Title: PV neurons and cortical temporal coding in complex scene analysis

Abbreviated title: PV neurons and cortical coding

Authors: 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 02215,

[2]Center for Systems Neuroscience, Boston University, Boston, Massachusetts, United States of America, 02215

[3]Hearing Research Center, Boston University, Boston, Massachusetts, United States of America, 02215

[4]Department of Biomedical Engineering, Boston University, Boston, Massachusetts, United States of America, 02215

[5]Department of Comparative Biosciences, University of Illinois, Urbana, Illinois, United States of America, 61820

[6]Department of Bioengineering, University of Illinois, Urbana, Illinois, United States of America, 61820

*Corresponding author

Email: kamalsen@bu.edu (K.S.)

Conflicts of interest: The authors declare no conflicts of interest.

Acknowledgements: This research project was made possible by the National Science Foundation (#IIS-1835270) and the National Institute of Health (#1R34NS111742-01). We would also like to thank Alberto Cruz-Martín, Michael Economo, and Conor Houghton for comments and suggestions on the manuscript.

# 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 and compete 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 suppressing PV neurons degrades cortical discrimination of dynamic sounds in a cocktail party-like setting via changes in rapid temporal modulations in rate and spike timing. These findings suggest a critical role of PV neurons in cortical temporal coding underlying complex scene analysis.


# Significance Statement

One of the most impressive examples of sensory perception by the brain is its ability to analyze complex scenes, e.g., following what a friend is saying at a party amongst other speakers. Although some humans can solve this problem with relative ease, it remains very difficult for

humans with hearing impairments, ADHD, and autism; and assistive devices, e.g., hearing aids. The brain mechanisms underlying complex scene analysis remain poorly understood. We investigated this by recording neural activity in auditory cortex in a complex auditory scene. When we suppressed the activity of PV neurons, cortical performance decreased via changes in response timing. Our findings suggest that PV neurons improve the brain's ability to analyze complex scenes by enhancing cortical response timing.

# 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 related to 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 (deCharms and Zador, 2000). Previous studies have demonstrated that both rate and spike timing-based coding is present in the cortex (Lu et al., 2001; Zuo et al., 2015; Yao and Sanes, 2021). However, a mechanistic understanding of how cortical circuits implement 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 (Pfeffer et al., 2013; Tremblay et al., 2016). This concerted approach has allowed systems neuroscience to identify key contributions of specific cell types to cortical circuit including identifying a breadth of diversity imparted by cortical inhibitory neuron types (e.g., parvalbumin-expressing (PV),

72 somatostatin-expressing (SOM) and vasoactive intestinal peptide-expressing (VIP) neurons)
73 (Tremblay et al., 2016). However, the specific roles of these diverse cell types in contributing to
74 the cortical code remain unexplored.

75 A potentially powerful strategy for unraveling cell-type specific contributions to cortical
76 coding is to probe cortical circuits with experimental paradigms that fully engage cortical circuits
77 in solving problems that require a significant cortical contribution. An important example of such
78 a problem is complex scene analysis, e.g., recognizing objects in a scene cluttered with multiple
79 objects at different spatial locations. The brain displays an astonishing ability to navigate such
80 complex scenes in everyday settings, an impressive feat yet to be matched by state-of-the-art
81 machines. How cortical circuits and cell-types contribute to this powerful computational ability
82 remains unclear.

83 PV neurons are a prominent group of inhibitory neurons in cortex. Several studies have
84 investigated the role played by PV neurons in the generation of oscillations (Antonoudiou et al.,
85 2020) and spike synchronization (Jang et al., 2020). However, the influence of PV neurons
86 influence on cortical coding at the single neuron level is unknown. The biophysical properties of
87 PV neurons are well suited for rapid temporal processing (Tremblay et al., 2016), motivating
88 several open questions: do PV neurons play a critical role in influencing timing-based codes for
89 cortical coding of dynamic stimuli in complex scenes? Are such temporal representations robust
90 to competing stimuli at other locations in space? Here, we address these questions in auditory
91 cortex using a combination of high-density electrophysiology, optogenetic tools for suppression
92 of PV neurons, and a family of novel spike distance metrics (Kreuz et al., 2013; Satuvuori and
93 Kreuz, 2018) to dissect specific contributions of PV neurons to the cortical code.

The auditory cortex is well suited to investigate these issues. It is thought to play a key role in solving the "cocktail party problem" (Narayan et al., 2007; Mesgarani and Chang, 2012), one of the most impressive examples of complex scene analysis. Here, we integrate a cocktail party-like paradigm (Maddox et al., 2012) with optogenetic suppression 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 suppressing PV neurons degrades cortical discrimination performance in A1 via spike timing-based coding in a cocktail party-like environment, suggesting that PV neurons play a critical role in cortical temporal coding underlying complex scene analysis.

# Materials and methods

## Subjects

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

## Surgery

Mice were surgically implanted with a head-plate as described previously (James et al., 2019; Gritton et al., 2020). Briefly, under isoflurane anesthesia, stereotaxic surgery was performed install a headplate, electrode, and optical fiber. The custom head-plate was mounted anterior to bregma allowing access to A1 caudally. The headplate was anchored to the skull with 3 stainless steel screws and dental cement. A fourth screw was connected to a metal pin and placed in the skull above the contralateral cerebellum to serve as the reference. A craniotomy was made above the right auditory cortex (AP -2.3 to -3.6, ML + 4.0 to +4.5, DV). Using a stereotaxic arm, a 32-contact linear probe (Neuronexus, Ann Arbor, MI; model: a 4x8-5mm-100-400-177-CM32) with 100µm spacing between electrode contacts and 400µm spacing between shanks, was positioned into A1, perpendicular to the cortical surface. Because of the curvature of the A1 cortical surface, not all four shanks could be placed at precisely the same depth during each experiment. Probes were advanced until all electrode contacts were within the cortical tissue and shanks were positioned along the rostro-caudal axis of A1 (Fig 1A-C). An optical fiber, 200µm in diameter, was placed medially to the 4 shanks and positioned between the two innermost shanks terminating at the cortical surface (Fig 1A). After implantation, mice were allowed to recover for 4-7 days before undergoing habituation to being head-fixed as described below.

## Habituation

Following surgery and complete recovery, mice were first handled for several days before being head-fixed to the recording apparatus. Mice were gradually exposed to longer restraint periods at the same time of day as subsequent recording sessions. Each animal received at least 6 habituation sessions prior to the first recording day. Under head-fixed conditions, mice were

135 loosely wrapped with a piece of lab tissue (Kimwipes: Kimberly-Clark, Irving, TX) to encourage
136 reduced movement. At the conclusion of habituation, between three and five days of recording
137 sessions were done per animal and had the same duration as habituation sessions.

138

## 139 Recording sessions and data acquisition

140 All recordings were made with a Tucker Davis Technologies RZ2 recording system in an
141 electrically-shielded sound attenuation chamber. Broadband neural signals at 24414.0625 Hz were
142 recorded for each of the 32 channels. Local field potentials (LFPs) were band-pass filtered between
143 1 and 300 Hz, notch-filtered at 60 Hz, and digitized at 3051.8 Hz.

144 Recording sessions consisted of both non-optogenetic and optogenetic trials in random
145 order. Inter-trial interval was 5 seconds, with 3s of stimulus playback followed by 2s of silence.
146 Mice were exposed to target-alone ("clean") trials and target-masker ("masked") combinations. 10
147 trials were given per target identity for all possible combinations of target location, masker location
148 (including "clean" trials), and optogenetic suppression of PV neurons. Thus, animals received a
149 total of 800 trials per ~60 minute recording session.

150

### 151 Auditory stimuli

152 All auditory stimuli were generated in Matlab and auditory stimuli consisted of either
153 target, masker or combination stimuli played from the four ES-1 electrostatic speakers. Target
154 stimuli consisted of white noise modulated in time by human speech envelopes extracted from
155 recordings of Harvard sentences. Masker stimuli consisted of 10 unique tokens of unmodulated

156 white noise. Before speaker calibration, all stimuli were generated with the same RMS value, and
157 sampling frequency was 195312 Hz to capture the frequency range of hearing in mice. Stimuli
158 were loaded onto a custom RPvdsEx circuit on an RZ6 Multi I/O processor, which was connected
159 to two PM2R multiplexers that controlled the location of target and masker stimuli during trials.

160 During recordings, the stimuli were presented 18 cm from the mouse's head using four ES-
161 1 electrostatic speakers, driven by two ED-1 speaker drivers. The four speakers were arranged
162 around the mouse at four locations on the azimuthal plane: directly in front (0°), two contralateral
163 (45° and 90°) and 1 ipsilateral (-90°) to the right auditory cortex recording area. Before recording
164 sessions, stimulus intensity was calibrated using a conditioning amplifier and microphone (Brüel
165 and Kjaer, amplifier: model 2690, and microphone model: 4939-A-011). For 7 of the 9 mice, all
166 stimuli were at a measured 75dB intensity at the mouse's head. For the remaining 2 mice, stimulus
167 intensity was set to 70dB. Stimulus playback lasted 3s with a 1ms cosine ramp at onset and offset.

168

## 169 Optogenetic stimulation

170 Laser light for optogenetic stimulation of nucleus basalis was delivered through a
171 multimode optically-shielded 200µm fiber (Thorlabs, Newton, NJ; model: BFH48-200), coupled
172 to a 532nm DPSS laser (Shanghai Laser Ltd., Shanghai, China; model: BL532T3-200FC). Laser
173 power was calibrated to 10mW at the fiber tip using a light meter calibrated for 532nm wavelength
174 (PM100D, Thorlabs, Newton, NJ). The intensity was determined based on optogenetic cortical PV
175 suppression studies using Archaerhodopsin from the literature (Aizenberg et al., 2015; Zhu et al.,
176 2015). During optogenetic trials, the laser was turned on 50ms before stimulus onset and co-
177 terminated with auditory stimuli (Fig 1D). Square light pulses lasting 3.05s were delivered via

178 TTL trigger from the RZ2 recording system to the laser diode controller (ADR-1805). Optogenetic
179 trials were randomized throughout the recording session such that animals received all
180 stimulus/masker pairs from each location with and without laser. Recordings were done in
181 successive blocks of increasing optogenetic suppression strengths of 2mW, 5mW, and 10mW,
182 with each block lasting ~60 minutes and having their own set of control trials. These laser strengths
183 are consistent with past studies to ensure that optogenetic suppression did not result in epileptiform
184 activity in cortex (Aizenberg et al., 2015; Phillips and Hasenstaub, 2016).

185

186 # Statistical analysis

187 ## Spike extraction and clustering

188 Kilosort2 (https://github.com/MouseLand/Kilosort) was used to detect multi-units within
189 the recordings (Pachitariu et al., 2016). Before spike detection and sorting, the broadband signal
190 was band-passed between 300 and 5000 Hz using a 3$^{rd}$-order Butterworth filter. Kilosort results
191 were then loaded onto Phy2 (https://github.com/cortex-lab/phy), a graphical user interface for
192 examining spike cluster information (Rossant et al., 2016). Clusters with artifact-like spike
193 waveforms were deemed as noise, while the remaining clusters were deemed as either single or
194 multi-units. The spikes toolbox (https://github.com/cortex-lab/spikes) was then used to import the
195 cluster information from Phy to Matlab and extract spike waveforms from the high-passed signal
196 (Pachitariu et al., 2016). Clusters were assigned to recording channels based on which site yielded
197 the largest average spike amplitude. To remove any remaining artifacts from laser onset and offset,
198 all spikes with waveforms above an absolute threshold of 1500 μV or a positive value above 750

199 μV were discarded. Finally, clusters were deemed single-unit if less than 5% of their inter-spike
200 intervals were below 2ms, while the remaining clusters were deemed multi-unit (Fig 1E).

201 Clusters were deemed auditory-responsive if the mean z-score during non-optogenetic
202 trials was larger than an absolute value of +1 between 10 and 30ms after stimulus onset or during
203 the entire 3s of stimulus playback, based on the average response up to 1 second before stimulus
204 onset. To further refine our dataset, we excluded units that showed a lot of artifacts or a weak
205 auditory response. For each unit, we estimated the mean firing rate at each configuration by
206 binning the 10 trials per target into a PSTH with a bin size of 20ms. We then compared the
207 maximum evoked firing rate to spontaneous activity at each spatial grid configuration using a two-
208 sample t-test. Configurations with responses that yielded p-values above 0.01 were deemed as
209 "weak" responses and thus excluded from further analysis.

210

## 211 Neural discriminability performance using SPIKE-distance

212 Neural discrimination performance was calculated using a template-matching approach
213 similar to our previous study (Maddox et al., 2012). Briefly, spike trains were classified to one of
214 the two target stimuli based on whose template, one from each stimulus, yielded a smaller spike
215 distance. For each target-masker configuration, 100 iterations of template matching were done. In
216 each iteration, one of the 10 spike trains for each target was chosen as a template, and all remaining
217 trials were matched to each template to determine target identity. All possible pairs of templates
218 were used across the 100 iterations to calculate an average value of neural discriminability. SPIKE-
219 distance (Satuvuori and Kreuz, 2018) calculates the dissimilarity between two spike trains based

on differences in spike timing and instantaneous firing rate without additional parameters. For one spike train in a 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 $\Delta t_P$ represents the distance between the preceding spike from train 1 ($t_P^{(1)}$) and the nearest spike from train 2, $\Delta t_F$ 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:

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

236 cSPIKE (http://wwwold.fi.isc.cnr.it/users/thomas.kreuz/Source-Code/cSPIKE.html) was used to
237 calculate all spike train distances between all possible spike train pairs for all spatial grid
238 configurations (Satuvuori and Kreuz, 2018).

239 Spatial grid "hotspots" of high neural discriminability were determined using three criteria:
240 1) mean performance must be above 70% during control trials, 2) performance must be significant
241 via two-sample t-test ($p < 0.05$) relative to a null distribution generated by template-matching spike
242 trains from each target with themselves, and 3) the effect size between the two distributions must
243 be greater than 1:

244
$$d = \frac{\bar{x}_1 - \bar{x}_0}{\sqrt{\frac{(n_1 - 1)s_1^2 - (n_0 - 1)s_0^2}{n_1 + n_0 - 2}}}$$

245 where values with subscript 0 represent the mean, standard deviation, and number of template-
246 matching iterations for the null performance distribution. Additionally, configurations where at
247 least 3 trials for one target showed zero spiking were excluded from analysis, as the lower number
248 of trials would result in a coarser estimation of performance. To determine how firing rate
249 modulation, spike timing, and average firing rate contribute to discriminability, we used different
250 distance measures as inputs to the template matching approach. For all hotspots, performances
251 using inter-spike interval (ISI)-distance, rate-independent (RI)-SPIKE-distance, and spike count
252 distance, the absolute difference in spike count between trains, were also calculated and compared
253 to SPIKE-distance-based values. We utilized paired t-tests between control and optogenetic values
254 to determine statistical significance ($p < 0.05$), with tests done separately for clean and masked
255 trials.

256

## Comparisons between different optogenetic strengths

To determine the effect of optogenetic suppression strength on neural responses, we calculated the change in firing rate and performance between conditions within each experimental block. For each block, we first found the spatial grid hotspots using our performance and effect size criteria. Using a one-way ANOVA with laser power as the between-subjects factor, we compared the average difference in firing rate during the 50ms after laser onset, the change in mean evoked firing rate, and the change in discriminability performance between control and optogenetic conditions within each experimental block. For all three measures, data from clean and masked trials were pooled before determining statistical significance. Post-hoc multiple comparisons were then done for ANOVAs that yielded laser power as a significant factor.

## Sparseness

To determine how the sensitivity of neural responses changes with optogenetic suppression, 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 10ms were used, and the driven responses of both target stimuli were concatenated in the time axis before calculations.

## ISI-distance

To determine how optogenetic suppression 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 suppression affects spike timing, 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.

# 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 (Maddox et al., 2012) 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°) to the right auditory cortex recording area. Target stimuli consisted of white noise modulated by human speech envelopes extracted from a speech corpus (Rothauser, 1969). 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 target-masker combinations (Figs 2Ai-Ci).

## 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 2A); and with the masker, for all possible combinations of target-masker locations ("Masked" condition, Figs 2B-C). Values near 100% and 50% represent high discrimination performance and chance performance, respectively. We refer to the matrix of performance values from all speaker configurations as "spatial grids",

317 and positions of high performance (≥ 70%), which were also statistically significant ($p < 0.05$) and
318 had a relatively large effect size ($d \geq 1$), as "hotspots". These hotspots represent locations of
319 enhanced discriminability between the two targets using a spike distance-based classifier to
320 determine how well target identity can be predicted given the spike train from that site based on
321 dissimilarities in spike timing and instantaneous rate (Maddox et al., 2012) (Fig 2D: see Methods,
322 'Neural discriminability using SPIKE-distance').

323 Fig 2A illustrates spike trains from an example site that shows high discriminability under
324 both target-only conditions (Figs 2A, 2D, and 2E: black); and for a specific spatial configuration
325 in the presence of a competing noise masker (Figs 2B, 2D, and 2E: red). In the masked condition,
326 discriminability depends strongly on the spatial configuration of the target and masker, indicating
327 that the response at this site is spatial configuration-sensitive (Figs 2C-E, red versus green).

328 We next quantified the proportion of neurons that displayed hotspots. We used automated
329 detection and sorting algorithms to identify 135 units from n = 9 animals used in this study
330 (Pachitariu et al., 2016; Rossant et al., 2016). Consistent with previous studies from auditory
331 cortex, we found 69 of these 135 units, or ~51% of units, were significantly sound-modulated in
332 response to the auditory stimulus. To further quantify these units, we focused our analysis on those
333 cells with statistically significant responses to auditory stimuli (see Methods: 'Spike extraction
334 and clustering'). Further analysis revealed that 33 of these 69 units showed hotspots at one or more
335 spatial configurations tested or ~48% of all auditory responsive units. Using these criteria, we
336 found a total of 80 spatial hotspots in the clean condition and 38 spatial hotspots in the masked
337 condition from 33 (out of 135) recorded units.

338

339 **Suppression of PV neurons reduces discrimination performance at**
340 **hotspots.**

341 To investigate the role of PV interneurons in auditory discrimination performance, we
342 compared responses at hotspots, with and without optogenetic suppression of PV neurons in A1
343 (Fig 1D). Optogenetic suppression occurred on 50% of trials, randomly interleaved, throughout
344 the recording sessions. Suppression was achieved using a 2mW, 5mW or 10mW, light output from
345 a 532nm laser that began 50ms prior to the auditory stimulus and consisted of continuous
346 illuminations that co-terminated with sound offset. All three intensities enhanced the firing rate
347 and reduced overall performance in a level dependent manner although the effects were most
348 robust at the 10mW intensity (Extended Data Figure 3-1). Thus, the remaining analysis will focus
349 on results from the 10mW intensity.

350 As seen in the individual examples in Fig 3A, and described in detail below, we found that
351 PV suppression altered the temporal dynamics of modulations in firing rate during stimuli. To
352 better understand how optogenetic suppression impacted performance at hotspots, we compared
353 performance across sites in the presence or absence of optogenetic suppression. We found that
354 hotspot performance consistently decreased in both clean and masked conditions during
355 suppression (Fig 3B-C).

356 To further quantify the effect of optogenetic suppression on spiking during stimulus
357 playback, we compared the driven firing rate (FR) and temporal sparseness of activity of units in
358 the presence or absence of PV suppression. We found that the driven FR increased significantly in
359 both the clean and masked condition (Fig 3D). Temporal sparseness is a measure of the distribution
360 of a unit's response over time, and sparse coding is thought to increase the efficiency of encoding

361 by minimizing the number of spikes necessary to convey information about the stimulus (Vinje
362 and Gallant, 2000). We found there was a significant reduction in temporal sparseness for both the
363 cleaned and masked conditions (Fig 3E). These results indicate that PV suppression strongly
364 affects both firing rate and temporal sparseness during the stimulus.

365 Based on our findings that PV suppression reduced neural discriminability across the
366 spatial grid, we next assessed the relationship between FR and performance by correlating activity
367 between the two measures; we found no significant relationship between driven FR and
368 performance in either the optogenetic ($n = 118$ configurations; $r = 0.0140$, $p = 0.8811$) or the non-
369 optogenetic ($r = 0.0320$, $p = 0.731$) conditions. Therefore, although optogenetic suppression
370 elevates firing rate and reduces performance, changes in firing rate independently does not account
371 for the degradation in discrimination performance with PV suppression.

372 When we considered how sparseness was correlated with performance, we found no
373 significant relationship between sparseness and performance under control ($n = 118$
374 configurations; $r = 0.166$, $p = 0.0720$) and a weakly positive relationship under the laser ($r = 0.184$,
375 $p = 0.0460$) condition, indicating that changes in sparseness could not fully account for the
376 degradation in discrimination performance with PV suppression.

377

## 378 Reduced discrimination performance driven by changes in temporal
## 379 dynamics of rapid firing rate modulation and spike timing.

380 To determine the extent to which changes in the temporal dynamics of rapid firing rate
381 modulation, spike timing, and average firing rate changes that occur during suppression might

382 affect performance, we calculated different performance metrics across all hotspots. Specifically,
383 we used inter-spike interval (ISI)-distance, rate-independent (RI)-SPIKE-distance, or spike count,
384 as the basis for discriminability between spike trains. ISI-distance calculates the distance between
385 two spike trains based on the average dissimilarity in instantaneous firing rate, while RI-SPIKE-
386 distance measures spike timing dissimilarity between two trains without being sensitive to firing
387 rate differences (Satuvuori and Kreuz, 2018). Spike count distance is the absolute difference in the
388 number of spikes between trains, effectively measuring differences in total firing rate. We found
389 that both ISI-distance-based (Fig 4A) and RI-SPIKE-distance-based (Fig 4B) performances
390 showed significant decreases with optogenetic suppression. Spike count distance-based
391 performance, however, did not show significant changes in performance during the optogenetic
392 condition (Fig 4C). These findings suggest that discrimination based on both rapid rate modulation
393 and spike timing-based coding are profoundly influenced by optogenetic suppression of PV
394 neurons, but not by average firing rate or total spike count.

395 Overall, these findings demonstrate that the observed decrease in discrimination
396 performance upon PV suppression is due to effects on the temporal dynamics of rapid rate
397 modulation and spike timing, and not due to changes on global rate measures including average
398 rate or total spike count (Fig 3C, 4D). Taken together, these results suggest that PV neurons
399 improve cortical discrimination of dynamic stimuli in complex scenes by enhancing rapid temporal
400 modulations in firing rate and spike timing reproducibility and help maintain this temporal code
401 in the presence of competing stimuli at other spatial locations, providing a robust foundation for
402 cortical complex scene analysis.

403

# Discussion

## Diversity of cortical cell types and the cortical code.

One of the most striking features of the cerebral cortex is the tremendous diversity of its cell types (Scala et al., 2020). Understanding the computational role of such diversity in cortical coding is central to systems neuroscience. Addressing this central question requires understanding cell type-specific contributions to the cortical code at both the single neuron and the population level. A small number of previous studies have demonstrated a role of specific cell types in cortical population coding; specifically, the generation of oscillations (Cardin et al., 2009; Sohal et al., 2009), and synchrony across cortical layers and areas (Xiang et al., 1998; Bruno and Simons, 2002). However, cell type-specific contributions to the cortical code at the single and multi-unit level, a fundamental aspect of the cortical code, remain poorly understood. In this study, we addressed this fundamental gap by investigating the role of PV neurons in cortical coding of a complex scene, i.e., a cocktail party-like setting, in mouse A1.

## 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 (Britten et al., 1992; Wang et al., 2007). 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 suppression 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 suppression of PV neurons generally degraded discrimination performance in both clean and masked conditions by affecting rapid firing rate modulations and spike timing. Overall, our findings suggest that PV neurons enhance both rate modulation and spike timing-based cortical coding.

Previous studies have demonstrated that auditory cortical neurons can employ both rate and spike timing-based codes (Lu et al., 2001; Yao and Sanes, 2021); and provided insight into the roles of inhibitory neurons in shaping frequency tuning (Moore and Wehr, 2013; Li et al., 2014; Li et al., 2015; Kato et al., 2017), frequency discrimination (Aizenberg et al., 2015; Briguglio et al., 2018), adaptation (Natan et al., 2015), sparseness (Liang et al., 2019), and gap encoding (Keller et al., 2018). An influential review on neural coding also defined a precise notion of a "temporal code" as one that contains information in spike timing beyond rate modulations (Theunissen and Miller, 1995). 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 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 (Tremblay et al., 2016).

## 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(Maddox et al., 2012). 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.

## 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 (Narayan et al., 2005); whereas inhibition across different channels can lead to the formation of hotspots (Dong et al., 2016). Our results suggest that PV neurons mediate within-channel inhibition, corresponding to I neurons in the schematic model (Fig 5A). This is consistent with our finding that although suppressing PV neurons reduced discrimination performance, it did not completely eliminate the presence of hotspots on the spatial grids. Moreover, many hotspots with the highest discrimination performance before PV suppression were the locations with the highest level of performance in the PV suppressed condition (Fig 3Bii). This suggests that of PV

469 neurons alone do not control spatial configuration sensitivity in the spatial grids, but instead
470 decrease discrimination performance across the majority of hotspots. Based on these observations,
471 we hypothesize that a separate cell-type (X neurons in Fig 5) mediate cross-channel inhibition,
472 resulting in the generation of hotspots and the specific pattern of spatial configuration sensitivity
473 on spatial grids. This is also consistent with our observation that a few sites showed emergent
474 hotspots upon optogenetic suppression of PV neurons (3 of 83 in the clean condition and 9 of 45
475 in the masked condition) (see Target 0˚, Masker 0˚ location in Fig 3B: 67% compared to 70%
476 under laser condition for example). Such emergent hotspots may arise via unmasking of network
477 connections upon suppression of PV neurons (Fig 5B), consistent with network connectivity
478 between excitatory pyramidal, PV and SOM neurons (Pfeffer et al., 2013). A candidate cell type
479 that correspond to X cells are somatostatin-positive (SOM) neurons, which have been implicated
480 in di-synaptic feedback inhibition (Kapfer et al., 2007; Silberberg and Markram, 2007) and
481 surround suppression (Adesnik et al., 2012; Lakunina et al., 2020).

482

483 # References

484 Adesnik H, Bruns W, Taniguchi H, Huang ZJ, Scanziani M (2012) A neural circuit for spatial
485 summation in visual cortex. Nature 490:226-231.
486 Aizenberg M, Mwilambwe-Tshilobo L, Briguglio JJ, Natan RG, Geffen MN (2015) Bidirectional
487 Regulation of Innate and Learned Behaviors That Rely on Frequency Discrimination by
488 Cortical Inhibitory Neurons. PLoS Biol 13:e1002308.

489 Antonoudiou P, Tan YL, Kontou G, Upton AL, Mann EO (2020) Parvalbumin and Somatostatin
490 Interneurons Contribute to the Generation of Hippocampal Gamma Oscillations. J
491 Neurosci 40:7668-7687.
492 Briguglio JJ, Aizenberg M, Balasubramanian V, Geffen MN (2018) Cortical Neural Activity
493 Predicts Sensory Acuity Under Optogenetic Manipulation. J Neurosci 38:2094-2105.
494 Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a
495 comparison of neuronal and psychophysical performance. J Neurosci 12:4745-4765.
496 Bruno RM, Simons DJ (2002) Feedforward mechanisms of excitatory and inhibitory cortical
497 receptive fields. J Neurosci 22:10966-10975.
498 Cardin JA, Carlen M, Meletis K, Knoblich U, Zhang F, Deisseroth K, Tsai LH, Moore CI (2009)
499 Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature
500 459:663-667.
501 deCharms RC, Zador A (2000) Neural representation and the cortical code. Annu Rev Neurosci
502 23:613-647.
503 Dong J, Colburn HS, Sen K (2016) Cortical Transformation of Spatial Processing for Solving the
504 Cocktail Party Problem: A Computational Model(1,2,3). eNeuro 3.
505 Gritton HJ, Nocon JC, James NM, Lowet E, Abdulkerim M, Sen K, Han X (2020) Oscillatory
506 activity in alpha/beta frequencies coordinates auditory and prefrontal cortices during
507 extinction learning. bioRxiv.
508 James NM, Gritton HJ, Kopell N, Sen K, Han X (2019) Muscarinic receptors regulate auditory
509 and prefrontal cortical communication during auditory processing. Neuropharmacology
510 144:155-171.

511 Jang HJ, Chung H, Rowland JM, Richards BA, Kohl MM, Kwag J (2020) Distinct roles of
512 parvalbumin and somatostatin interneurons in gating the synchronization of spike times in
513 the neocortex. Sci Adv 6:eaay5333.
514 Kapfer C, Glickfeld LL, Atallah BV, Scanziani M (2007) Supralinear increase of recurrent
515 inhibition during sparse activity in the somatosensory cortex. Nat Neurosci 10:743-753.
516 Kato HK, Asinof SK, Isaacson JS (2017) Network-Level Control of Frequency Tuning in Auditory
517 Cortex. Neuron 95:412-423 e414.
518 Keller CH, Kaylegian K, Wehr M (2018) Gap encoding by parvalbumin-expressing interneurons
519 in auditory cortex. J Neurophysiol 120:105-114.
520 Kreuz T, Chicharro D, Houghton C, Andrzejak RG, Mormann F (2013) Monitoring spike train
521 synchrony. J Neurophysiol 109:1457-1472.
522 Lakunina AA, Nardoci MB, Ahmadian Y, Jaramillo S (2020) Somatostatin-Expressing
523 Interneurons in the Auditory Cortex Mediate Sustained Suppression by Spectral Surround.
524 J Neurosci 40:3564-3575.
525 Li LY, Xiong XR, Ibrahim LA, Yuan W, Tao HW, Zhang LI (2015) Differential Receptive Field
526 Properties of Parvalbumin and Somatostatin Inhibitory Neurons in Mouse Auditory
527 Cortex. Cereb Cortex 25:1782-1791.
528 Li LY, Ji XY, Liang F, Li YT, Xiao Z, Tao HW, Zhang LI (2014) A feedforward inhibitory circuit
529 mediates lateral refinement of sensory representation in upper layer 2/3 of mouse primary
530 auditory cortex. J Neurosci 34:13670-13683.
531 Liang F, Li H, Chou XL, Zhou M, Zhang NK, Xiao Z, Zhang KK, Tao HW, Zhang LI (2019)
532 Sparse Representation in Awake Auditory Cortex: Cell-type Dependence, Synaptic
533 Mechanisms, Developmental Emergence, and Modulation. Cereb Cortex 29:3796-3812.

534 Lu T, Liang L, Wang X (2001) Temporal and rate representations of time-varying signals in the
535 auditory cortex of awake primates. Nat Neurosci 4:1131-1138.
536 Maddox RK, Billimoria CP, Perrone BP, Shinn-Cunningham BG, Sen K (2012) Competing sound
537 sources reveal spatial effects in cortical processing. PLoS Biol 10:e1001319.
538 Mesgarani N, Chang EF (2012) Selective cortical representation of attended speaker in multi-talker
539 speech perception. Nature 485:233-236.
540 Moore AK, Wehr M (2013) Parvalbumin-expressing inhibitory interneurons in auditory cortex are
541 well-tuned for frequency. J Neurosci 33:13713-13723.
542 Narayan R, Ergun A, Sen K (2005) Delayed inhibition in cortical receptive fields and the
543 discrimination of complex stimuli. J Neurophysiol 94:2970-2975.
544 Narayan R, Best V, Ozmeral E, McClaine E, Dent M, Shinn-Cunningham B, Sen K (2007) Cortical
545 interference effects in the cocktail party problem. Nat Neurosci 10:1601-1607.
546 Natan RG, Rao W, Geffen MN (2017) Cortical Interneurons Differentially Shape Frequency
547 Tuning following Adaptation. Cell Rep 21:878-890.
548 Natan RG, Briguglio JJ, Mwilambwe-Tshilobo L, Jones SI, Aizenberg M, Goldberg EM, Geffen
549 MN (2015) Complementary control of sensory adaptation by two types of cortical
550 interneurons. Elife 4.
551 Pachitariu M, Steinmetz N, Kadir S, Carandini M, Harris KD (2016) Kilosort: realtime spike-
552 sorting for extracellular electrophysiology with hundreds of channels. bioRxiv.
553 Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M (2013) Inhibition of inhibition in visual cortex:
554 the logic of connections between molecularly distinct interneurons. Nat Neurosci 16:1068-
555 1076.

556 Phillips EA, Hasenstaub AR (2016) Asymmetric effects of activating and inactivating cortical
557 interneurons. Elife 5.
558 Rossant C, Kadir SN, Goodman DFM, Schulman J, Hunter MLD, Saleem AB, Grosmark A,
559 Belluscio M, Denfield GH, Ecker AS, Tolias AS, Solomon S, Buzsaki G, Carandini M,
560 Harris KD (2016) Spike sorting for large, dense electrode arrays. Nat Neurosci 19:634-
561 641.
562 Rothauser EH, Chapman, W. D., Guttman, N., Nordby, K. S., Silbiger, H. R., Urbanek, G. E., &
563 Weinstock, M. (1969) I.E.E.E. recommended practice for speech quality measurements.
564 IEEE Trans Audio Electroacoust 17:225-246.
565 Satuvuori E, Kreuz T (2018) Which spike train distance is most suitable for distinguishing rate
566 and temporal coding? J Neurosci Methods 299:22-33.
567 Scala F, Kobak D, Bernabucci M, Bernaerts Y, Cadwell CR, Castro JR, Hartmanis L, Jiang X,
568 Laturnus S, Miranda E, Mulherkar S, Tan ZH, Yao Z, Zeng H, Sandberg R, Berens P,
569 Tolias AS (2020) Phenotypic variation of transcriptomic cell types in mouse motor cortex.
570 Nature.
571 Silberberg G, Markram H (2007) Disynaptic inhibition between neocortical pyramidal cells
572 mediated by Martinotti cells. Neuron 53:735-746.
573 Sohal VS, Zhang F, Yizhar O, Deisseroth K (2009) Parvalbumin neurons and gamma rhythms
574 enhance cortical circuit performance. Nature 459:698-702.
575 Theunissen F, Miller JP (1995) Temporal encoding in nervous systems: a rigorous definition. J
576 Comput Neurosci 2:149-162.
577 Tremblay R, Lee S, Rudy B (2016) GABAergic Interneurons in the Neocortex: From Cellular
578 Properties to Circuits. Neuron 91:260-292.

579 Vinje WE, Gallant JL (2000) Sparse coding and decorrelation in primary visual cortex during
580 natural vision. Science 287:1273-1276.
581 Wang L, Narayan R, Grana G, Shamir M, Sen K (2007) Cortical discrimination of complex natural
582 stimuli: can single neurons match behavior? J Neurosci 27:582-589.
583 Xiang Z, Huguenard JR, Prince DA (1998) Cholinergic switching within neocortical inhibitory
584 networks. Science 281:985-988.
585 Yao JD, Sanes DH (2021) Temporal Encoding is Required for Categorization, But Not
586 Discrimination. Cereb Cortex 31:2886-2897.
587 Zhu Y, Qiao W, Liu K, Zhong H, Yao H (2015) Control of response reliability by parvalbumin-
588 expressing interneurons in visual cortex. Nat Commun 6:6802.
589 Zuo Y, Safaai H, Notaro G, Mazzoni A, Panzeri S, Diamond ME (2015) Complementary
590 contributions of spike timing and spike rate to perceptual decisions in rat S1 and S2 cortex.
591 Curr Biol 25:357-363.
592

593 **Figure 1. Methodology and experimental design. A:** Illustration depicting recording electrode
594 location and optical fiber placement. Subjects were implanted with a 4-shank, 32-channel electrode
595 and optogenetic fiber in right hemisphere of A1. Each shank contained 8 sites per shank with
596 100µm spacing between electrode contacts. **B:** Representative local field potential (LFP) activity
597 from one mouse. **C:** Average z-scored spiking activity from all subjects. Z-scoring was done by
598 grouping all channels with the same electrode depth (n = 9 mice, 36 shanks). **D:** Schematic for
599 control and optogenetic trial presentation. During 50% of all trials, a 532nm laser would turn on
600 50ms before sound stimulus onset and turn off during sound offset. **E:** Example mean single unit
601 waveform and correlogram from data set. Dashed lines in mean waveform represent one standard
602 deviation above and below the mean, while scale bars are equal to 200µV and 1ms. Dashed red
603 lines in correlogram represent ISIs of ±2ms.

604

605 **Figure 2. Cortical discrimination in a cocktail party paradigm in mouse A1. A:** Illustration
606 depicting the stimulus configuration for clean trials originating at +90˚ azimuth (**Ai**) and responses
607 to both target stimuli (T) (**Aii**). Auditory stimuli were presented from speakers at 4 locations.
608 Target stimuli consisted of white noise modulated by human speech envelopes extracted from
609 recordings of speech sentences (see Methods, 'Auditory stimuli'). As shown in Aii, responses
610 during clean trials exhibit spike timing and rapid firing rate modulation that follow the amplitude
611 envelope of both target stimuli. **Bi:** Stimulus configuration for trials where targets (T) played at
612 +90˚ and a competing masking stimulus (M) played at -90˚. Masking stimuli consisted of
613 unmodulated white noise with the same onset and offset times as target stimuli. **Bii**: Responses to
614 masked trials shown in Bi. In this configuration, spike timing and firing rate modulation follow
615 both target stimuli, despite the presence of the competing masker. **Ci:** Stimulus configuration for
616 trials where targets played at 0˚ and maskers played at +45˚. **Cii**: Responses to target-masker
617 configuration shown in Ci. For this configuration, spike timing and firing rate modulation do not
618 follow either target stimulus, resulting in similar responses between target identities. **D:** Neural
619 discriminability performance for all possible target-masker location configurations, referred to as
620 the "spatial grid", for example cell featured in A-C. Outlined spots indicate configurations shown
621 in A-C, matched by the outline color. Performance is calculated using a template-matching
622 approach based on differences in instantaneous firing rate and spike timing similarities between
623 trains (see Methods, 'Neural discriminability using SPIKE-distance'). The top grid shows
624 discriminability for "clean" trials on top, while the bottom grid shows discriminability for
625 "masked" trials. All blocks are color-coded according to the color axis shown to the right of the
626 masked grid. Configurations with high performance (≥ 70%), a large effect size ($d \geq 1$) and a
627 significant difference ($p < 0.05$) from the chance distribution, e.g., the configurations outlined in

black and red, are referred to as “hotspots”. **E:** Effect sizes for each spatial grid configuration in D (see Methods, ‘Neural discriminability using SPIKE-distance’), with the same outlines corresponding to examples in A-C. Positive values represent an increase in performance relative to a chance distribution, while negative values represent a decrease in performance relative to chance.

**Figure 3. Effects of suppressing PV neurons on cortical discrimination. A:** Responses during clean stimulus trials originating at 45˚ from one example cell during control (**Ai**) and optogenetic (**Aii**) conditions. PV suppression resulted in an increase in the baseline firing rate, effectively reducing the modulation depth of the neural response. Inset (**Aiii**) shows a zoomed-in portion of the response between 0.1 and 0.4s after sound onset, as outlined in Ai and Aii. Responses during optogenetic trials show less phase-locking to stimulus features, compared to the control. **B:** Spatial grids from same example cell during control (**Bi**) and optogenetic (**Bii**) conditions, with the clean configurations at 45˚ outlined to correspond to responses shown in Ai-ii. Performance is color-coded according to the axis shown to the right of Bii. In this example, performance decreased during optogenetic suppression for both clean (see all clean spots in Bi versus Bii) and masked trials (note hotspots at target 45˚, masker 90˚ and target 45˚, masker -90˚ in Bi versus the same spots in Bii). **C:** Paired comparisons of SPIKE-distance-based performance from control and PV-suppressed trials at the same spatial grid location. Paired t-tests yielded a significant decrease in performance for both clean (n = 80 configurations; p = 5.99e-05) and masked (n = 38 configurations; p = 9.82e-04) trials during optogenetic suppression, indicating that optogenetic suppression reduced performance by affecting spike timing and rapid firing rate modulation during stimulus presentation. **D:** Paired comparisons between conditions for mean firing rate during stimulus playback. Paired t-tests yielded a significant increase in mean firing rate for both clean (n = 80 configurations; p = 1.60e-11) and masked (n = 38 configurations; p = 0.0084) trials during optogenetic suppression. **E:** Paired temporal sparseness comparisons between conditions. Temporal sparseness measures the distribution of a unit's response over time. Paired t-tests yielded a significant decrease in sparseness during both clean (p = 1.30e-14) and masked (p = 0.0325) optogenetic trials.

**Figure 4. Effects of suppressing 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 (n = 80 configurations; p = 2.05e-05) and masked (n = 38 configurations; p = 7.05e-05) 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 = 1.25e-07) and masked (p = 3.84e-05) 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 clean trials (p = 0.3561) but showed a weakly significant decrease for masked (p = 0.0480) trials. **D:** Summary figure showing contributions from spike distance measures presented in Figure 4C and 5A-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 only weakly significant during masked trials, indicating that total spike count is insufficient to explain the decrease in overall discrimination performance.

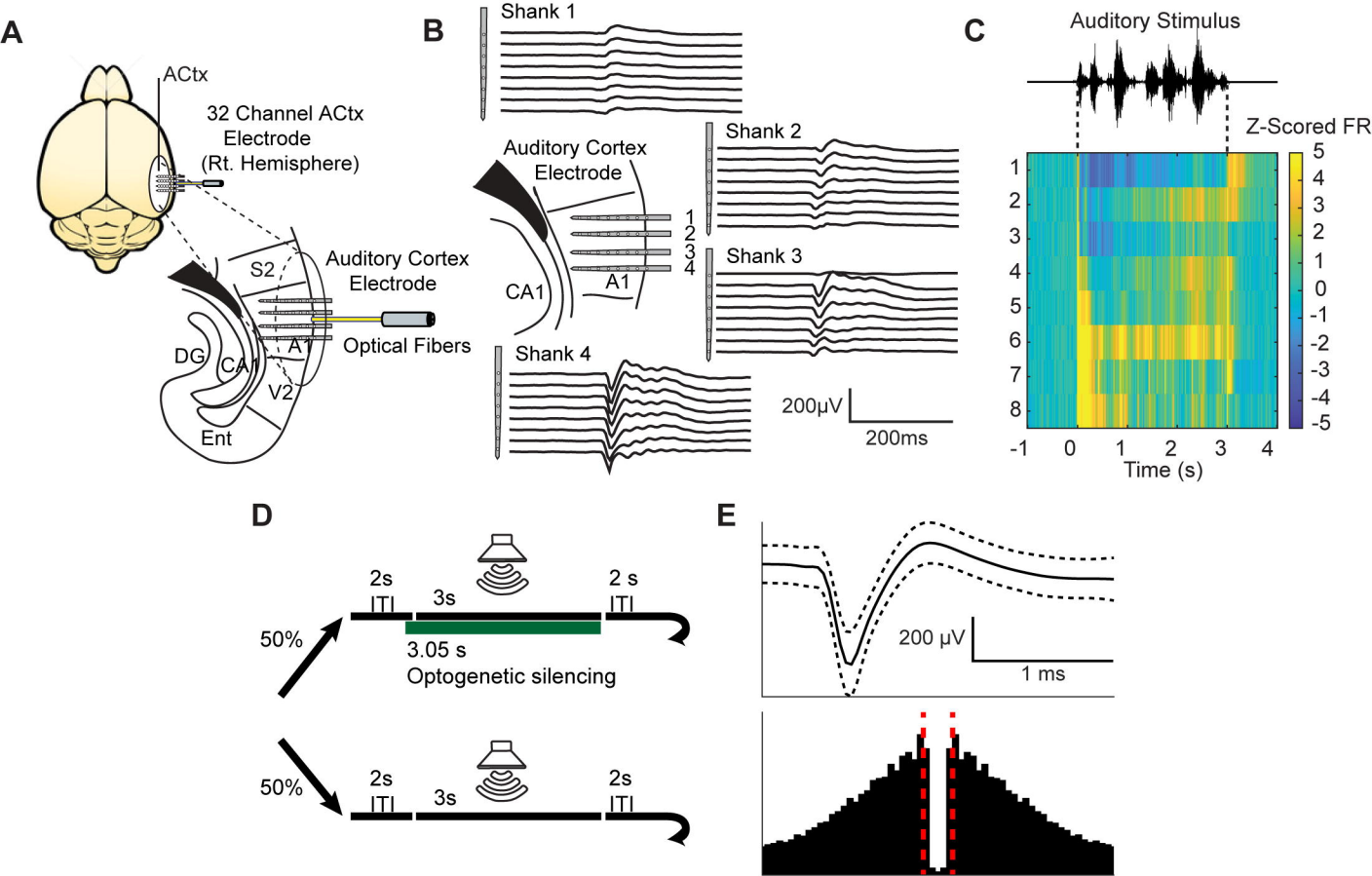
675 **Figure 5. Cortical circuits for complex scene analysis.** Schematic model representing cortical
676 circuit under the control condition (**A**) and in the laser condition (**B**). C cells represent cortical
677 units, R cells represent excitatory relay units, I cells represent within-channel inhibitory units, and
678 X cells represent cross-channel inhibitory units.

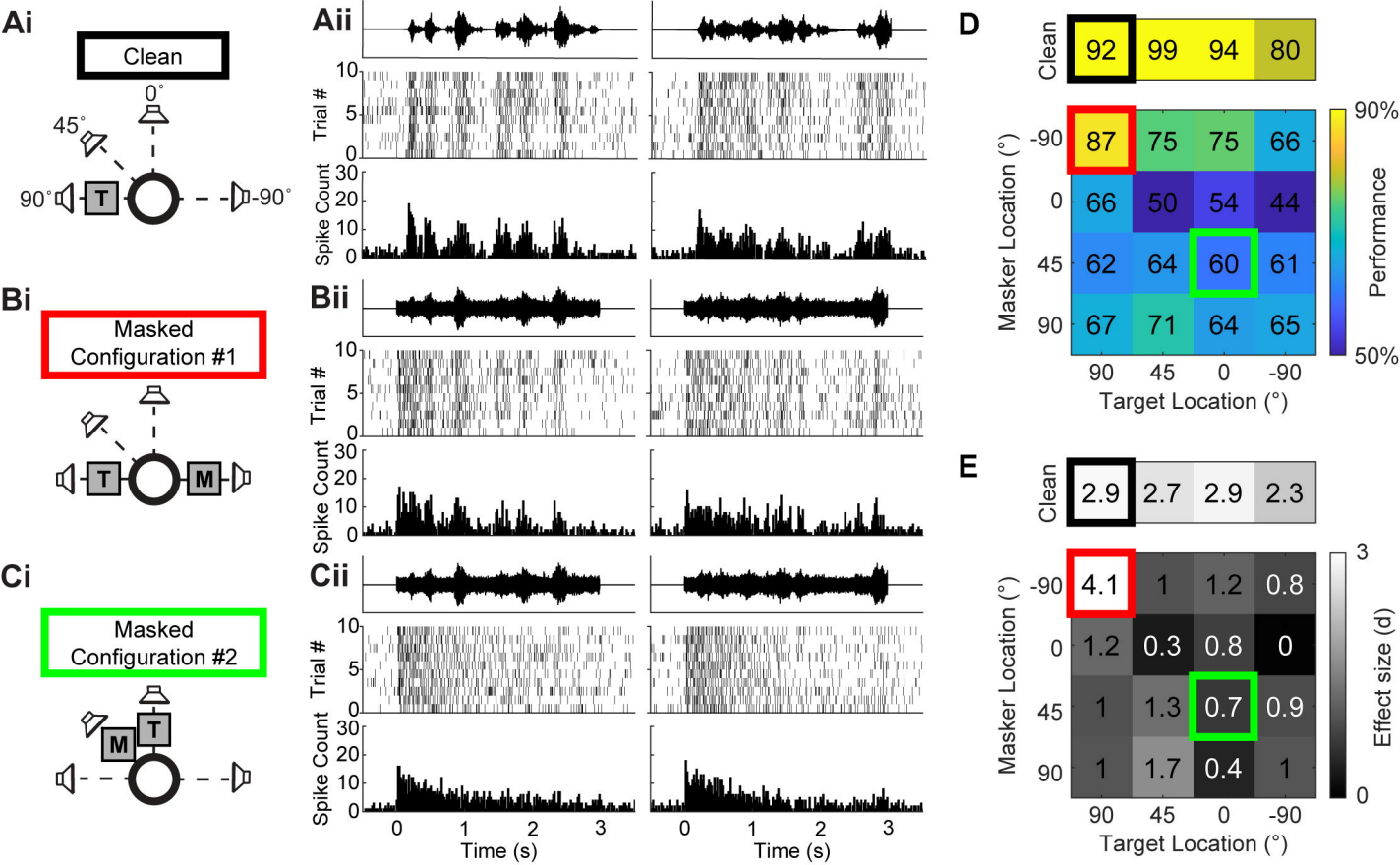
679

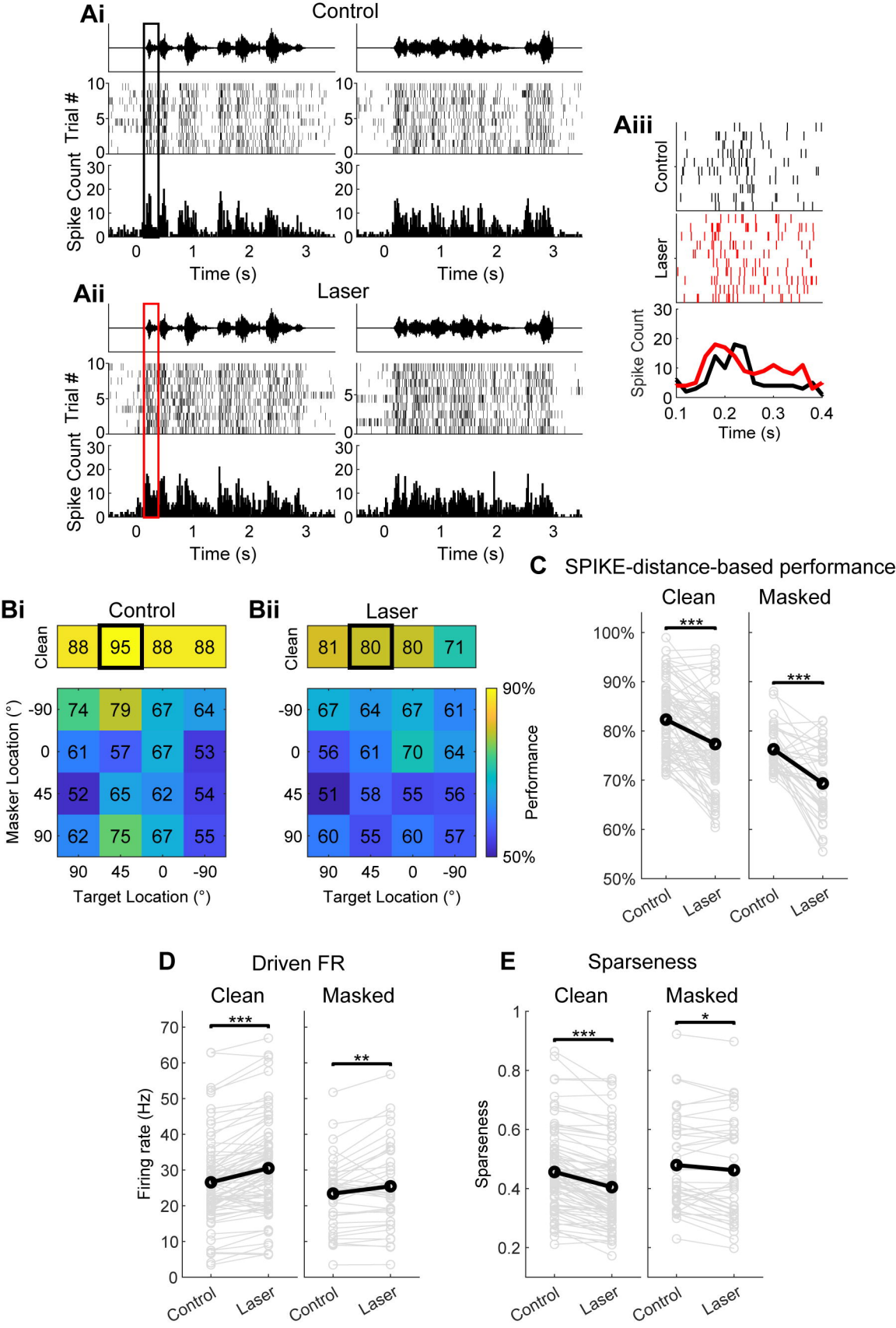
680 **Extended Data Figure 3-1. Effects of optogenetic strength on firing rate and discriminability.**
681 **A:** Traces of average spontaneous firing rate from 50ms before to 50ms after laser onset for control
682 trials and all laser strengths. During this time window, no auditory stimuli were presented to the
683 mice. The control trace was estimated by pooling all non-optogenetic trials from each of the three
684 optogenetic blocks. **Bi:** Scatter plot of spontaneous firing rate during the 50ms after laser onset in
685 control and optogenetic conditions for all three laser strengths. Each point shows the average
686 spontaneous firing rate at a given spatial grid hotspots in both conditions, with 96 hotspots from
687 the 2mW block (black), 94 hotspots from the 5mW block (red), and 118 hotspots from the 10mW
688 block (green). Blue dashed line represents unity values. **Bii:** Difference in spontaneous firing rate
689 between conditions for each laser strength (mean ± s.e.m.). One-way ANOVA yielded laser
690 strength as a significant factor ($F(2,305) = 5.5648$; $p = 0.0042$), with a significant post-hoc
691 difference in spontaneous firing rate between 2mW and 10mW trials ($p = 0.0026$). **C:** Traces of
692 average evoked firing rate from 50ms before to 50ms after sound stimulus onset for control trials
693 and all laser strengths. The control trace was estimated by pooling all control trials from each of
694 the three optogenetic blocks. **Di:** Scatter plot of evoked firing rate during the 3s of stimulus
695 playback in control and optogenetic conditions for all three laser strengths. Open dots represent
696 data from clean trials while filled dots represent data from masked trials. **Dii:** Difference in evoked
697 firing rate between conditions for each laser strength (mean ± s.e.m.). Clean and masked firing
698 rates were pooled before one-way ANOVA, which yielded laser strength as a significant factor
699 ($F(2,305) = 9.8290$; $p = 7.30e-05$). Post-hoc comparisons found a significance difference in evoked
700 firing rate between 2mW and 10mW trials ($p = 2.79e-05$), with differences trending significant
701 between 2mW and 5mW trials ($p = 0.0746$) and between 5mW and 10mW trials ($p = 0.0864$). The
702 effects of optogenetically suppressing PV neurons on both spontaneous and evoked firing rate

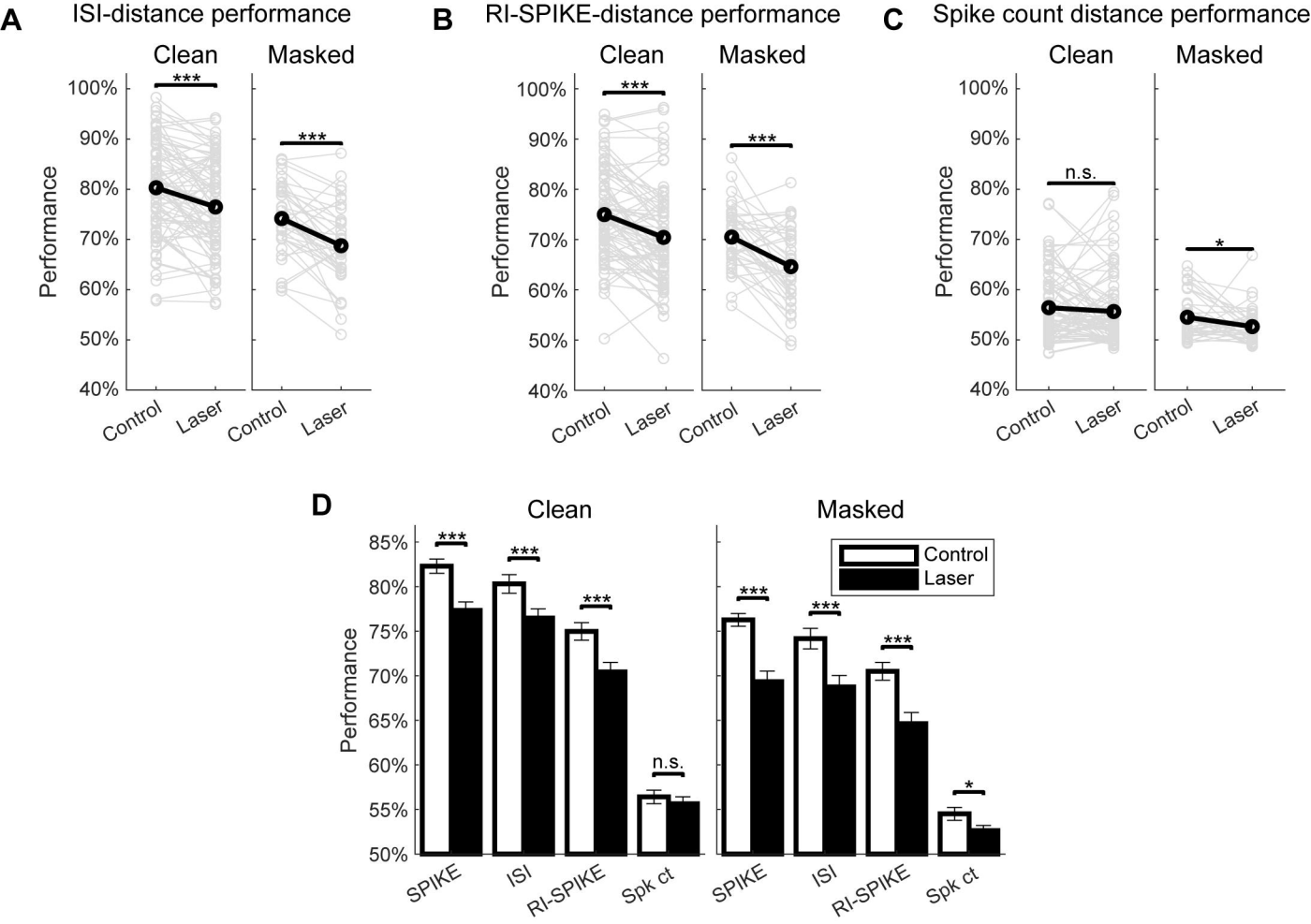
703 were consistent with past findings in A1 (Aizenberg et al., 2015; Phillips and Hasenstaub, 2016;
704 Natan et al., 2017). **Ei:** Scatter plot of neural discriminability performance in control and
705 optogenetic conditions for all three laser strengths. Open dots represent data from clean trials while
706 filled dots represent data from masked trials. **Eii:** Difference in discriminability between
707 conditions for each laser strength (mean ± s.e.m.). Clean and masked performances were pooled
708 before one-way ANOVA, which yielded laser strength as a significant factor ($F(2,305) = 4.1818$;
709 $p = 0.0162$). In post-hoc tests, we found that optogenetic suppression during 10mW trials yielded
710 a significantly larger decrease in performance than in 2mW trials ($p = 0.0107$).

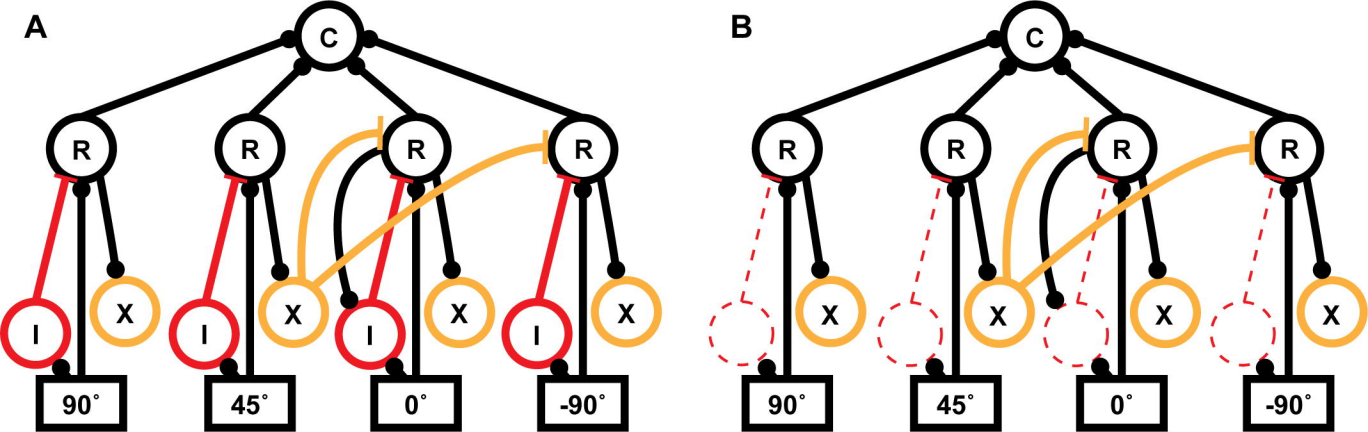
711


A
ACtx
32 Channel ACtx
Electrode
(Rt. Hemisphere)
S2
Auditory Cortex
Electrode
Optical Fibers
DG
CA1
A1
V2
Ent
B
Shank 1
Shank 2
Shank 3
Shank 4
Auditory Cortex
Electrode
CA1
A1
1
2
3
4
200μV
200ms
C
Auditory Stimulus
Z-Scored FR
1
2
3
4
5
6
7
8
5
4
3
2
1
0
-1
-2
-3
-4
-5
-1
0
1
2
3
4
Time (s)
D
2s
ITI
3s
2 s
ITI
50%
3.05 s
Optogenetic silencing
50%
2s
ITI
3s
2s
ITI
E
200 μV
1 ms


Ai
Clean
0°
45°
90°
-90°
T
Bi
Masked Configuration #1
T
M
Ci
Masked Configuration #2
M
T
Aii
Bii
Cii
Trial #
Spike Count
Time (s)
D
Clean
92 99 94 80
Masker Location (°)
-90 87 75 75 66
0 66 50 54 44
45 62 64 60 61
90 67 71 64 65
90 45 0 -90
Target Location (°)
90%
50%
Performance
E
Clean
2.9 2.7 2.9 2.3
-90 4.1 1 1.2 0.8
0 1.2 0.3 0.8 0
45 1 1.3 0.7 0.9
90 1 1.7 0.4 1
3
0
Effect size (d)


Ai
Control
Trial #
Spike Count
Time (s)
Aii
Laser
Aiii
Control
Laser
Spike Count
Time (s)
Bi
Control
Clean
88 95 88 88
Masker Location (°)
-90 74 79 67 64
0 61 57 67 53
45 52 65 62 54
90 62 75 67 55
90 45 0 -90
Target Location (°)
Bii
Laser
Clean
81 80 80 71
-90 67 64 67 61
0 56 61 70 64
45 51 58 55 56
90 60 55 60 57
90 45 0 -90
Target Location (°)
90%
Performance
50%
C
SPIKE-distance-based performance
Clean
Masked
100%
90%
80%
70%
60%
50%
***
***
Control
Laser
Control
Laser
D
Driven FR
Clean
Masked
Firing rate (Hz)
***
**
Control
Laser
Control
Laser
E
Sparseness
Clean
Masked
Sparseness
***
*
Control
Laser
Control
Laser


A
ISI-distance performance
Clean
Masked
Performance
100%
90%
80%
70%
60%
50%
40%
***
***
Control
Laser
Control
Laser
B
RI-SPIKE-distance performance
Clean
Masked
Performance
***
***
Control
Laser
Control
Laser
C
Spike count distance performance
Clean
Masked
Performance
n.s.
*
Control
Laser
Control
Laser
D
Clean
Masked
Control
Laser
Performance
85%
80%
75%
70%
65%
60%
55%
50%
***
***
***
n.s.
***
***
***
*
SPIKE
ISI
RI-SPIKE
Spk ct
SPIKE
ISI
RI-SPIKE
Spk ct


A
C
R
R
R
R
I
X
I
X
I
X
I
X
90˚
45˚
0˚
-90˚
B
C
R
R
R
R
X
X
X
X
90˚
45˚
0˚
-90˚