaay5333. Epub 2020/05/20. doi: 10.1126/sciadv.aay5333.
575 PubMed PMID: 32426459; PubMed Central PMCID: PMCPMC7176419.

576 10. Satuvuori E, Kreuz T. Which spike train distance is most suitable for distinguishing rate
577 and temporal coding? J Neurosci Methods. 2018;299:22-33. Epub 2018/02/21. doi:
578 10.1016/j.jneumeth.2018.02.009. PubMed PMID: 29462713.
579 11. Kreuz T, Chicharro D, Houghton C, Andrzejak RG, Mormann F. Monitoring spike train
580 synchrony. J Neurophysiol. 2013;109(5):1457-72. Epub 2012/12/12. doi: 10.1152/jn.00873.2012.
581 PubMed PMID: 23221419.
582 12. Narayan R, Best V, Ozmeral E, McClaine E, Dent M, Shinn-Cunningham B, et al. Cortical
583 interference effects in the cocktail party problem. Nat Neurosci. 2007;10(12):1601-7. doi:
584 10.1038/nn2009. PubMed PMID: 17994016.
585 13. Mesgarani N, Chang EF. Selective cortical representation of attended speaker in multi-
586 talker speech perception. Nature. 2012;485(7397):233-6. Epub 2012/04/24. doi:
587 10.1038/nature11020. PubMed PMID: 22522927; PubMed Central PMCID: PMCPMC3870007.
588 14. Maddox RK, Billimoria CP, Perrone BP, Shinn-Cunningham BG, Sen K. Competing sound
589 sources reveal spatial effects in cortical processing. PLoS Biol. 2012;10(5):e1001319. Epub
590 2012/05/09. doi: 10.1371/journal.pbio.1001319. PubMed PMID: 22563301; PubMed Central
591 PMCID: PMCPMC3341327.
592 15. Middlebrooks JC, Bremen P. Spatial stream segregation by auditory cortical neurons. J
593 Neurosci. 2013;33(27):10986-1001. Epub 2013/07/05. doi: 10.1523/JNEUROSCI.1065-13.2013.
594 PubMed PMID: 23825404; PubMed Central PMCID: PMCPMC3718378.
595 16. Li LY, Ji XY, Liang F, Li YT, Xiao Z, Tao HW, et al. A feedforward inhibitory circuit
596 mediates lateral refinement of sensory representation in upper layer 2/3 of mouse primary auditory
597 cortex. J Neurosci. 2014;34(41):13670-83. Epub 2014/10/10. doi: 10.1523/JNEUROSCI.1516-
598 14.2014. PubMed PMID: 25297094; PubMed Central PMCID: PMCPMC4188965.

599 17. Li LY, Xiong XR, Ibrahim LA, Yuan W, Tao HW, Zhang LI. Differential Receptive Field
600 Properties of Parvalbumin and Somatostatin Inhibitory Neurons in Mouse Auditory Cortex. Cereb
601 Cortex. 2015;25(7):1782-91. Epub 2014/01/16. doi: 10.1093/cercor/bht417. PubMed PMID:
602 24425250; PubMed Central PMCID: PMCPMC4459283.
603 18. Kato HK, Asinof SK, Isaacson JS. Network-Level Control of Frequency Tuning in
604 Auditory Cortex. Neuron. 2017;95(2):412-23 e4. Epub 2017/07/12. doi:
605 10.1016/j.neuron.2017.06.019. PubMed PMID: 28689982; PubMed Central PMCID:
606 PMCPMC5705232.
607 19. Moore AK, Wehr M. Parvalbumin-expressing inhibitory interneurons in auditory cortex
608 are well-tuned for frequency. J Neurosci. 2013;33(34):13713-23. Epub 2013/08/24. doi:
609 10.1523/JNEUROSCI.0663-13.2013. PubMed PMID: 23966693; PubMed Central PMCID:
610 PMCPMC3755717.
611 20. Aizenberg M, Mwilambwe-Tshilobo L, Briguglio JJ, Natan RG, Geffen MN. Bidirectional
612 Regulation of Innate and Learned Behaviors That Rely on Frequency Discrimination by Cortical
613 Inhibitory Neurons. PLoS Biol. 2015;13(12):e1002308. Epub 2015/12/03. doi:
614 10.1371/journal.pbio.1002308. PubMed PMID: 26629746; PubMed Central PMCID:
615 PMCPMC4668086.
616 21. Briguglio JJ, Aizenberg M, Balasubramanian V, Geffen MN. Cortical Neural Activity
617 Predicts Sensory Acuity Under Optogenetic Manipulation. J Neurosci. 2018;38(8):2094-105.
618 Epub 2018/01/26. doi: 10.1523/JNEUROSCI.2457-17.2017. PubMed PMID: 29367406; PubMed
619 Central PMCID: PMCPMC5824743.
620 22. Natan RG, Briguglio JJ, Mwilambwe-Tshilobo L, Jones SI, Aizenberg M, Goldberg EM,
621 et al. Complementary control of sensory adaptation by two types of cortical interneurons. Elife.

622 2015;4. Epub 2015/10/16. doi: 10.7554/eLife.09868. PubMed PMID: 26460542; PubMed Central
623 PMCID: PMCPMC4641469.
624 23. Liang F, Li H, Chou XL, Zhou M, Zhang NK, Xiao Z, et al. Sparse Representation in
625 Awake Auditory Cortex: Cell-type Dependence, Synaptic Mechanisms, Developmental
626 Emergence, and Modulation. Cereb Cortex. 2019;29(9):3796-812. Epub 2018/10/12. doi:
627 10.1093/cercor/bhy260. PubMed PMID: 30307493; PubMed Central PMCID:
628 PMCPMC6686756.
629 24. Keller CH, Kaylegian K, Wehr M. Gap encoding by parvalbumin-expressing interneurons
630 in auditory cortex. J Neurophysiol. 2018;120(1):105-14. Epub 2018/03/29. doi:
631 10.1152/jn.00911.2017. PubMed PMID: 29589814; PubMed Central PMCID:
632 PMCPMC6093948.
633 25. Rothauser EH, Chapman, W. D., Guttman, N., Nordby, K. S., Silbiger, H. R., Urbanek, G.
634 E., & Weinstock, M. I.E.E.E. recommended practice for speech quality measurements. IEEE Trans
635 Audio Electroacoust. 1969;17(3):225-46. doi: 10.1109/TAU.1969.1162058. .
636 26. Pachitariu M, Steinmetz N, Kadir S, Carandini M, Harris KD. Kilosort: realtime spike-
637 sorting for extracellular electrophysiology with hundreds of channels. bioRxiv. 2016. doi:
638 10.1101/061481.
639 27. Rossant C, Kadir SN, Goodman DFM, Schulman J, Hunter MLD, Saleem AB, et al. Spike
640 sorting for large, dense electrode arrays. Nat Neurosci. 2016;19(4):634-41. Epub 2016/03/15. doi:
641 10.1038/nn.4268. PubMed PMID: 26974951; PubMed Central PMCID: PMCPMC4817237.
642 28. James NM, Gritton HJ, Kopell N, Sen K, Han X. Muscarinic receptors regulate auditory
643 and prefrontal cortical communication during auditory processing. Neuropharmacology.

644 2019;144:155-71. Epub 2018/10/24. doi: 10.1016/j.neuropharm.2018.10.027. PubMed PMID:
645 30352212; PubMed Central PMCID: PMCPMC6400225.
646 29. Vinje WE, Gallant JL. Sparse coding and decorrelation in primary visual cortex during
647 natural vision. Science. 2000;287(5456):1273-6. PubMed PMID: 10678835.
648 30. Buran BN, von Trapp G, Sanes DH. Behaviorally gated reduction of spontaneous discharge
649 can improve detection thresholds in auditory cortex. J Neurosci. 2014;34(11):4076-81. Epub
650 2014/03/14. doi: 10.1523/JNEUROSCI.4825-13.2014. PubMed PMID: 24623785; PubMed
651 Central PMCID: PMCPMC3951702.
652 31. Luo TZ, Maunsell JHR. Attentional Changes in Either Criterion or Sensitivity Are
653 Associated with Robust Modulations in Lateral Prefrontal Cortex. Neuron. 2018;97(6):1382-93
654 e7. Epub 2018/03/06. doi: 10.1016/j.neuron.2018.02.007. PubMed PMID: 29503191; PubMed
655 Central PMCID: PMCPMC5866245.
656 32. Scala F, Kobak D, Bernabucci M, Bernaerts Y, Cadwell CR, Castro JR, et al. Phenotypic
657 variation of transcriptomic cell types in mouse motor cortex. Nature. 2020. Epub 2020/11/14. doi:
658 10.1038/s41586-020-2907-3. PubMed PMID: 33184512; PubMed Central PMCID:
659 PMCPMC8113357.
660 33. Cardin JA, Carlen M, Meletis K, Knoblich U, Zhang F, Deisseroth K, et al. Driving fast-
661 spiking cells induces gamma rhythm and controls sensory responses. Nature. 2009;459(7247):663-
662 7. Epub 2009/04/28. doi: 10.1038/nature08002. PubMed PMID: 19396156; PubMed Central
663 PMCID: PMCPMC3655711.
664 34. Sohal VS, Zhang F, Yizhar O, Deisseroth K. Parvalbumin neurons and gamma rhythms
665 enhance cortical circuit performance. Nature. 2009;459(7247):698-702. Epub 2009/04/28. doi:
666 10.1038/nature07991. PubMed PMID: 19396159; PubMed Central PMCID: PMCPMC3969859.

35. Bruno RM, Simons DJ. Feedforward mechanisms of excitatory and inhibitory cortical receptive fields. J Neurosci. 2002;22(24):10966-75. Epub 2002/12/18. PubMed PMID: 12486192; PubMed Central PMCID: PMCPMC6758449.

36. Xiang Z, Huguenard JR, Prince DA. Cholinergic switching within neocortical inhibitory networks. Science. 1998;281(5379):985-8. Epub 1998/08/14. doi: 10.1126/science.281.5379.985. PubMed PMID: 9703513.

37. Britten KH, Shadlen MN, Newsome WT, Movshon JA. The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci. 1992;12(12):4745-65. PubMed PMID: 1464765.

38. Wang L, Narayan R, Grana G, Shamir M, Sen K. Cortical discrimination of complex natural stimuli: can single neurons match behavior? J Neurosci. 2007;27(3):582-9. PubMed PMID: 17234590.

39. Theunissen F, Miller JP. Temporal encoding in nervous systems: a rigorous definition. J Comput Neurosci. 1995;2(2):149-62. Epub 1995/06/01. doi: 10.1007/BF00961885. PubMed PMID: 8521284.

40. Narayan R, Ergun A, Sen K. Delayed inhibition in cortical receptive fields and the discrimination of complex stimuli. J Neurophysiol. 2005;94(4):2970-5. PubMed PMID: 15917327.

41. Dong J, Colburn HS, Sen K. Cortical Transformation of Spatial Processing for Solving the Cocktail Party Problem: A Computational Model(1,2,3). eNeuro. 2016;3(1). Epub 2016/02/13. doi: 10.1523/ENEURO.0086-15.2015. PubMed PMID: 26866056; PubMed Central PMCID: PMCPMC4745179.

689 42. Kapfer C, Glickfeld LL, Atallah BV, Scanziani M. Supralinear increase of recurrent
690 inhibition during sparse activity in the somatosensory cortex. Nat Neurosci. 2007;10(6):743-53.
691 Epub 2007/05/23. doi: 10.1038/nn1909. PubMed PMID: 17515899; PubMed Central PMCID:
692 PMCPMC3518866.
693 43. Silberberg G, Markram H. Disynaptic inhibition between neocortical pyramidal cells
694 mediated by Martinotti cells. Neuron. 2007;53(5):735-46. Epub 2007/03/03. doi:
695 10.1016/j.neuron.2007.02.012. PubMed PMID: 17329212.
696 44. Adesnik H, Bruns W, Taniguchi H, Huang ZJ, Scanziani M. A neural circuit for spatial
697 summation in visual cortex. Nature. 2012;490(7419):226-31. Epub 2012/10/13. doi:
698 10.1038/nature11526. PubMed PMID: 23060193; PubMed Central PMCID: PMCPMC3621107.
699 45. Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M. Inhibition of inhibition in visual cortex:
700 the logic of connections between molecularly distinct interneurons. Nat Neurosci.
701 2013;16(8):1068-76. Epub 2013/07/03. doi: 10.1038/nn.3446. PubMed PMID: 23817549;
702 PubMed Central PMCID: PMCPMC3729586.
703 46. Gritton HJ, Nocon JC, James NM, Lowet E, Abdulkerim M, Sen K, et al. Oscillatory
704 activity in alpha/beta frequencies coordinates auditory and prefrontal cortices during extinction
705 learning. bioRxiv. 2020. doi: 10.1101/2020.10.30.362962.
706 47. Zhu Y, Qiao W, Liu K, Zhong H, Yao H. Control of response reliability by parvalbumin-
707 expressing interneurons in visual cortex. Nat Commun. 2015;6:6802. Epub 2015/04/15. doi:
708 10.1038/ncomms7802. PubMed PMID: 25869033.
709 48. Morrill RJ, Hasenstaub AR. Visual Information Present in Infragranular Layers of Mouse
710 Auditory Cortex. J Neurosci. 2018;38(11):2854-62. Epub 2018/02/15. doi:

711 10.1523/JNEUROSCI.3102-17.2018. PubMed PMID: 29440554; PubMed Central PMCID:
712 PMCPMC5852663.

713

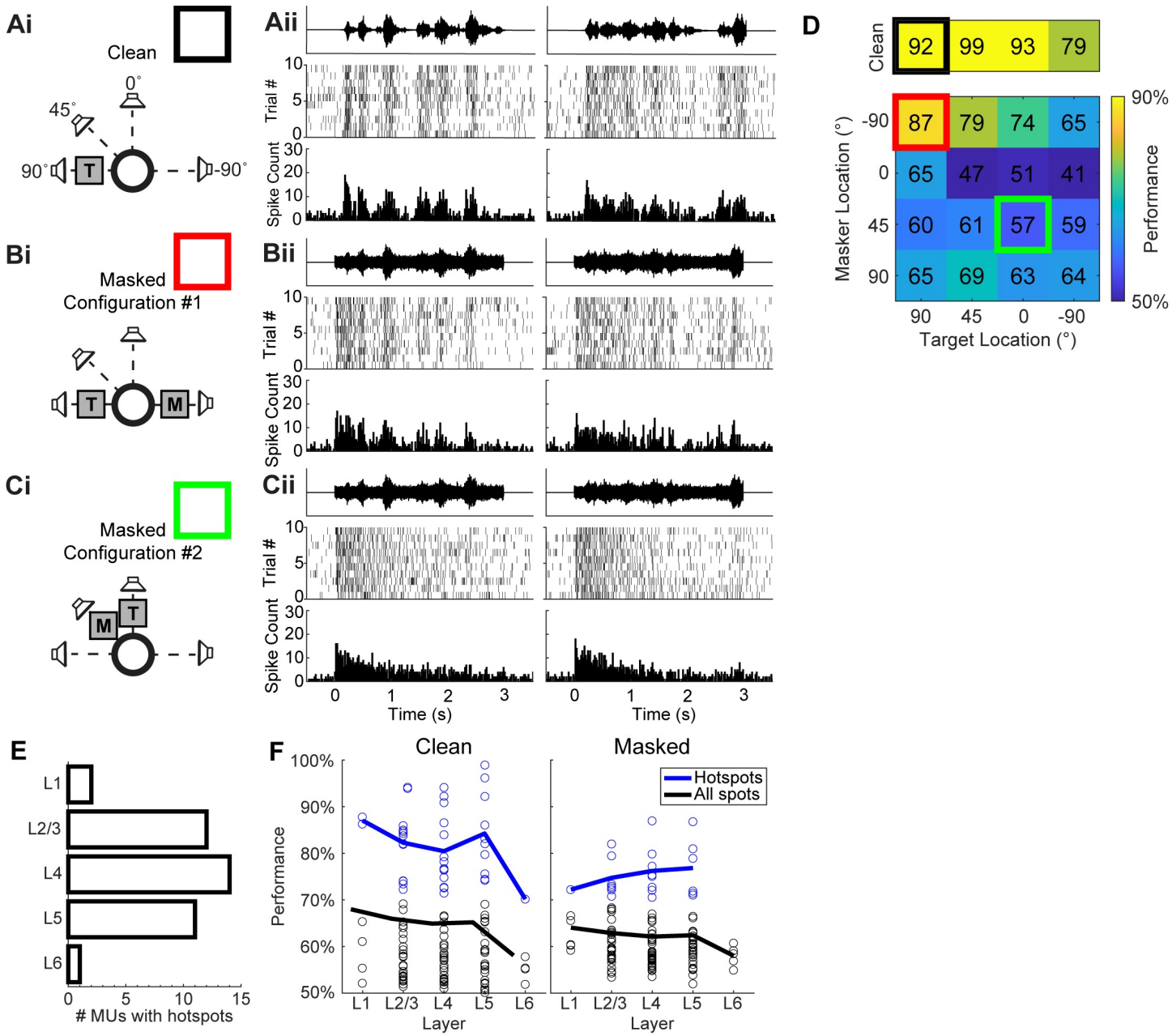


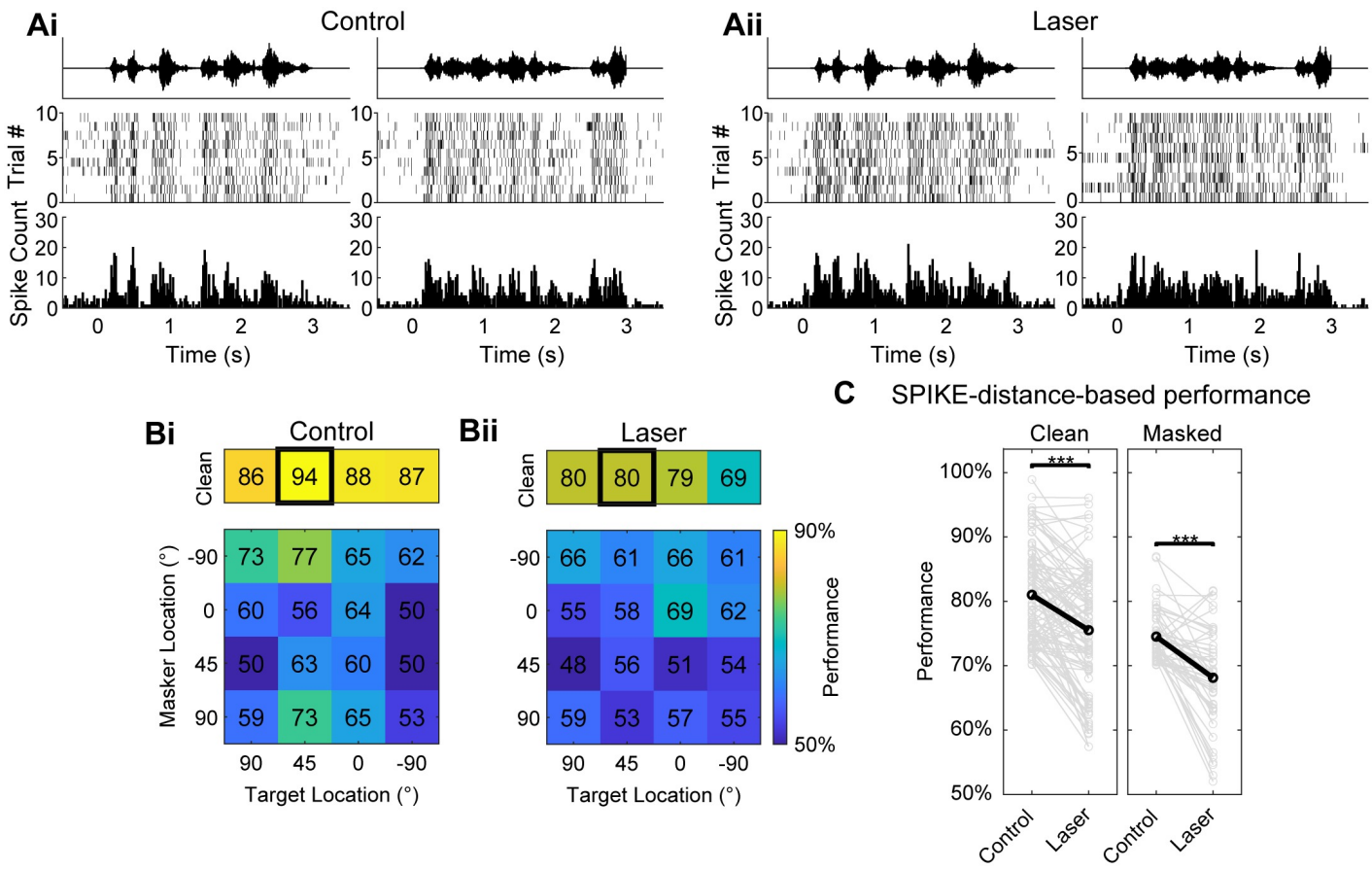
Fig 1. Cortical discrimination in a cocktail party paradigm in mouse A1.

A: Illustration depicting the stimulus configuration for clean trials originating at +90˚ azimuth (**Ai**) and responses to both target stimuli (T) (**Aii**). Auditory stimuli were presented from speakers at 4 locations. Target stimuli consisted of white noise modulated by human speech envelopes extracted from recordings of speech sentences (see Methods, 'Auditory stimuli'). As shown in Aii, responses during clean trials exhibit spike timing and rapid firing rate modulation that follow the amplitude

envelope of both target stimuli. **Bi:** Stimulus configuration for trials where targets (T) played at +90˚ and a competing masking stimulus (M) played at -90˚. Masking stimuli consisted of unmodulated white noise with the same onset and offset times as target stimuli. **Bii**: Responses to masked trials shown in Bi. In this configuration, spike timing and firing rate modulation follow both target stimuli, despite the presence of the competing masker. **Ci:** Stimulus configuration for trials where targets played at 0˚ and maskers played at +45˚. **Cii**: Responses to target-masker configuration shown in Ci. For this configuration, spike timing and firing rate modulation do not follow either target stimulus, resulting in similar responses between target identities. **D:** Neural discriminability performance for all possible target-masker location configurations, referred to as the "spatial grid", for example cell featured in A-C. Outlined spots indicate configurations shown in A-C, matched by the outline color. Performance is calculated using a template-matching approach based on differences in instantaneous firing rate and spike timing similarities between trains (see Methods, 'Neural discriminability using SPIKE-distance'). The top grid shows discriminability for "clean" trials on top, while the bottom grid shows discriminability for "masked" trials. All blocks are color-coded according to the color axis shown to the right of the masked grid. Configurations with relatively high performance (above 70%), e.g., the configuration outlined in red, are referred to as "hotspots". **E.** Number of spatial configuration-sensitive sites per layer across all animals. Layers were identified using current-source density analysis on the local field potential to identify the L4 granular layer (see Methods, 'Current source density analysis'). **F:** Performance during non-optogenetic trials for clean (left) and masked (right) stimuli, with the lines showing mean performance per layer across all spots (black) or hotspots only (blue). ANOVA by layer was separately done for clean and masked trials and included all spots. The ANOVA for clean trials did not reveal a significant effect of layer ($p(4,519) = 0.0505$), while the ANOVA for

743 masked did ($p(4,2092) = 0.0033$). Post-hoc Tukey-Kramer comparisons revealed a significant

744 difference in mean masked performance between L2/3 and L6 ($p = 0.0041$).



745 **Fig 2. Effects of silencing PV neurons on cortical discrimination.**

746 **A:** Responses during clean stimulus trials originating at 45˚ from one example cell during control
747 (**Ai**) and optogenetic (**Aii**) conditions. PV silencing resulted in an increase in the baseline firing
748 rate effectively reducing the modulation depth of the neural response. **B:** Spatial grids from
749 example cell during control (**Bi**) and optogenetic (**Bii**) conditions. Performance is color-coded
750 according to the axis shown to the right of Bii. The outlined configurations correspond to clean
751 trials at 45˚ shown in Ai-ii. In this example, performance decreased during optogenetic silencing
752 for both clean (see all clean spots in Bi versus Bii) and masked trials (note hotspots at target 45˚,
753 masker 90˚ and target 45˚, masker -90˚ in Bi versus the same spots in Bii). **C:** Paired comparisons
754 of SPIKE-distance-based performance from control and PV-inactivated trials at the same spatial
755 grid location. Paired t-tests yielded a significant decrease in performance for both clean (p =

3.129e-10) and masked ($p = 1.3046e-08$) trials during optogenetic silencing, indicating that optogenetic silencing reduced performance by affecting spike timing and rapid firing rate modulation during stimulus presentation.

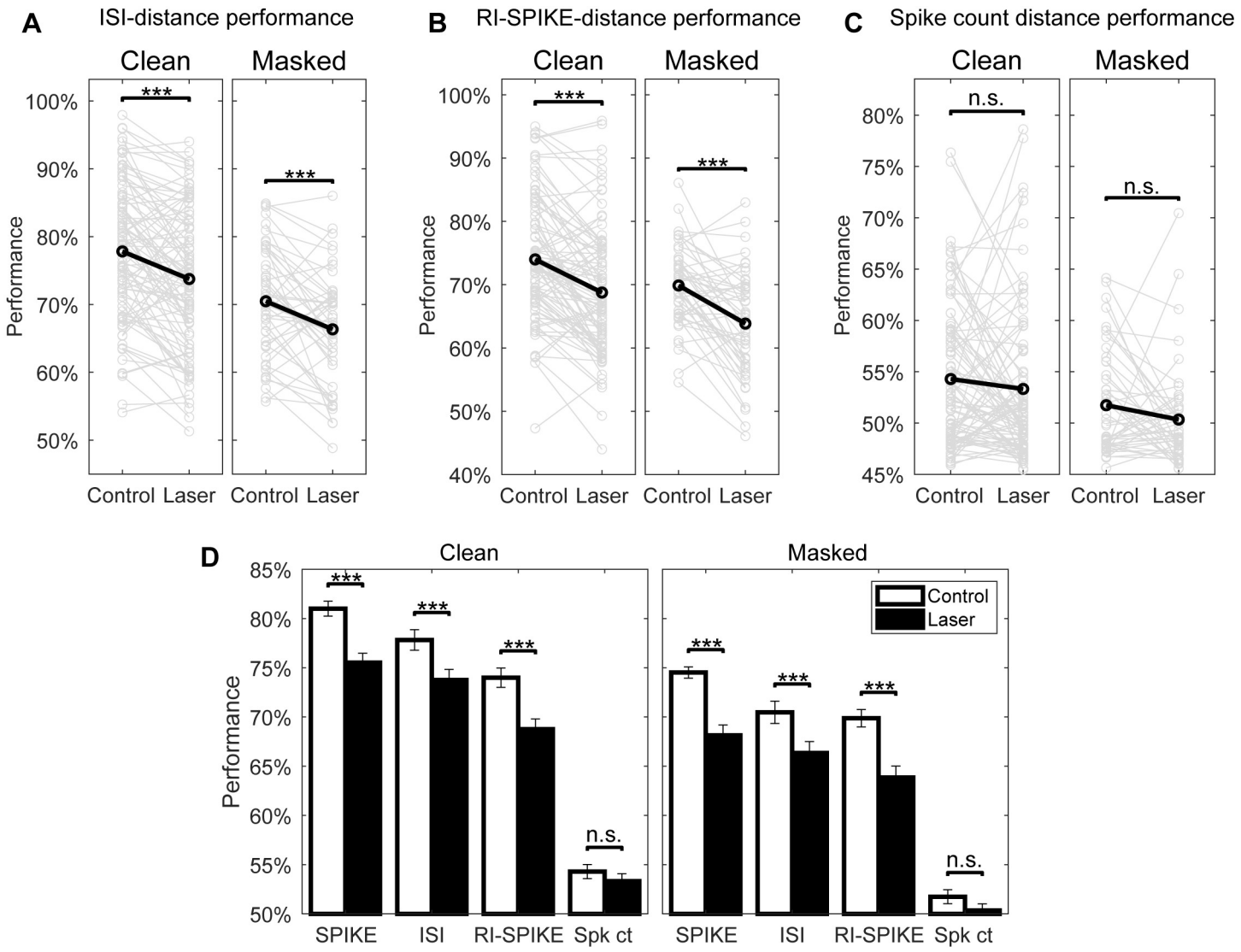
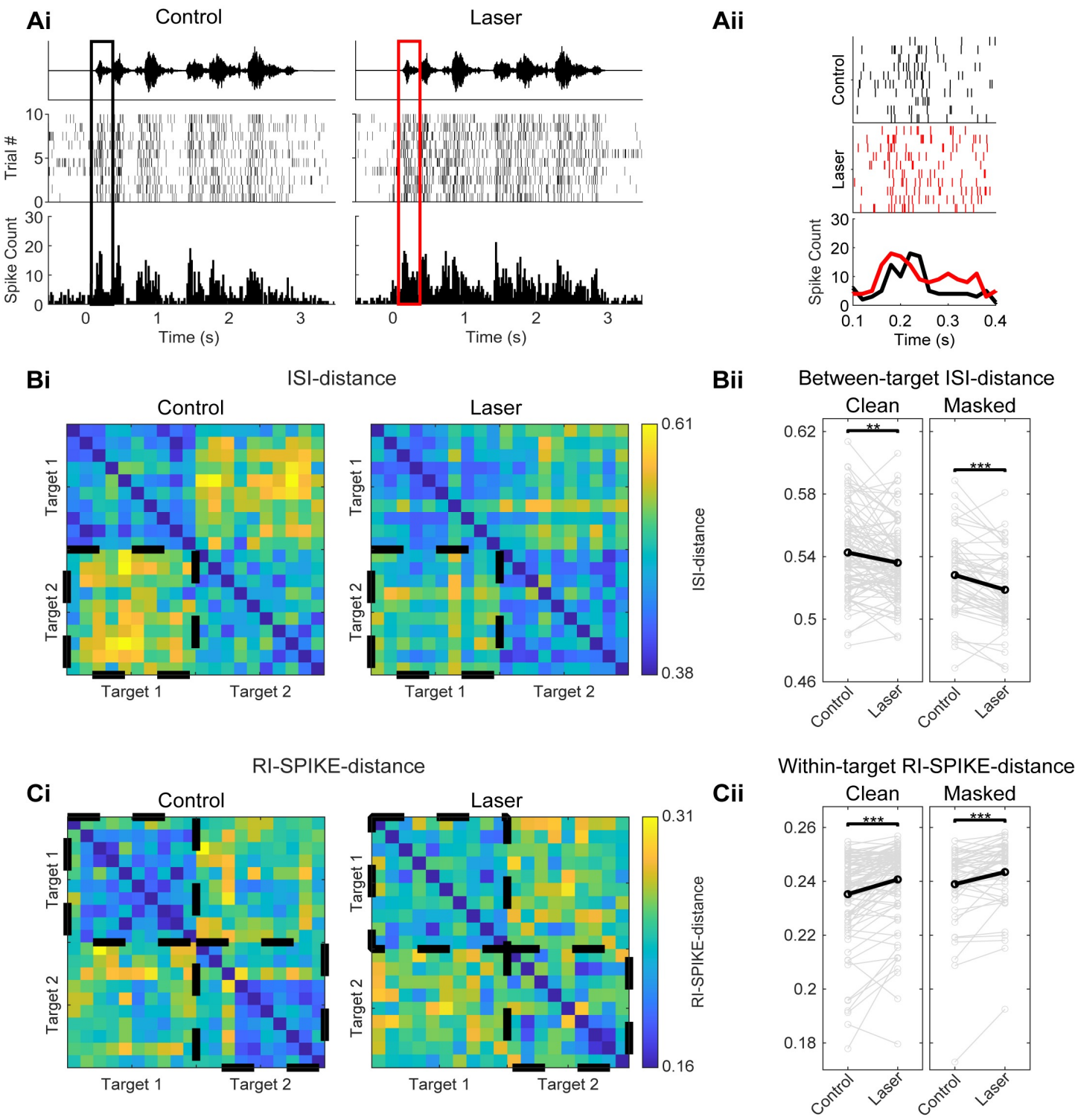


Fig 3. Effects of silencing PV neurons on spike timing and rate-based coding measures.

A: Performance based on ISI-distance, which measures differences between trains in instantaneous firing rate only (see Methods, 'ISI-distance'). Paired t-tests showed a significant decrease in performance for both clean (p = 3.879e-06) and masked (p = 4.325e-04) trials. **B:** Performance based on RI-SPIKE-distance, which measures differences between trains in spike timing only (see Methods, 'RI-SPIKE-distance'). Paired t-tests showed a significant decrease in performance for both clean (p = 1.579e-09) and masked (p = 1.864e-06) trials. **C:** Performance based on differences in total spike count between spike trains. Paired t-tests did not yield a significant change in

performance for both clean ($p = 0.2415$) and masked ($p = 0.1170$) trials. **D:** Summary figure showing contributions from spike-timing measures presented in Figure 2C and 3A-C on the same scale and axis. Changes in spike timing and instantaneous firing rate-based measures (RI-SPIKE and ISI, respectively) reflect the decrease in overall discriminability performance (SPIKE), indicating both components contribute to overall discrimination performance. In contrast, performance based on spike count was near chance level, and differences in performance based on total spike count was not significant, indicating that total spike count is insufficient to explain the decrease in overall discrimination performance.



775 **Fig 4. Effects of silencing PV neurons on distinct components of discrimination: between-**

776 **target difference and within-target reproducibility.**

A: Response to target 1 at 45˚ during both control (left, 'Control') and optogenetic (right, 'Laser') trials from the same cell as in Figure 2Ai-ii, with inset (**Aii**) showing the response from 0.1 to 0.4 seconds after stimulus onset for both non-optogenetic (black) and optogenetic (red) trials. Inset shows an increase in baseline spiking during optogenetic trials relative to the control response. **Bi:** Example ISI-distance matrix of all possible spike train pairs from one target-masker configuration, with the color scale for both conditions shown on the right. ISI-distance increases with differences in instantaneous firing rate, and a value of 0 is obtained for identical spike trains (e.g. the blue main diagonal in all distance matrices represents a spike train's distance with itself). Black dashed outline indicates all distances between responses to different targets. In this example, the mean ISI-distance between different target stimuli decreased during optogenetic silencing. **Bii:** Mean between-target ISI-distance comparisons between conditions for all control hotspots. Paired t-tests yielded a significant decrease in distance during optogenetic silencing for both clean ($p = 0.0058$) and masked ($p = 1.098e-04$) trials; thus, optogenetic silencing decreased differences in rapid firing rate modulations between target stimuli and reduced neural discriminability due to decreased representation of distinct stimuli. **Ci:** Example RI-SPIKE-distance matrix of all possible spike train pairs from one target-masker configuration, with the color scale for both conditions shown on the right. As with ISI-distance, RI-SPIKE-distance is 0 for identical spike trains and increases with dissimilarities in spike timing. Black dashed outlines indicate all distances within responses to the same target stimulus. In this example, the mean RI-SPIKE-distance within the same target increased during optogenetic silencing. **Cii:** Mean within-target RI-SPIKE-distance comparisons between conditions for all control hotspots. Paired t-tests yielded a significant increase in distance during optogenetic silencing for both clean ($p = 1.101e-09$) and masked ($p = 5.602e-04$) trials;

799 thus, optogenetic silencing increased the variability in spike timing within responses to the same

800 target stimulus and reduced neural discriminability by decreasing reproducibility across trials.

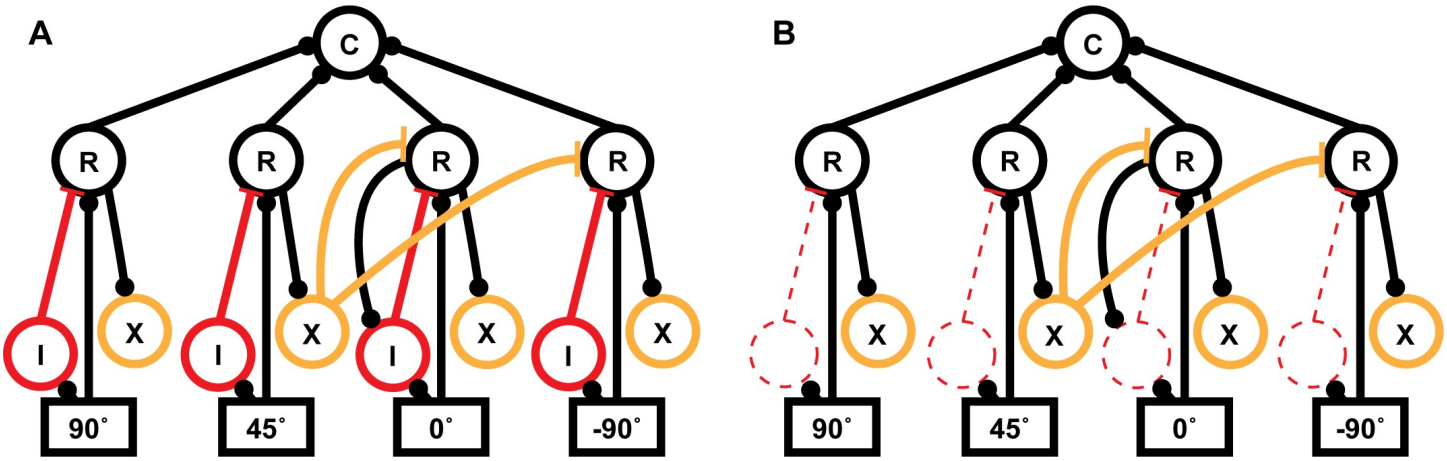


Fig 5. Cortical circuits for complex scene analysis.

Schematic model representing cortical circuit under the control condition **A**) and in the laser condition (**B**). C cells represent cortical units, R cells represent excitatory relay units, I cells represent within-channel inhibitory units, and X cells represent cross-channel inhibitory units.

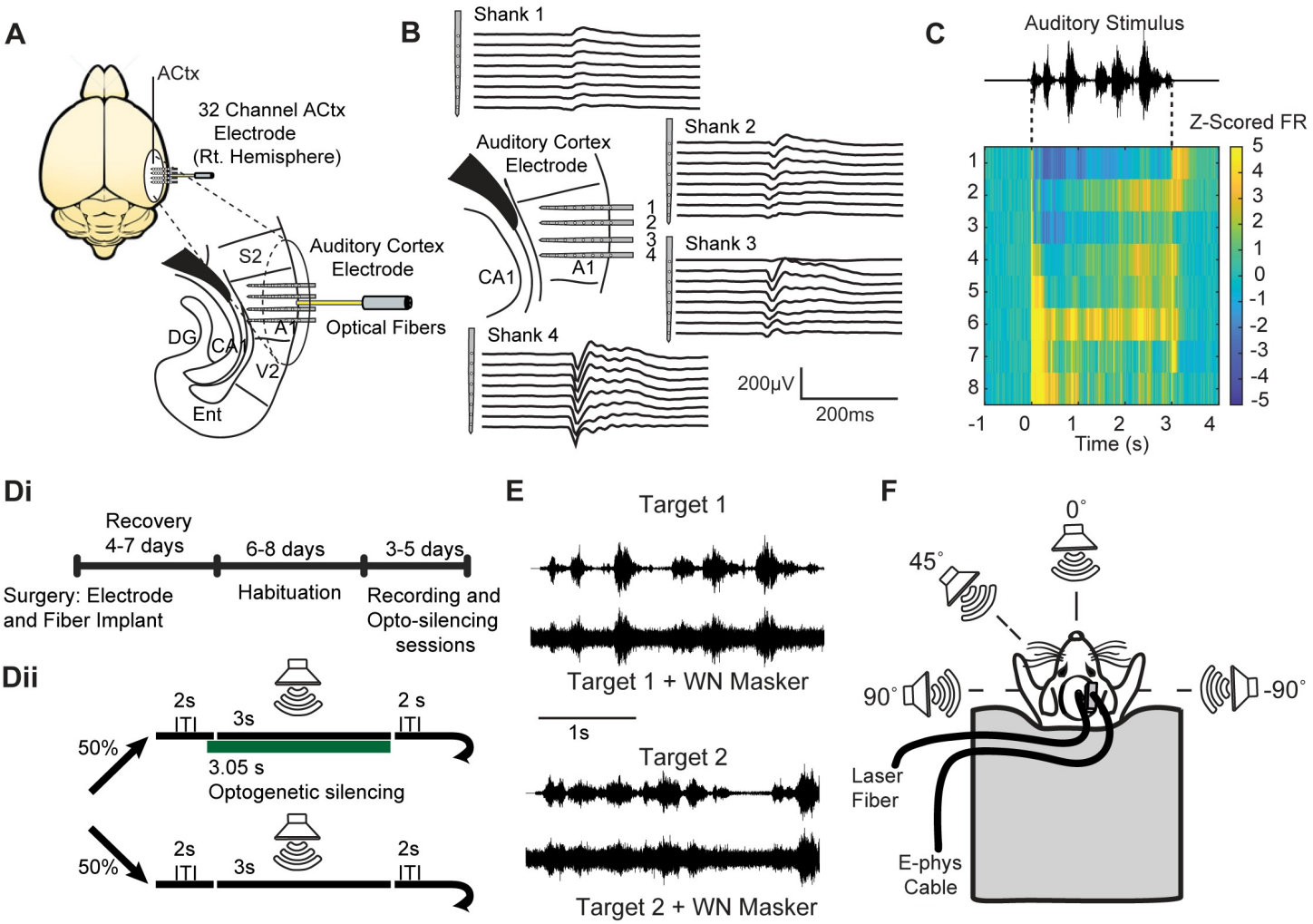


Fig 6. Methodology and experimental design.

A: Illustration depicting recording electrode location and optical fiber placement. Subjects were implanted with a 4-shank, 32-channel electrode and optogenetic fiber in right hemisphere of A1. Each shank contained 8 sites per shank with 100 μm spacing between electrode contacts. **B:** Representative local field potential (LFP) activity from one mouse. LFPs were recorded from all sites and were used for current source density analysis for layer determination (see Methods, 'Current source density analysis'). **C:** Average z-scored spiking activity from all subjects and all shanks normalized by depth (n=9 mice, 36 shanks). **D:** Illustration of experimental timeline and trial configuration. (**Di**): Mice underwent surgery for the implantation, followed by a period of

recovery, before habituation to being head-fixed. The duration of habituation gradually increased across days until habituation sessions matched recording sessions in terms of duration. (**Dii**): Trials consisted of 10 repetitions of all possible target/target-masker/and location combinations randomized for both non-optogenetic trials and optogenetic trials. A total of 800 trials were performed in each recording session. Trials consisted of 3 seconds of stimulus playback followed by 2 seconds of silence. For optogenetic silencing trials, laser onset began 50ms before auditory stimulus onset and ended with stimulus offset. **E:** Diagrams of the 4 possible auditory combinations used to determine neural discrimination performance. Auditory stimuli consisted of one of two target stimuli alone (clean trials), or with a competing masking stimulus (masked trials) originating from the same speaker or any of the other speaker locations. Target stimuli consisted of white noise modulated by human speech envelopes while masker stimuli consisted of unmodulated white noise. **F:** Illustration depicting the recording set-up and speaker configuration for all subjects used in this study. Auditory stimuli were presented at 4 locations, 18 cm from the mouse's head using 4 electrostatic speakers.