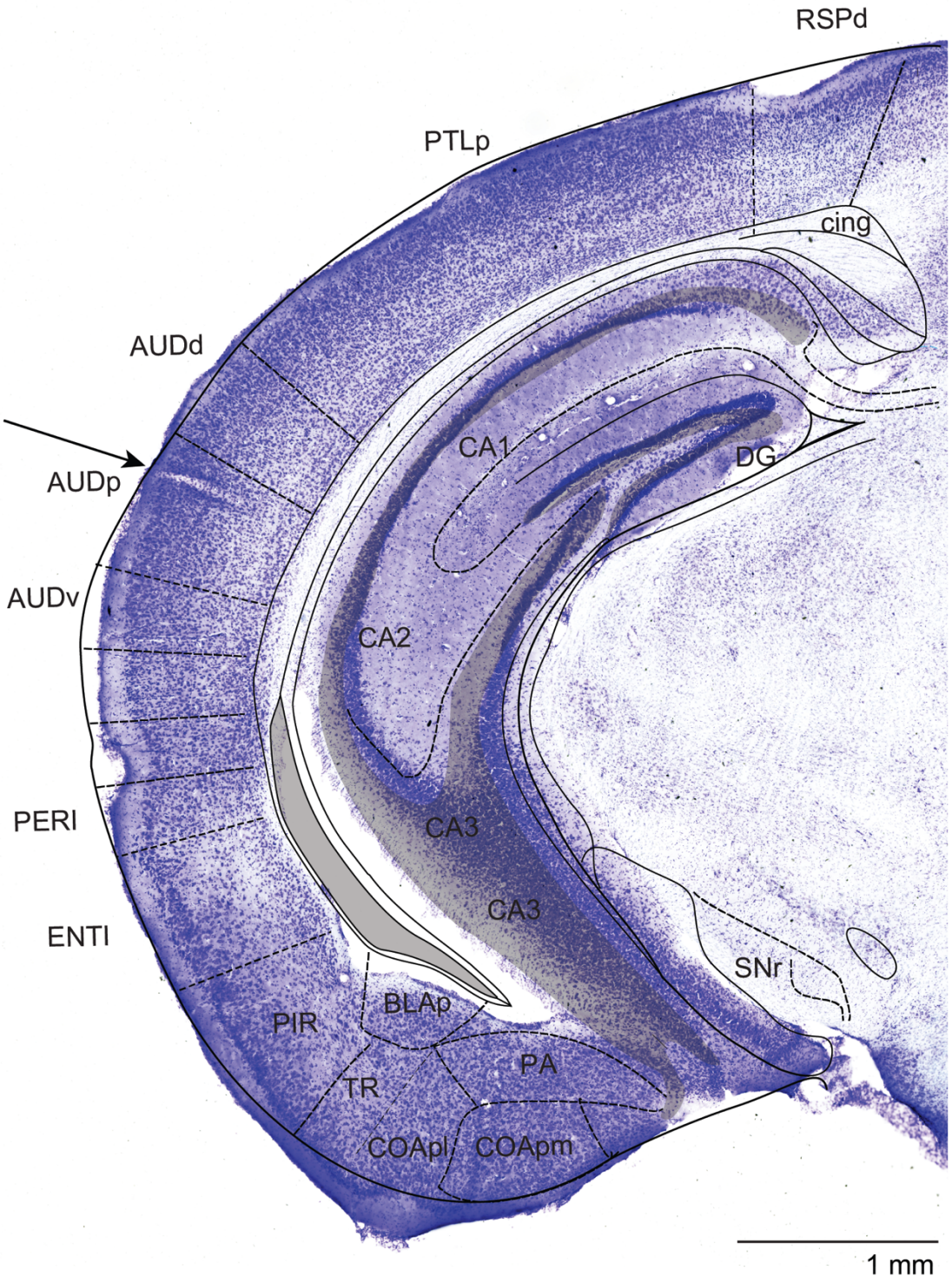
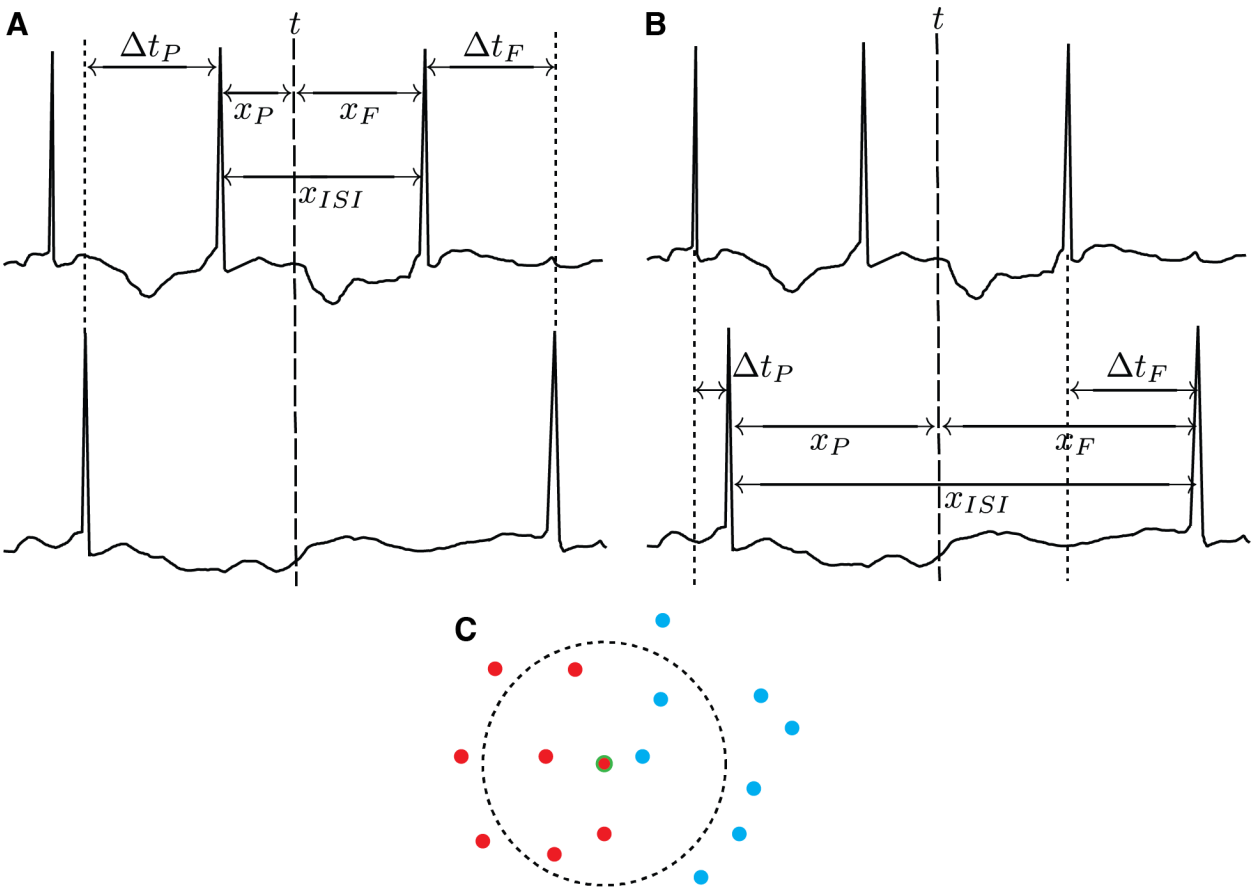
Subject	L1	L2/3	L4	L5	L6	Total
1	0	2	4	1	0	7
2	0	3	0	3	0	6
3	0	0	0	1	0	1
4	0	1	0	0	0	1
5	0	2	0	3	0	5
6	0	1	0	1	0	2
7	0	0	1	0	0	1
8	0	3	1	2	0	6
9	0	0	2	1	0	3
10	0	2	0	0	0	2
11	1	3	2	0	1	7
12	0	1	1	1	1	4
Total	1	18	11	13	2	45

Supplemental Table 1. Count of single units with at least one configuration (target-only or target with competing masker) of high neural discriminability per subject and per layer.



Supplemental Figure 1. Electrode placement in mouse auditory cortex. Figure is reproduced from Supplementary Figure 1 in (1). Representative image from a nissl stained histological section showing electrode location. Arrow indicates location of electrode placement within coronal section that targets the primary auditory cortex (AUDp). Image is overlaid with topography based on the Allen mouse brain atlas. AUDd: dorsal auditory area; AUDp: primary auditory area; AUDv: ventral auditory area; BLAp: basolateral amygdalar nucleus, posterior part; CA1-3: fields CA1-3;

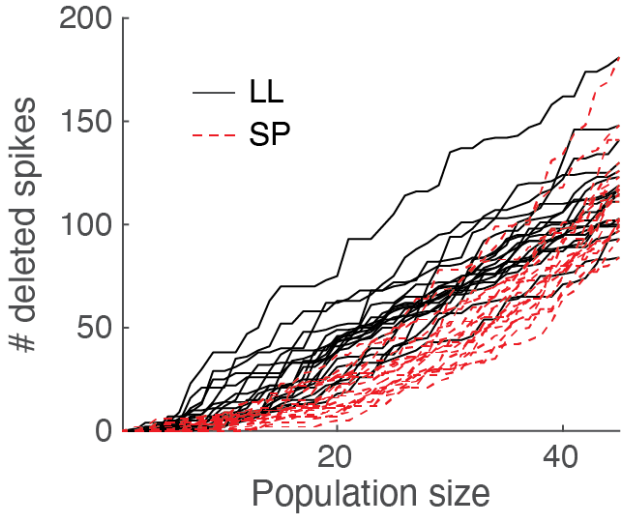
cing: cingulum bundle; COApl: cortical amygdalar area, lateral part; COApm: cortical amygdalar area, medial part; DG: dentate gyrus; ENTI: entorhinal area, lateral part; PA: posterior amygdalar nucleus; PERI: perirhinal area; PIR: piriform area; PTLp: posterior parietal association area; RSPd: retrosplenial area, dorsal part; SNr: substantia nigra, reticular part; TR: postpiriform transition area.



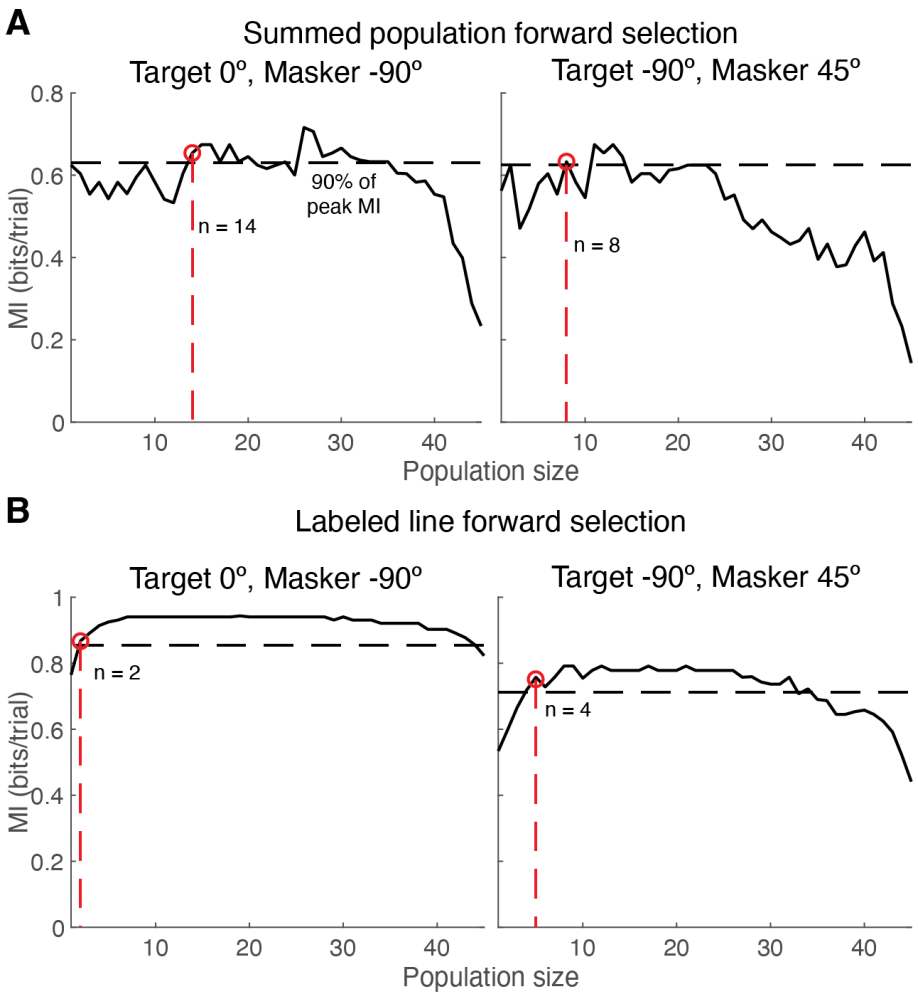
Supplemental Figure 2. Illustrating the calculation of the spike train distance and the mutual information. Two example spike trains demonstrating the different quantities required for calculating SPIKE-distance; **(A)** shows the SPIKE-distance quantities for spike train 1 (top), and **(B)** shows the quantities for spike train 2 (bottom). In Equation 3, these would correspond to the (1) and (2) superscripts, which have been left out here for clarity. **(C)** Calculating the mutual information from each data point is considered in turn; in our data, this would correspond to a spike train: this point then forms the seed point (circled in green). A ball is expanded around the seed point until it contains h-1 additional points, where h is a smoothing parameter. The number of data points corresponding to the same label as the seed point (h_i, red) is then counted. In our dataset, this label corresponds to the target identity; here, $h = 7$ and $h_i = 4$, the number of red points, excluding the seed. Now, the mutual information is:

$$I_0 = \frac{1}{n}\sum_{i=1}^{n} \log_2 \frac{n_x h_x(i)}{h} \tag{S1}$$

where n is the total number of data points and n_x is the number of different target labels. This is a bias estimator, and details on removing the bias and choosing h are given in (2).



Supplemental Figure 3. The number of coincident spikes increases with population size. y-axis measures the total amount of coincident spikes across all trials in both coding schemes (labeled line (LL), solid black lines; summed population (SP), dashed red lines) as a function of population size. Each curve represents a forward search for one configuration of target and masker location. In both schemes, a full population response (n = 45) resulted in an average deletion of ~10 spikes per trial.



Supplemental Figure 4. Forward search in the summed population approach shows non-monotonic changes in MI. (A) Forward selection searches in the summed population code for two masked configurations in Figure 4A where the best single-unit MI ($n = 1$) outperformed certain larger population sizes before reaching the optimal subpopulation K_{opt} (indicated by the red line and circle). Dashed black lines indicate 90% of the maximal MI value across all population sizes in each curve, which was used as the threshold to determine K_{opt}. **(B)** Forward selection searches in the labeled line code for the same two configurations in A. In this coding scheme, the MI saturates at very small values of *n*, and adding additional neurons to the optimal subpopulation results in very little change in the MI value before experiencing a decrease at large values of *n*.

Supplemental References

1. **Nocon JC, Gritton HJ, James NM, Mount RA, Qu Z, Han X, and Sen K**. Parvalbumin neurons enhance temporal coding and reduce cortical noise in complex auditory scenes. *Commun Biol* 6: 751, 2023.
2. **Witter J, and Houghton C**. A note on the unbiased estimation of mutual information. 2021, p. arXiv:2105.08682.