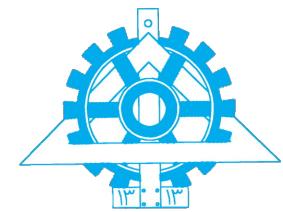




University of Tehran
College of Engineering
School of Electrical & Computer Eng



Principles of Cognitive Science

Assignment 2 Phase One

Name

Reza Chehreghani

Student ID

810101401

May 24, 2025

Contents

1	Spike Sorting from Scratch	1
1.1	Getting Started	1
1.2	Filtering the Data	3
1.3	Detecting the Spikes	4
1.4	Extracting Features	4
1.5	Clustering the Spikes	4
1.6	Discussion and Further Analysis	7
2	Spike Sorting with ROSS	11
2.1	Detection	11
2.2	Auto Sorting	12
2.3	Manual Sorting	13
2.4	Visualization	17
3	Analysis of Single Neuron Activity	21
3.1	PSTH Analysis	21
3.2	Fano Factor Analysis	22
3.2.1	Fano Factor	22
3.2.2	Mean-Matched Fano Factor	22
3.3	Category Classification Using SVM	24
3.4	Time-Time Decoding Analysis	25
3.5	Mutual Information Analysis Across Time	27
3.6	Category Discriminability Using d-prime	28
4	Analysis of Population Activity	30
4.1	Representational Dissimilarity Matrix (RDM) and Kendall's Tau Correlation	30
4.2	Generalized Linear Model (GLM) to Predict Ground Truth RDM	32

List of Figures

1	Raw extracellular voltage trace over time. The signal exhibits both low-frequency fluctuations and high-frequency components associated with neuronal spikes.	1
2	Histogram of recorded voltage amplitudes. The distribution is approximately Gaussian, indicative of background noise in the recording.	2
3	Comparison of raw (blue) and high-pass filtered (orange) signals. The filter suppresses low-frequency noise and emphasizes high-frequency spike components.	3
4	Extracted spike waveforms. Although the waveforms share a similar shape, amplitude variations suggest activity from multiple neurons.	4
5	Pairwise PCA scatter plots for $k = 1$ to 5. Each point represents a spike, colored by cluster assignment.	5
6	Pairwise PCA scatter plots for $k = 6$ to 10. Each point represents a spike, colored by cluster assignment.	6
7	Elbow method to determine the optimal number of clusters. A clear elbow appears at $k = 3$	7
8	t-SNE projection of spike waveforms for $k = 1$ to 5. Clusters are visualized using different colors.	8
9	t-SNE projection of spike waveforms for $k = 6$ to 10. Clusters are visualized using different colors.	9
10	Comparison of elbow method and silhouette score on t-SNE features. The silhouette score peaks at $k = 4$, suggesting four distinct clusters.	10
11	Spike detection in ROSS. Spikes are extracted from the filtered signal using an amplitude threshold based on estimated noise levels.	11
12	Automatic spike clustering using ROSS. Three initial clusters were identified based on waveform features.	12

13	Denoising process for cluster 1 using 85% threshold.	13
14	Denoising process for cluster 2 using 85% threshold.	14
15	Denoising process for cluster 3 using 85% threshold.	14
16	Final denoised clusters. Noise removal helped to better isolate spike waveforms from individual neurons.	15
17	Result of resorting Cluster 3. It was divided into three sub-clusters, increasing the total number of clusters to five.	16
18	Waveforms of the five clusters after spike sorting. Cluster 4 appears dominant in both frequency and spike count.	17
19	3D PCA plot showing temporal dynamics of spike clusters. Colors represent different clusters.	18
20	Raw signal with overlaid spike detections from all five clusters.	18
21	Example spike waveform from Cluster 1.	19
22	Example spike waveform from Cluster 2.	19
23	Example spike waveform from Cluster 3.	20
24	Example spike waveform from Cluster 4.	20
25	Example spike waveform from Cluster 5.	20
26	PSTH plots for neurons 5 and 62 across different stimulus categories. The horizontal axis represents time (ms), aligned to stimulus onset; the vertical axis shows the average firing rate (Hz).	21
27	Fano Factor over time for selected neurons. Higher values indicate greater variability in spike counts across trials.	22
28	Mean-Matched Fano Factor for each stimulus category over time. The MMFF accounts for differences in firing rate, providing a normalized measure of variability.	23
29	SVM classification results. The left plot shows the overall classification accuracy over time; the right plot displays recall values per category, reflecting how well each category was recognized.	24

30	Time-time decoding matrix with FDR-corrected significance. Each entry reflects decoding accuracy when the model is trained at time t_1 and tested at time t_2 . Significant regions ($p < 0.05$) are highlighted.	25
31	Mutual Information (MI) between neuronal activity and stimulus category as a function of time. The shaded region indicates the time period with statistically significant MI values (above the 95th percentile of the permutation-based null distribution).	27
32	Pairwise d-prime values for neuron 62. Each bar indicates the discriminability between a pair of stimulus categories. Higher values reflect better separability.	28
33	Kendall's tau correlation between the neural RDM and category-based ground truth RDM as a function of time. The peak alignment occurs around 100 ms post-stimulus.	30
34	Neural Representational Dissimilarity Matrix (RDM) at the time of peak category alignment (time slice 41, ≈ 100 ms). Lower dissimilarities (darker colors) within categories and higher dissimilarities (lighter colors) across categories indicate structured population encoding.	31
35	Variance explained (R^2 score) by a GLM trained on neural and visual feature RDMs to predict the ground truth RDM. The peak model performance occurs approximately 100–150 ms post-stimulus, highlighting the contribution of both neural and visual representations to category encoding.	32
36	PAC plots using the Modulation Index (MI) method for each stimulus category. The strength of coupling between low-frequency phase and high-frequency amplitude is visualized across frequency pairs.	34
37	PAC plots using the Canolty method. Compared to MI, this approach shows subtle differences across categories, with "Body" exhibiting stronger coupling in specific frequency regions.	35
38	Power spectrum across frequency bands for each stimulus category. All categories show dominance in low-frequency ranges, but the "Face" category exhibits reduced power in the high-frequency range compared to others.	36

Spike Sorting from Scratch

In this section, we implement a spike sorting pipeline from scratch to identify and classify neuronal spikes from extracellular recordings.

1.1 Getting Started

The dataset was first loaded and visualized. Figure 1 illustrates the recorded voltage amplitudes over time, showing the raw extracellular signal.

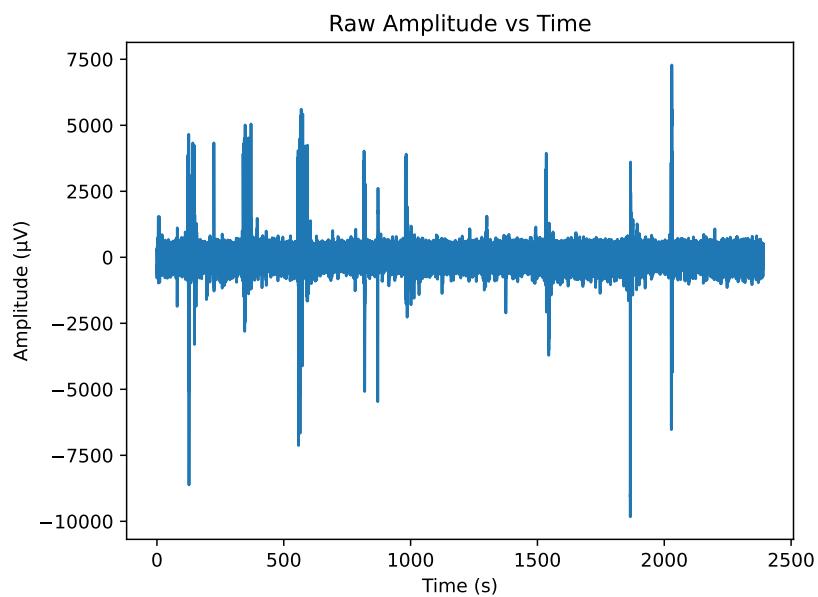


Figure 1: Raw extracellular voltage trace over time. The signal exhibits both low-frequency fluctuations and high-frequency components associated with neuronal spikes.

To better understand the signal distribution, a histogram of voltage amplitudes across the entire dataset was also plotted (Figure 2). The histogram suggests the presence of Gaussian-distributed background noise, typical in extracellular recordings.

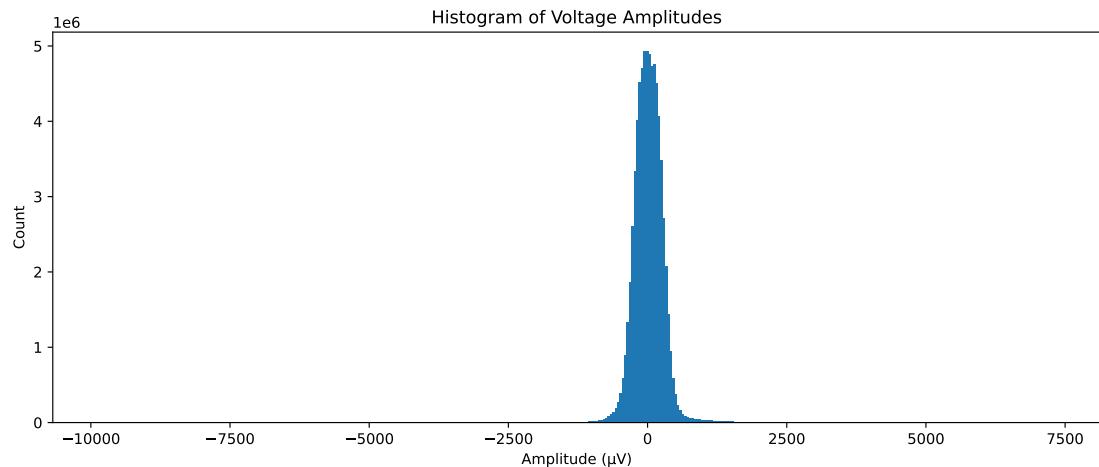


Figure 2: Histogram of recorded voltage amplitudes. The distribution is approximately Gaussian, indicative of background noise in the recording.

1.2 Filtering the Data

To isolate high-frequency neural activity, a 7th-order Butterworth high-pass filter with a cut-off frequency of 300 Hz was applied using zero-phase filtering (`filtfilt`). Figure 3 shows a comparison of a segment of raw and filtered signals.

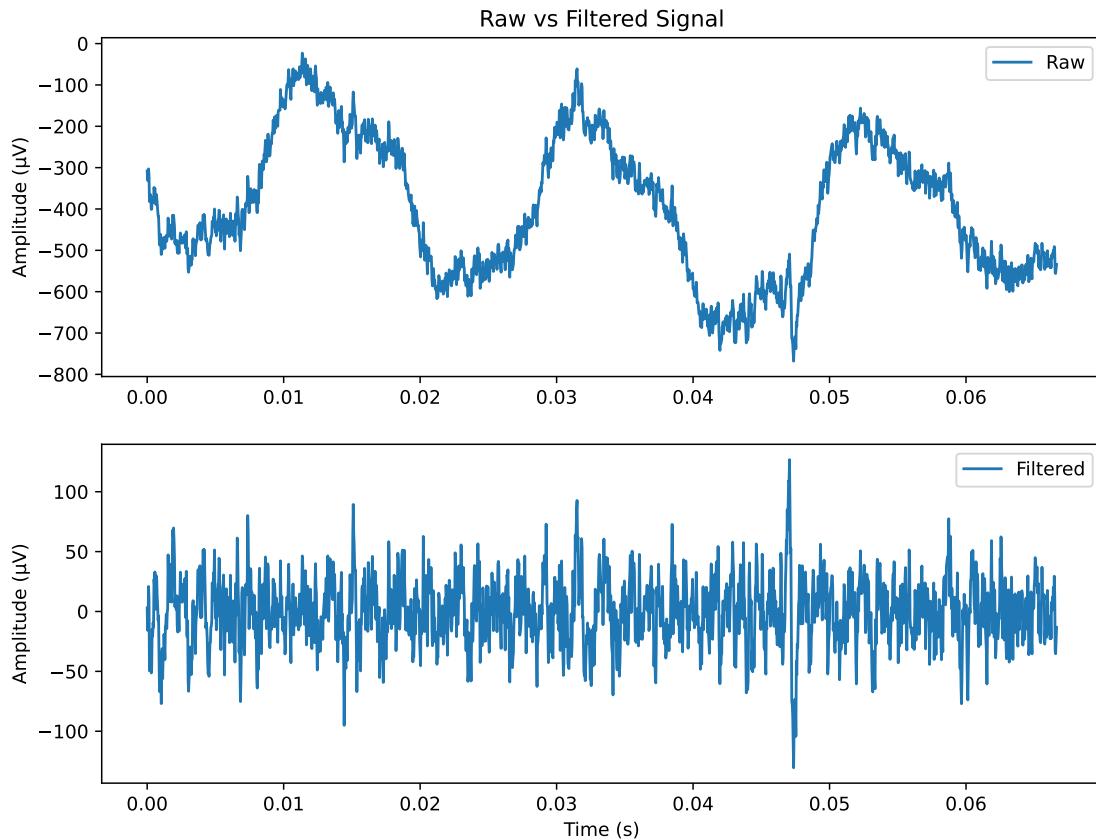


Figure 3: Comparison of raw (blue) and high-pass filtered (orange) signals. The filter suppresses low-frequency noise and emphasizes high-frequency spike components.

1.3 Detecting the Spikes

Spikes were detected by locating peaks that exceed a threshold $\theta = 5\sigma_n$, where $\sigma_n = \frac{\text{median}(|x|)}{0.6745}$. For each peak, a 4 ms window (2 ms before and after the peak, equivalent to 60 data points at 30 kHz) was extracted.

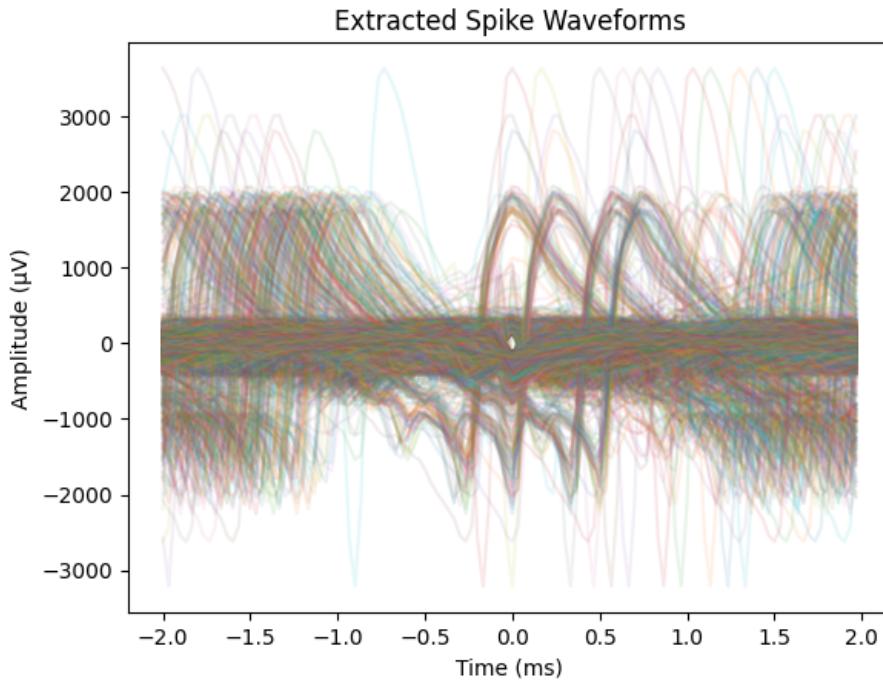


Figure 4: Extracted spike waveforms. Although the waveforms share a similar shape, amplitude variations suggest activity from multiple neurons.

1.4 Extracting Features

To reduce dimensionality and extract informative features, Principal Component Analysis (PCA) was applied to the spike waveform matrix. The first three principal components (PC1, PC2, and PC3) were selected for clustering.

1.5 Clustering the Spikes

K-means clustering was performed for values of k ranging from 1 to 10. The spike waveforms were projected onto the PCA components and plotted pairwise. The results for $k = 1$ to 5 are shown in Figure 5, and for $k = 6$ to 10 in Figure 6.

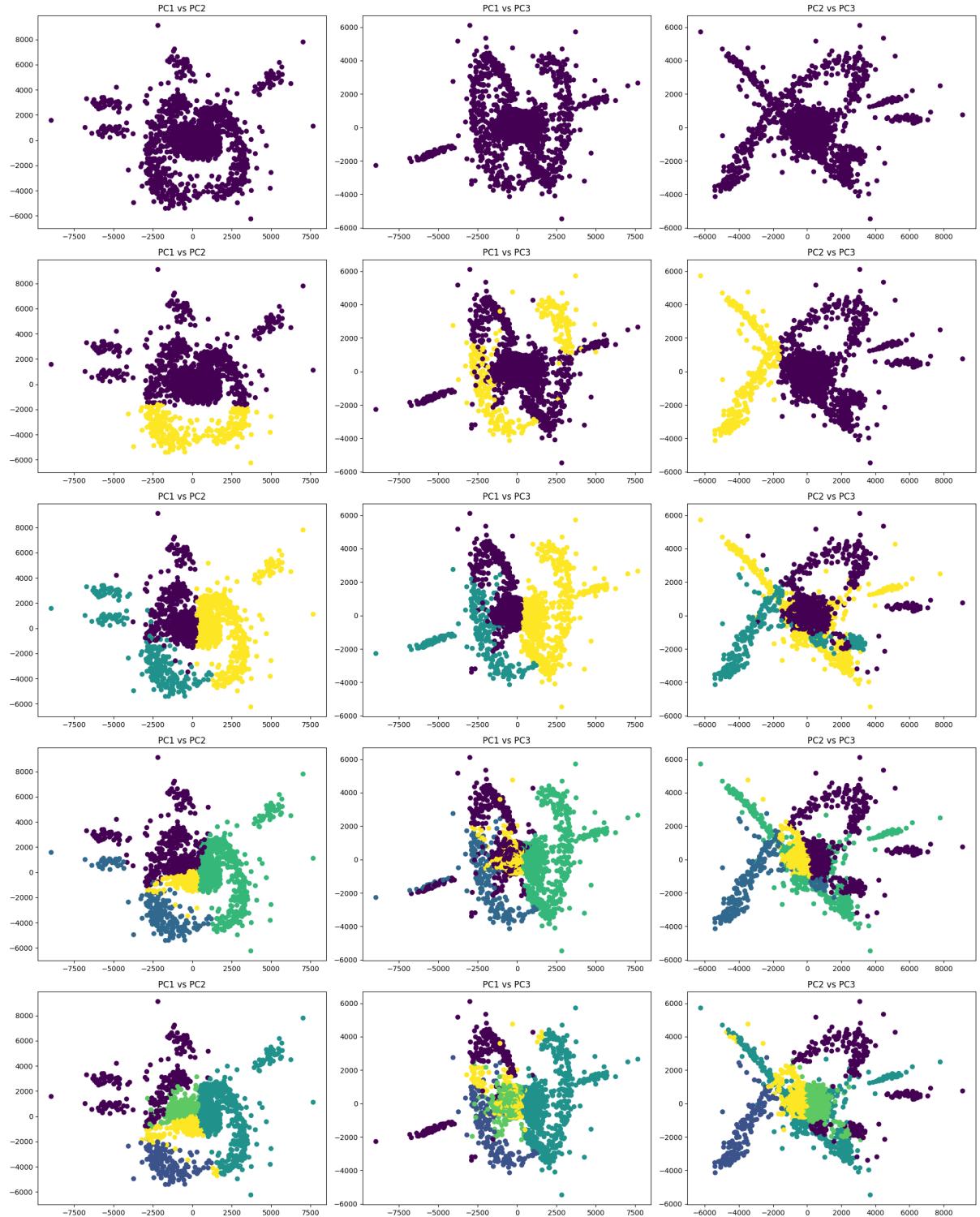


Figure 5: Pairwise PCA scatter plots for $k = 1$ to 5. Each point represents a spike, colored by cluster assignment.

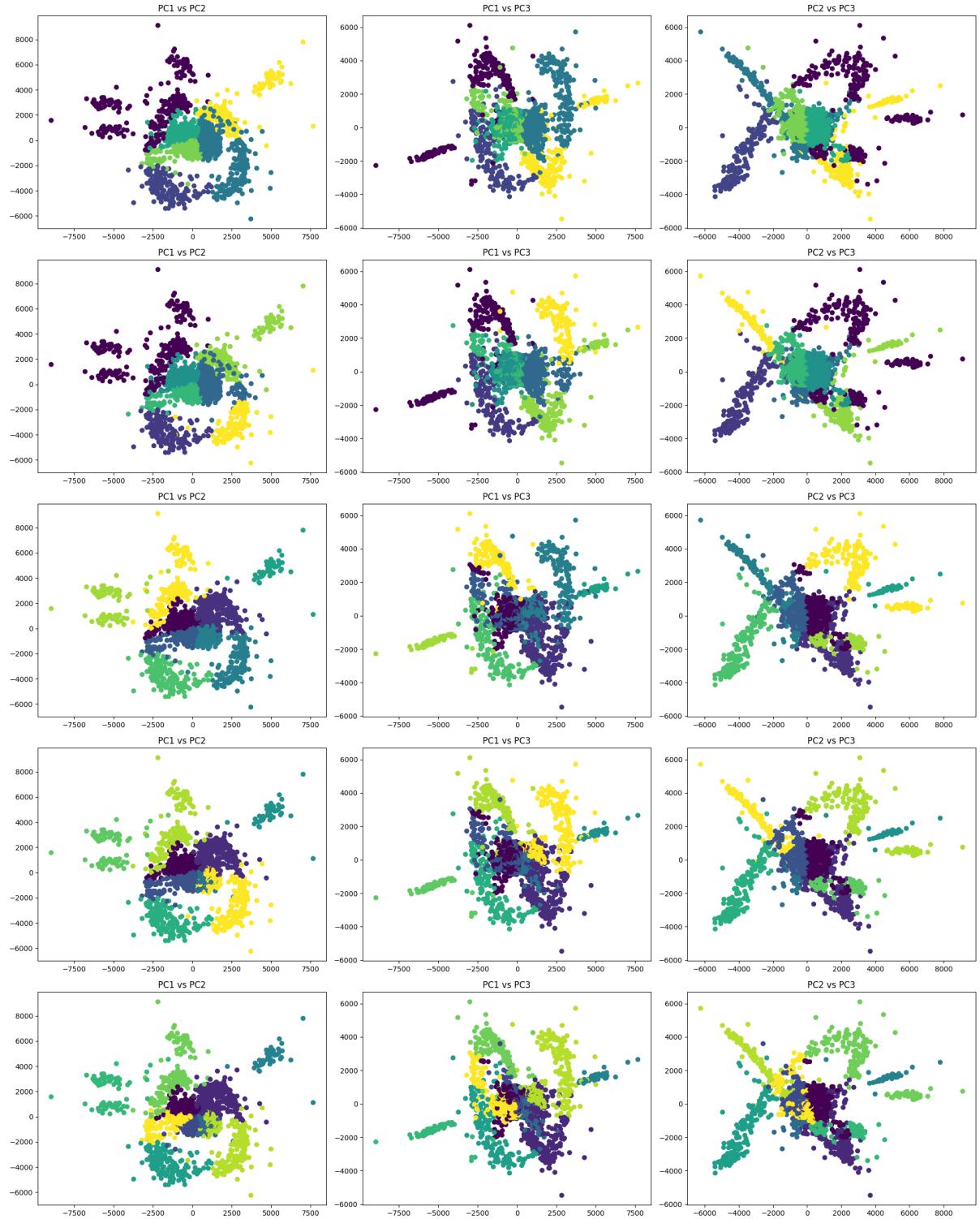


Figure 6: Pairwise PCA scatter plots for $k = 6$ to 10. Each point represents a spike, colored by cluster assignment.

To determine the optimal number of clusters, the elbow method was applied, suggesting that $k = 3$ is appropriate (Figure 7).

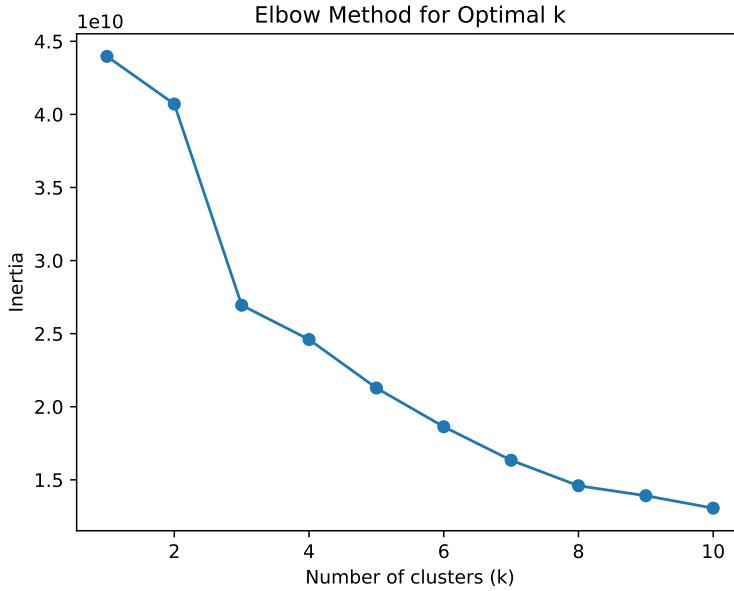


Figure 7: Elbow method to determine the optimal number of clusters. A clear elbow appears at $k = 3$.

1.6 Discussion and Further Analysis

1. To evaluate performance, I compared detected spikes with ground-truth events from `spikes.mat`. By finding the intersection of indices, I found 8 common spikes, confirming the temporal alignment given the same sampling rate of 30 kHz.
2. When using a new threshold defined as $\theta_{\text{new}} = 0.9 \times \max_t(X_t)$, only the highest amplitude spike was detected. This approach is ineffective since it ignores spikes with lower amplitudes, which often originate from more distant neurons or are affected by noise. Therefore, this thresholding method is not robust.
3. Using t-SNE for feature extraction instead of PCA, I repeated the clustering analysis for $k = 1$ to 10. Figures 8 and 9 show the clustering results in the t-SNE space.

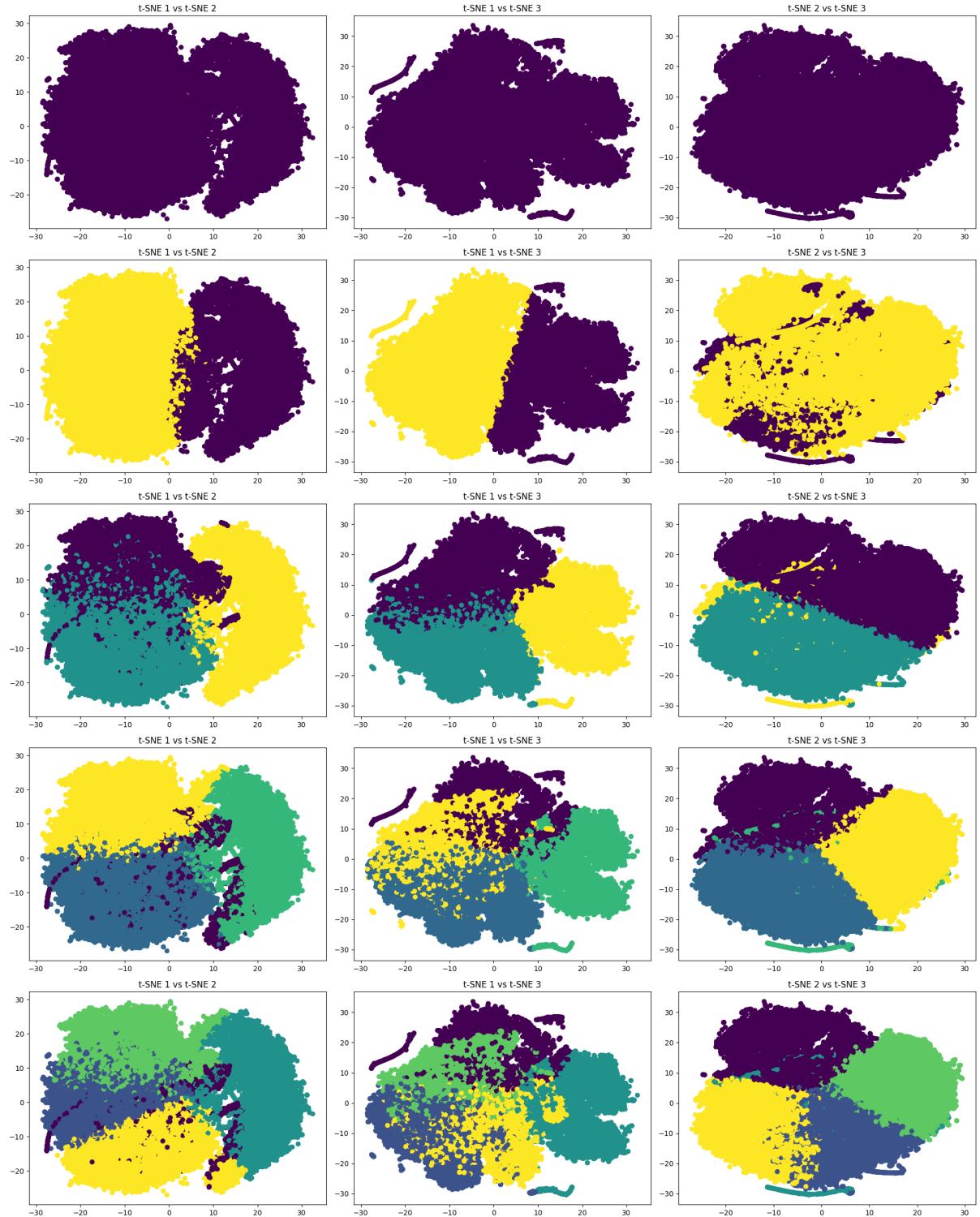


Figure 8: t-SNE projection of spike waveforms for $k = 1$ to 5. Clusters are visualized using different colors.

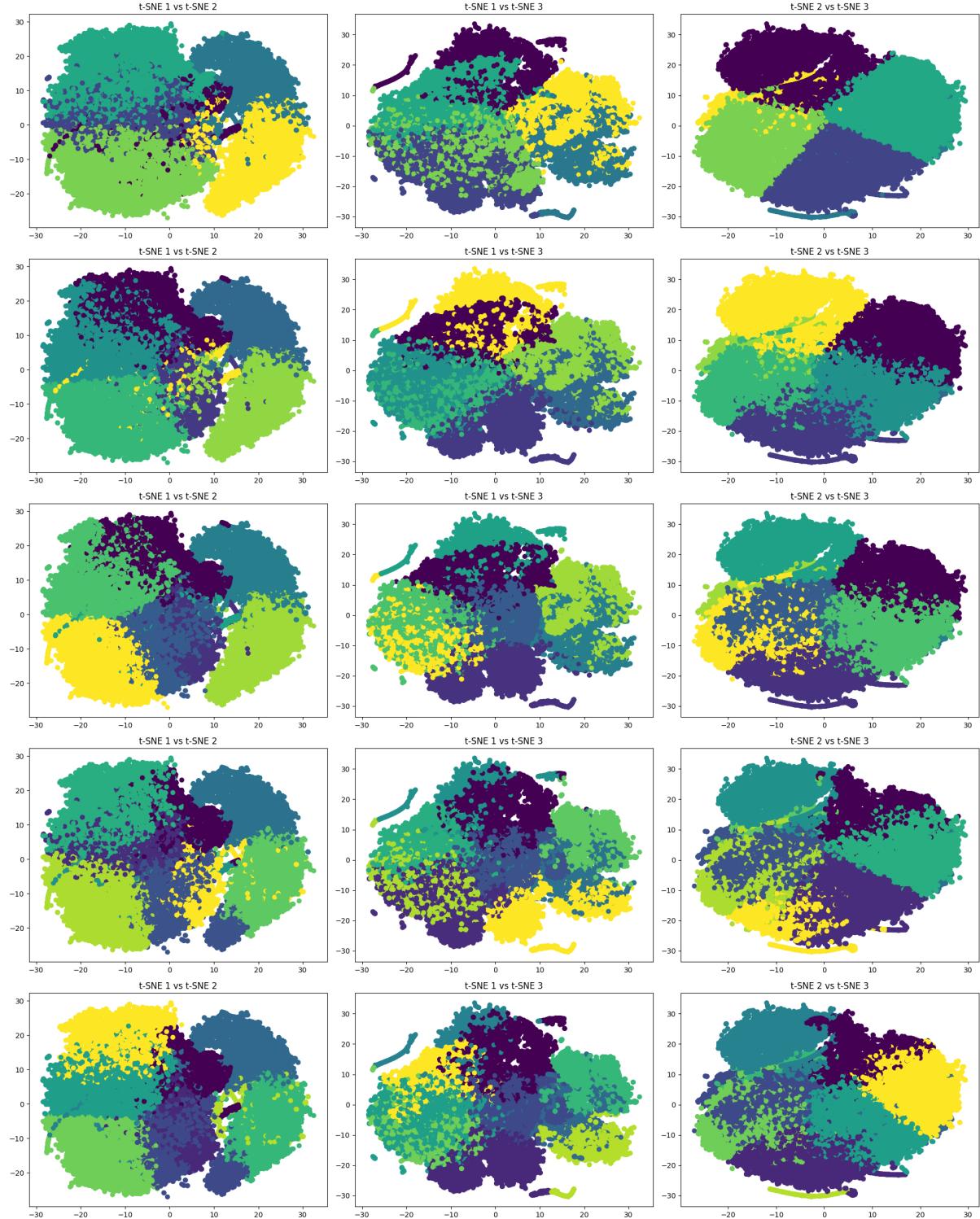


Figure 9: t-SNE projection of spike waveforms for $k = 6$ to 10. Clusters are visualized using different colors.

The elbow method on t-SNE features did not yield a clear optimal k , so I used the silhouette score, which indicated that $k = 4$ is optimal—slightly different from the PCA-based clustering result.

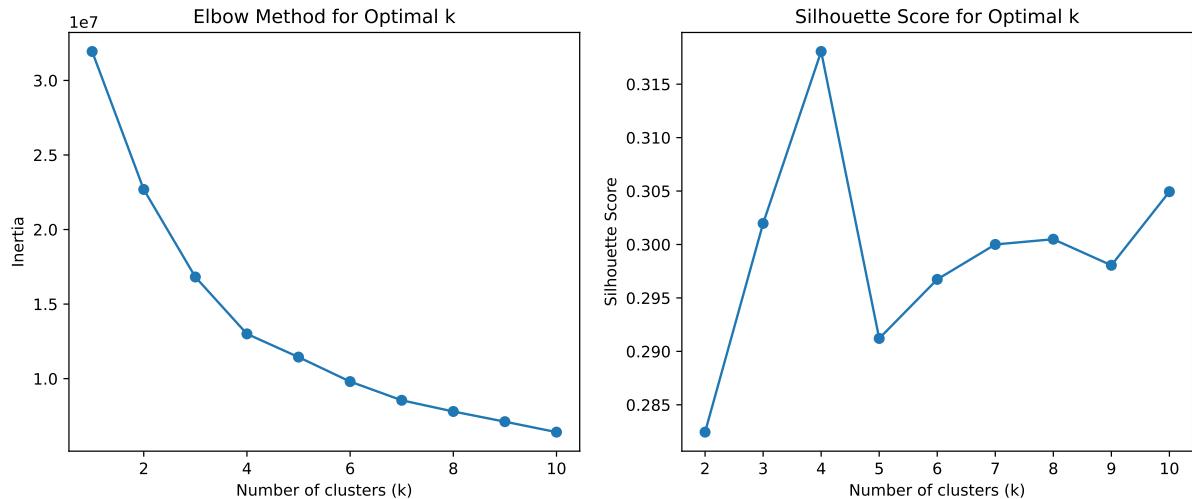


Figure 10: Comparison of elbow method and silhouette score on t-SNE features. The silhouette score peaks at $k = 4$, suggesting four distinct clusters.

Spike Sorting with ROSS

In this section, we utilize the ROSS (Robust Offline Spike Sorting) application to perform spike sorting through a semi-automated pipeline. The ROSS interface offers both automatic and manual tools for detection, clustering, and visualization of neuronal spikes.

2.1 Detection

The raw extracellular recording was loaded into ROSS. Spike detection was initiated after configuring appropriate parameters for filtering and thresholding. The results of the initial detection are shown in Figure 11.

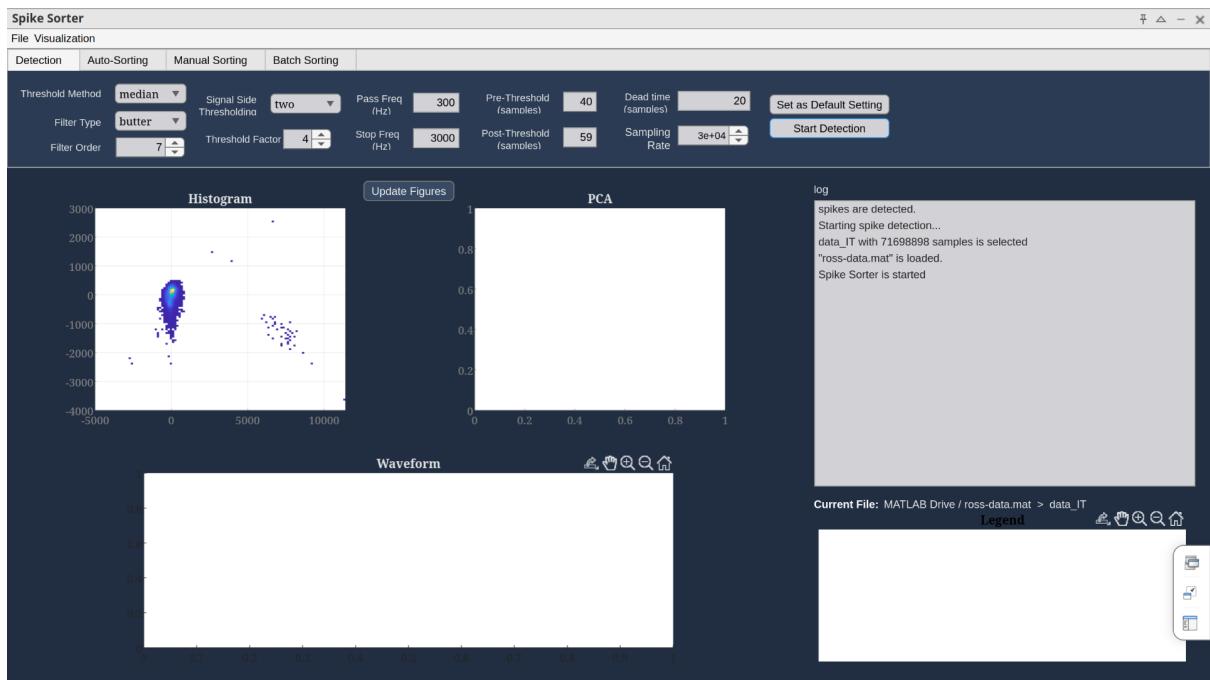


Figure 11: Spike detection in ROSS. Spikes are extracted from the filtered signal using an amplitude threshold based on estimated noise levels.

2.2 Auto Sorting

Following detection, the auto-sorting module was used. ROSS automatically grouped the detected spikes into three clusters. The initial clustering results are presented in Figure 12.

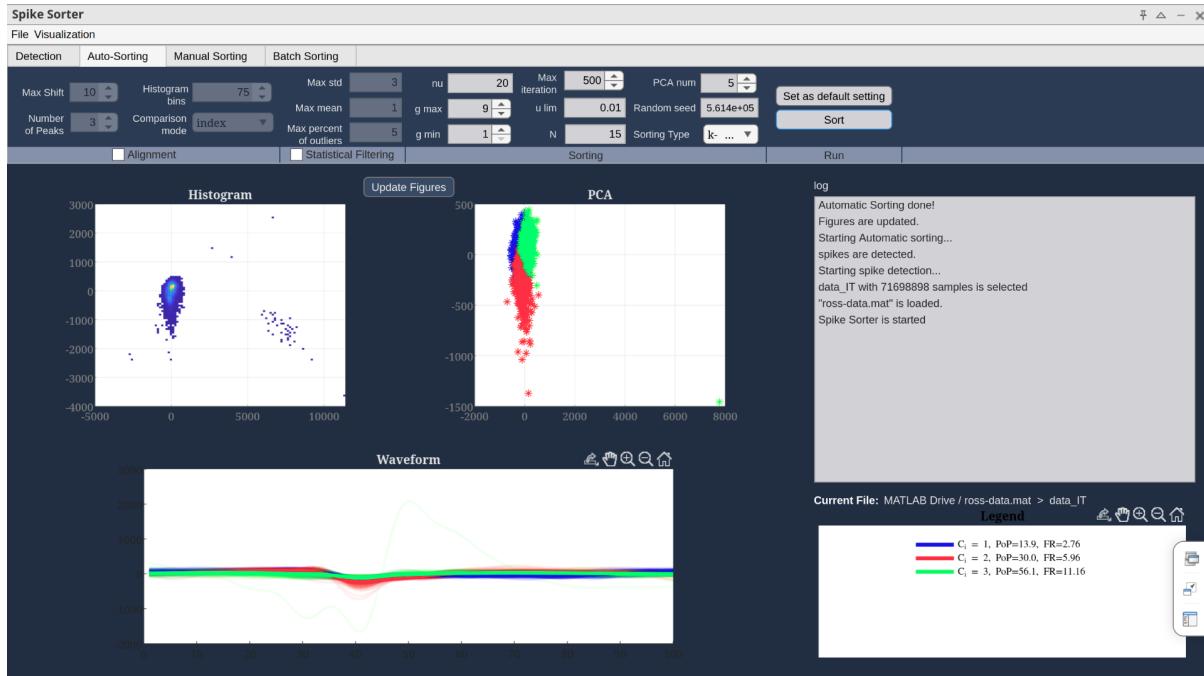


Figure 12: Automatic spike clustering using ROSS. Three initial clusters were identified based on waveform features.

2.3 Manual Sorting

1. **Denoising:** All three automatically identified clusters were manually denoised using a threshold of 85%. Each cluster was processed separately to remove noise and outliers.

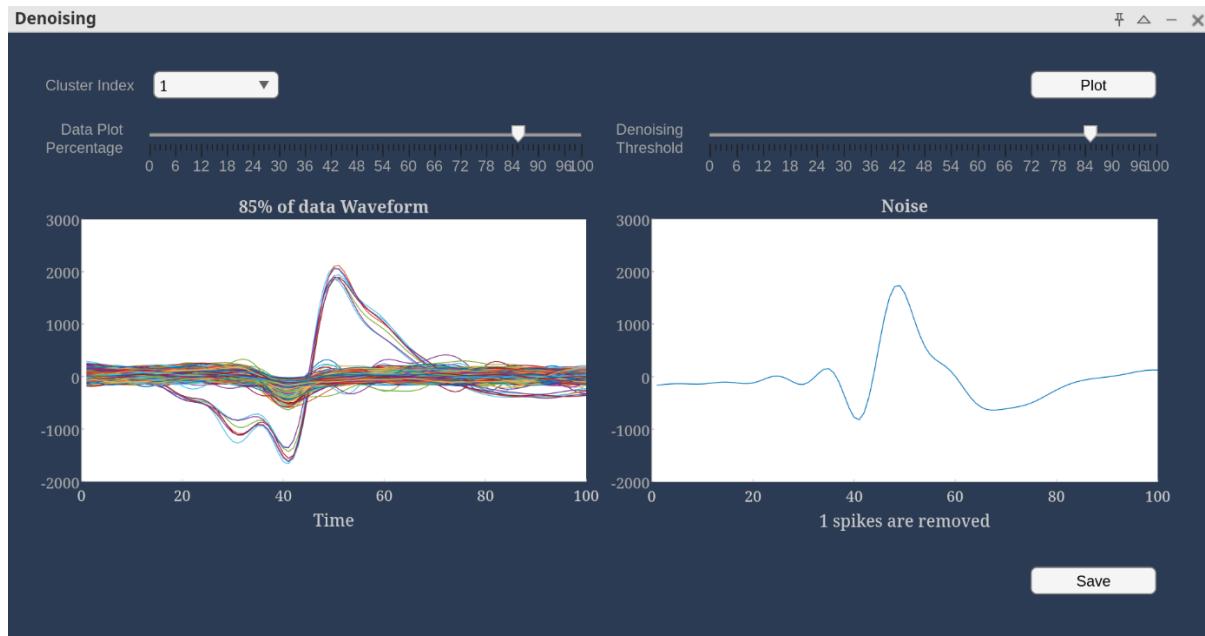


Figure 13: Denoising process for cluster 1 using 85% threshold.

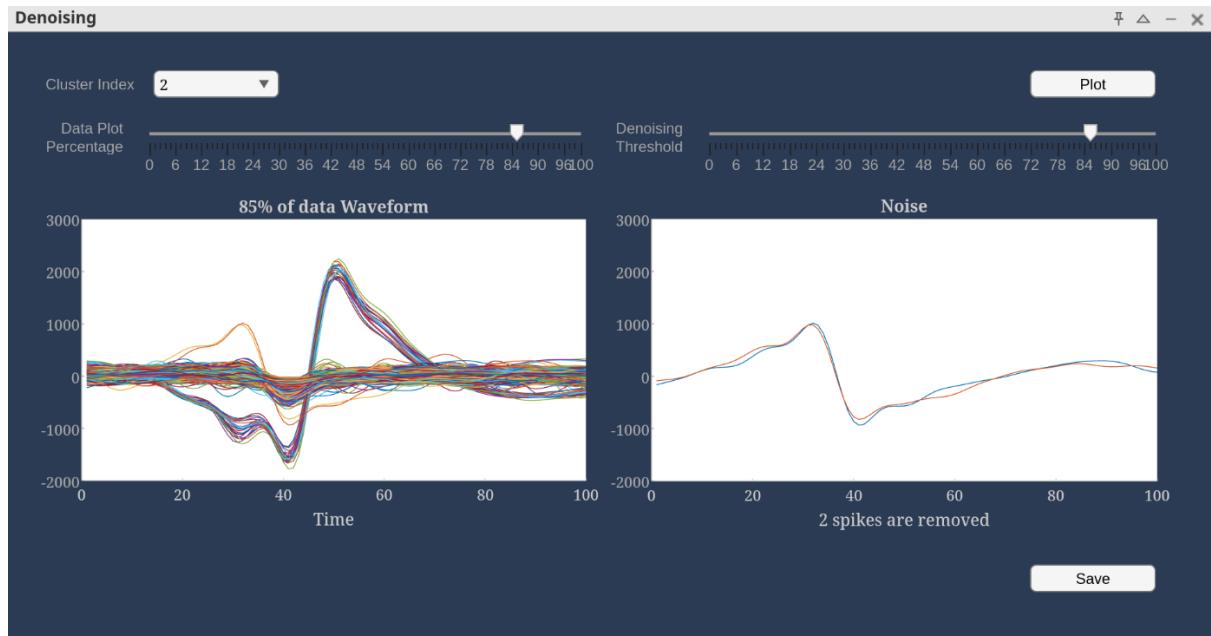


Figure 14: Denoising process for cluster 2 using 85% threshold.

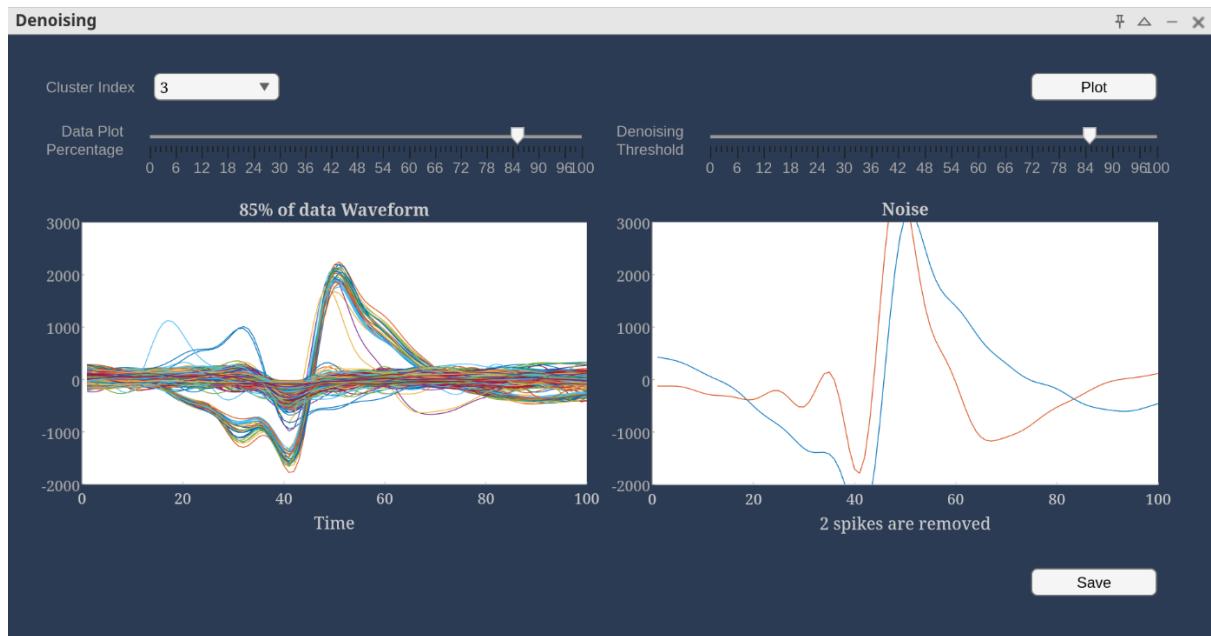


Figure 15: Denoising process for cluster 3 using 85% threshold.

The cleaned clusters after denoising are shown in Figure 16.

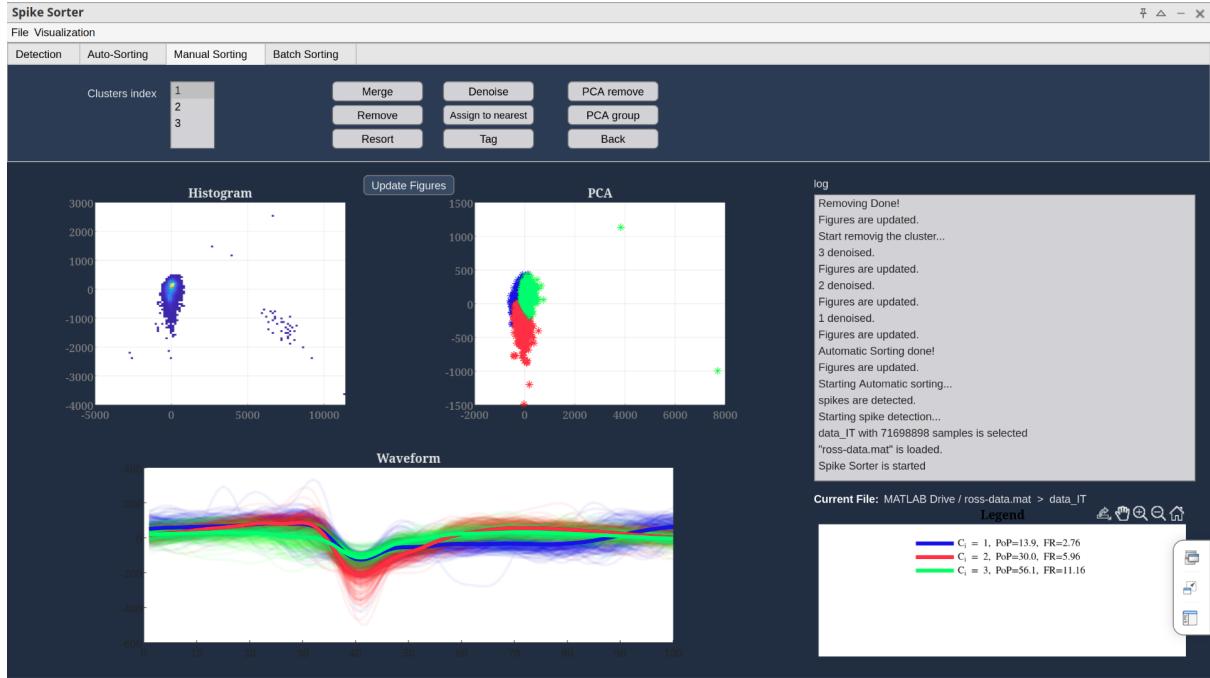


Figure 16: Final denoised clusters. Noise removal helped to better isolate spike waveforms from individual neurons.

2. Resorting: Cluster 3 had a high Percentage of Presence ($\text{PoP} = 56.1\%$), which is unusually high for a single neuron, suggesting that it might contain multiple neurons. Resorting this cluster resulted in its subdivision into three new clusters. After resorting, the total number of clusters increased to five (Figure 17).

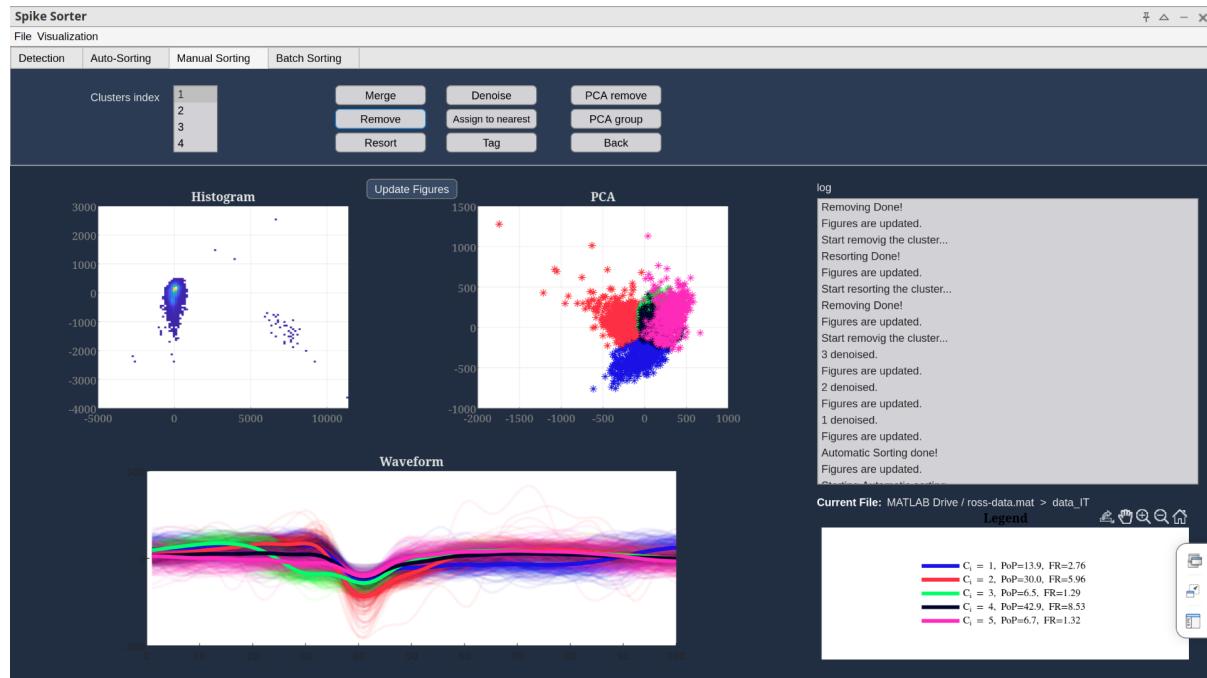


Figure 17: Result of resorting Cluster 3. It was divided into three sub-clusters, increasing the total number of clusters to five.

2.4 Visualization

1. **Waveform Visualization:** The spike waveforms for each of the five clusters were plotted. Cluster 4 (SP = 15281, PoP = 42.9%, FR = 8.53 Hz) had the largest number of spikes and the highest firing rate, indicating a highly active neuron (Figure 18).

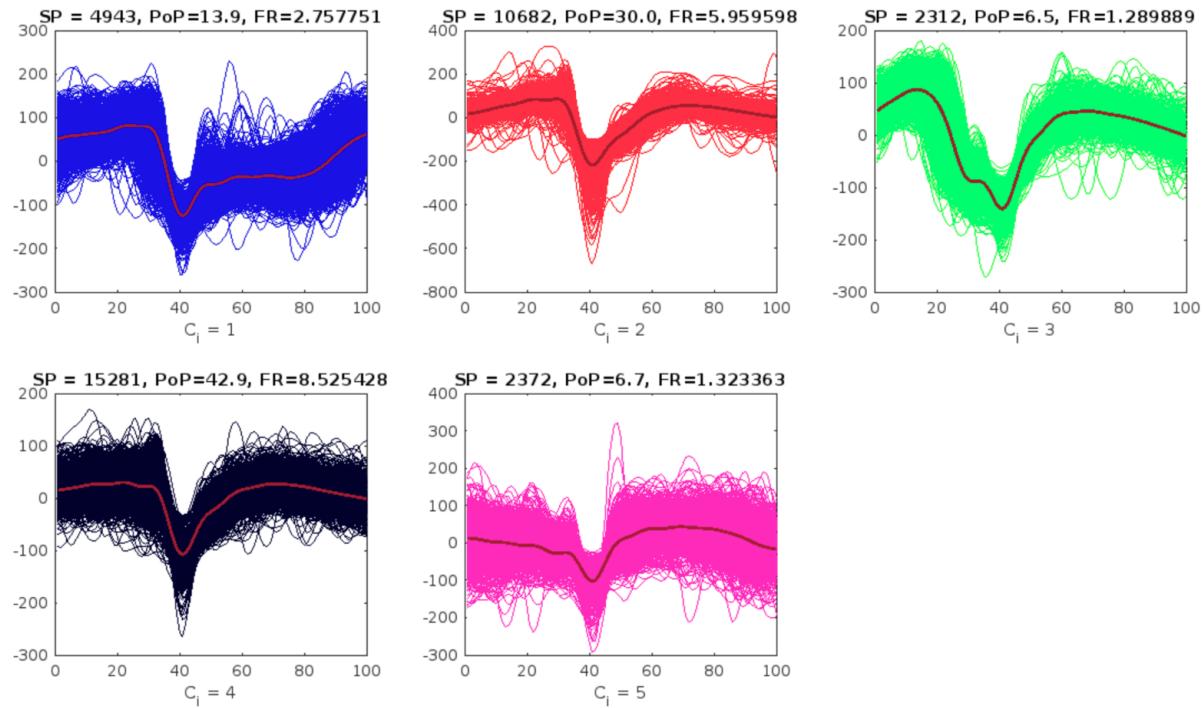


Figure 18: Waveforms of the five clusters after spike sorting. Cluster 4 appears dominant in both frequency and spike count.

2. 3D PCA Plot: A 3D PCA visualization of the first two components over time highlights the temporal structure and separation of the detected clusters (Figure 19).

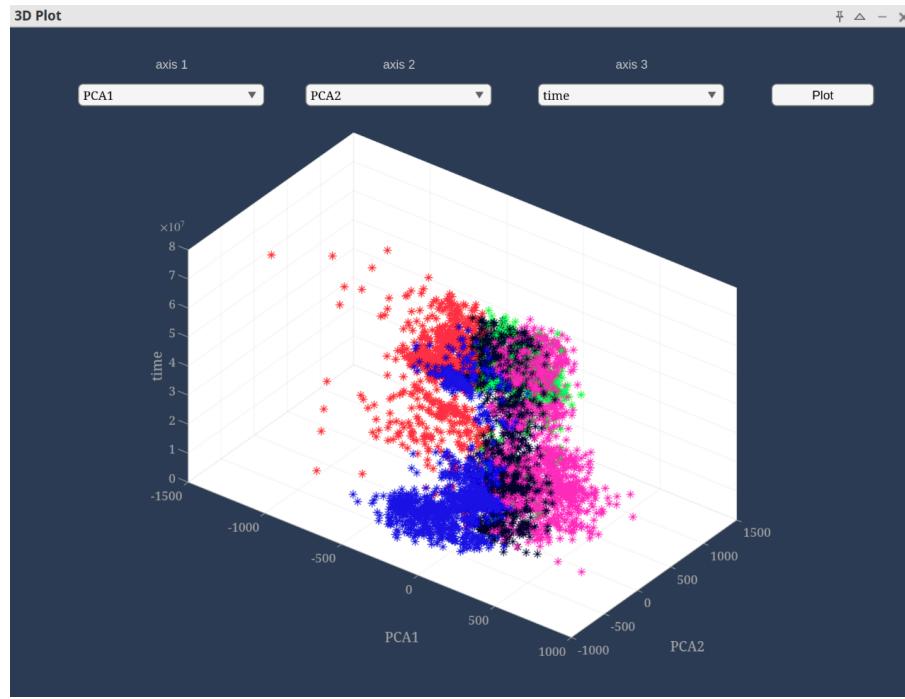


Figure 19: 3D PCA plot showing temporal dynamics of spike clusters. Colors represent different clusters.

3. Spike Localization on Raw Signal: The full raw signal was overlaid with markers indicating detected spikes. This allows visualization of spike occurrences relative to the ongoing signal dynamics (Figure 20).

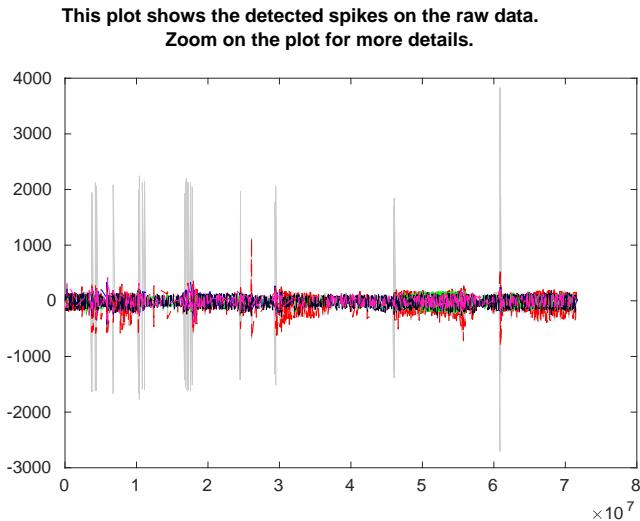


Figure 20: Raw signal with overlaid spike detections from all five clusters.

Zoomed-in segments of the raw signal highlight the waveform of individual spikes from each cluster.

This plot shows the detected spikes on the raw data.
Zoom on the plot for more details.

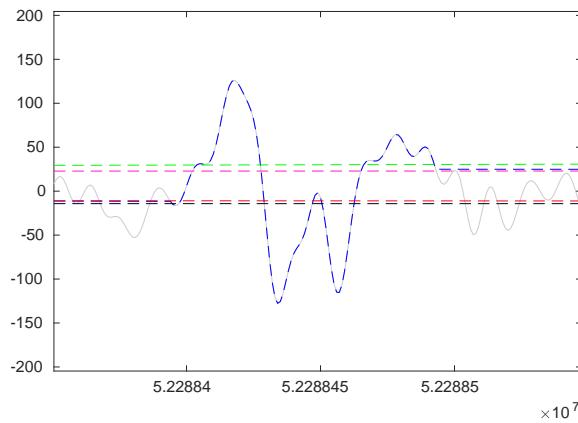


Figure 21: Example spike waveform from Cluster 1.

This plot shows the detected spikes on the raw data.
Zoom on the plot for more details.

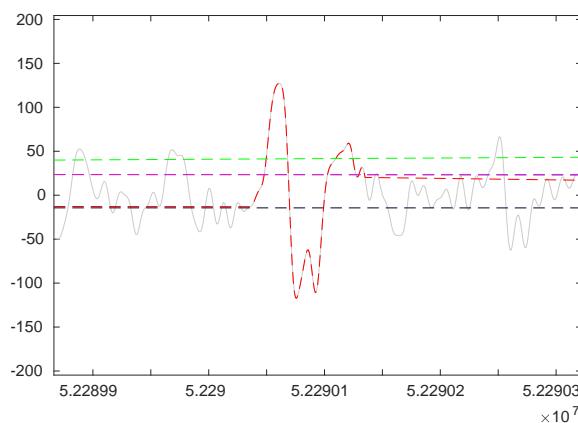


Figure 22: Example spike waveform from Cluster 2.

This plot shows the detected spikes on the raw data.
Zoom on the plot for more details.

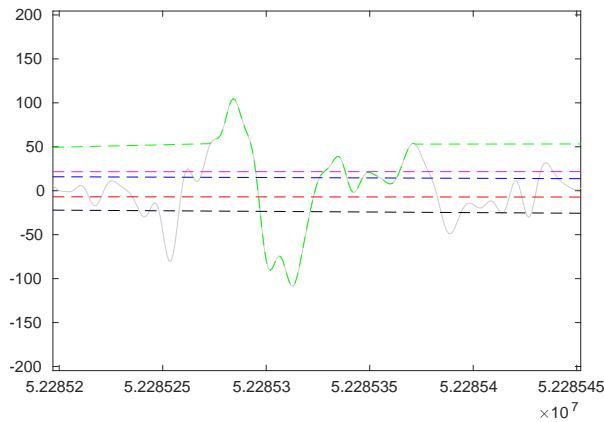


Figure 23: Example spike waveform from Cluster 3.

This plot shows the detected spikes on the raw data.
Zoom on the plot for more details.

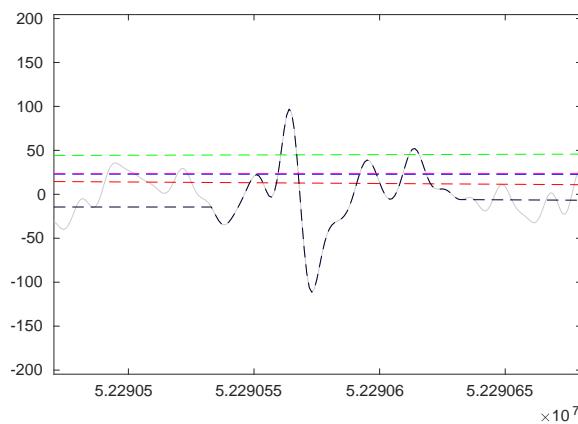


Figure 24: Example spike waveform from Cluster 4.

This plot shows the detected spikes on the raw data.
Zoom on the plot for more details.

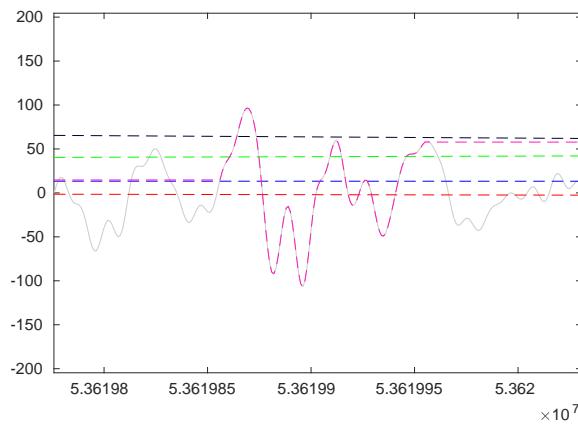


Figure 25: Example spike waveform from Cluster 5.

Analysis of Single Neuron Activity

3.1 PSTH Analysis

To investigate stimulus-specific firing behavior, I computed Peri-Stimulus Time Histograms (PSTHs) for different neurons across stimulus categories. PSTHs were generated by averaging spike counts across trials, aligned to stimulus onset. Figure 26 shows PSTHs for neurons 5 and 62.

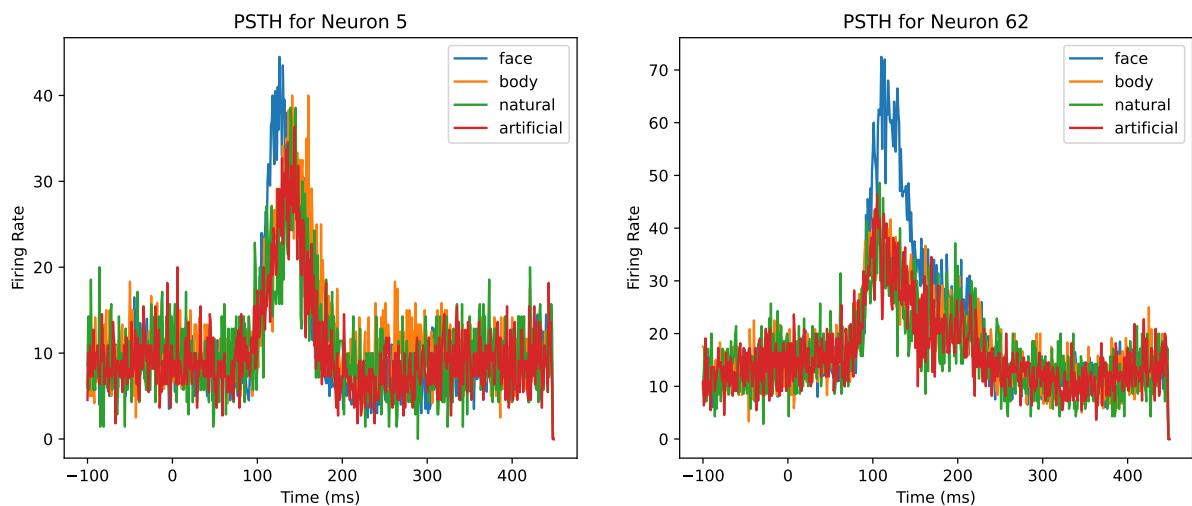


Figure 26: PSTH plots for neurons 5 and 62 across different stimulus categories. The horizontal axis represents time (ms), aligned to stimulus onset; the vertical axis shows the average firing rate (Hz).

Questions

- The neurons exhibit low baseline firing rates, but their activity increases approximately 100 ms after stimulus onset. This latency corresponds to visual stimulus presentation. The firing rates vary across stimulus categories, with both neurons showing signifi-

cantly elevated responses to face stimuli, suggesting face-selectivity.

3.2 Fano Factor Analysis

To quantify trial-to-trial variability in neural activity, I computed the Fano Factor for selected neurons and the Mean-Matched Fano Factor (MMFF).

3.2.1 Fano Factor

Spike counts were computed in sliding windows of 50 ms with a 5 ms step size. The Fano Factor was then calculated as the variance of spike counts divided by their mean for neurons 5, 31, 65, 68, 78, and 89. The results are shown in Figure 27.

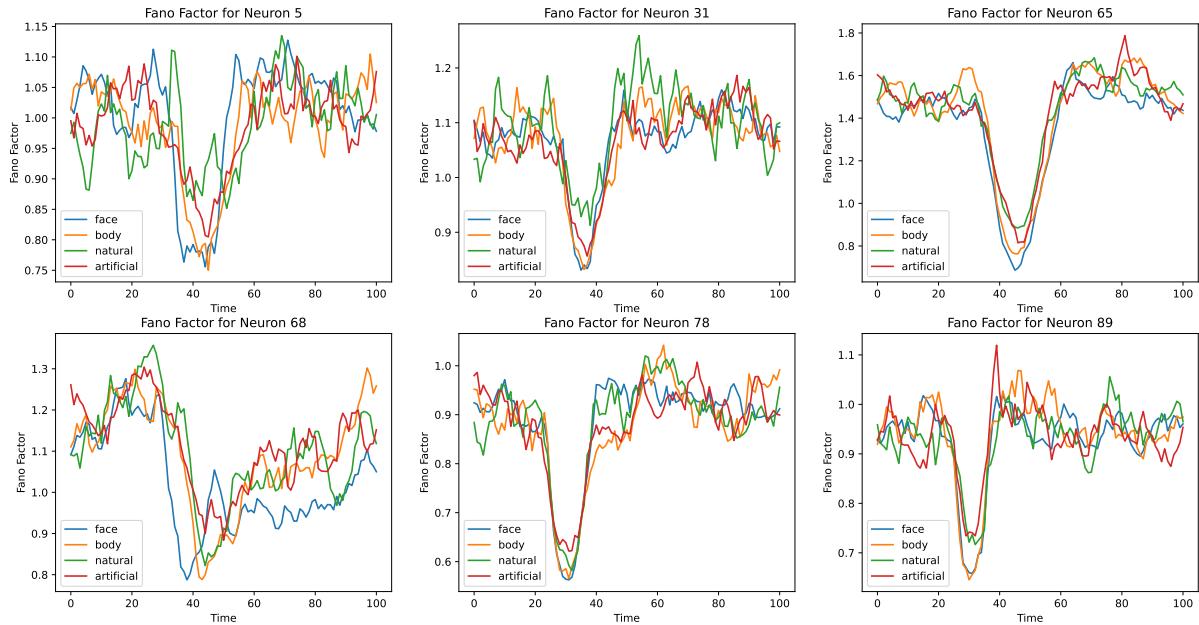


Figure 27: Fano Factor over time for selected neurons. Higher values indicate greater variability in spike counts across trials.

3.2.2 Mean-Matched Fano Factor

To control for differences in firing rate across categories, I computed the Mean-Matched Fano Factor (MMFF). This was achieved by plotting the variance of spike counts against their mean,

fitting a regression line, and using its slope as the MMFF. The MMFF was then calculated separately for each stimulus category, as shown in Figure 28.

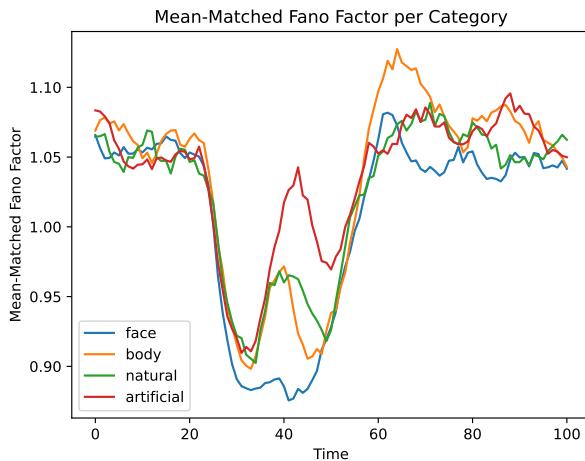


Figure 28: Mean-Matched Fano Factor for each stimulus category over time. The MMFF accounts for differences in firing rate, providing a normalized measure of variability.

Question

The MMFF analysis reveals that the neural response variability is lowest for the face category, followed by body and natural categories, with artificial stimuli exhibiting the highest variability. This suggests that neurons are more reliably and selectively tuned to face stimuli. The reduced variability implies consistent neural responses, reinforcing the role of these neurons in face detection and categorization.

3.3 Category Classification Using SVM

Each stimulus was assigned to one of four semantic categories: Face (0), Body (1), Natural (2), and Artificial (3). A Support Vector Machine (SVM) classifier was then trained using spike count features extracted from a fixed time window. The classification accuracy and recall for each category are shown in Figure 29.

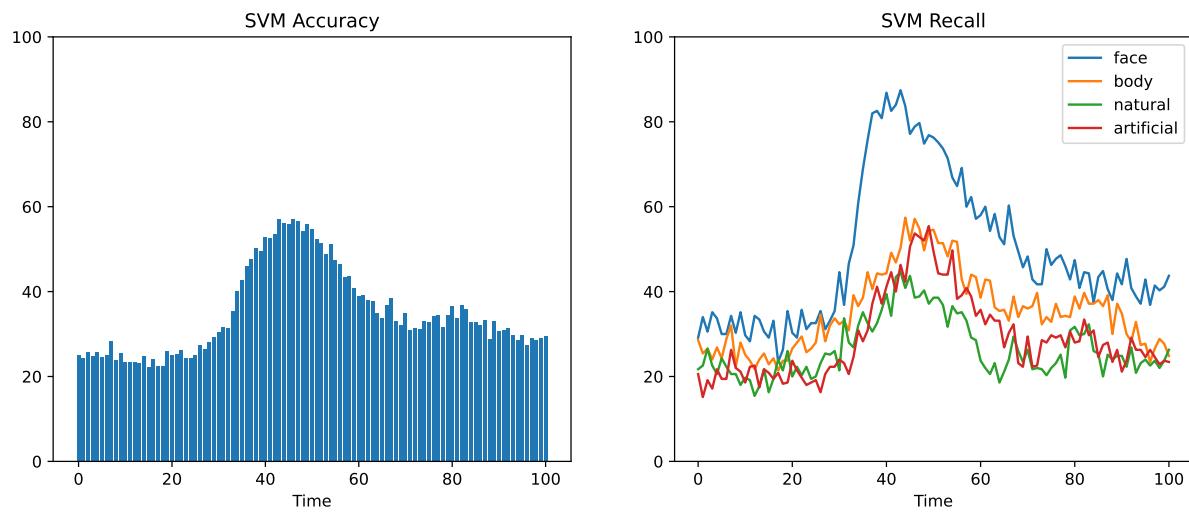


Figure 29: SVM classification results. The left plot shows the overall classification accuracy over time; the right plot displays recall values per category, reflecting how well each category was recognized.

Questions

- As shown in the figure, approximately 100 ms after stimulus onset, classification accuracy rises to nearly 60%. This is substantially higher than the chance level of 25%, indicating that stimulus identity can be predicted from neuronal responses with reasonable accuracy.
- The recall plot shows that face stimuli were most accurately recognized, followed by body, natural, and artificial categories. This suggests higher neural selectivity or discriminability for face-related visual inputs.

- These results support the idea that different visual categories elicit distinct patterns of neuronal firing. By learning these patterns, it is possible to decode perceptual content from population spike activity.

3.4 Time-Time Decoding Analysis

To assess the temporal dynamics of category-related information in neural activity, I performed a time-time decoding analysis. At each time point, an SVM classifier was trained and tested across all other time points, generating a 2D decoding matrix. To identify statistically significant regions, I used permutation testing to generate a null distribution, computed a p-value matrix, and applied False Discovery Rate (FDR) correction. The resulting time-time decoding matrix with significant regions is shown in Figure 30.

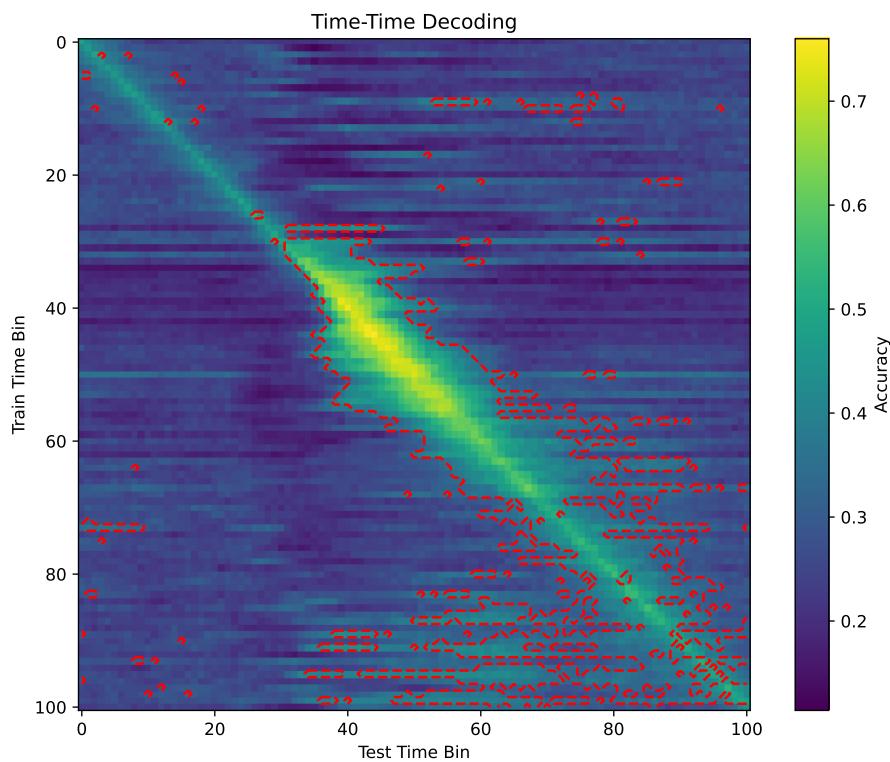


Figure 30: Time-time decoding matrix with FDR-corrected significance. Each entry reflects decoding accuracy when the model is trained at time t_1 and tested at time t_2 . Significant regions ($p < 0.05$) are highlighted.

Questions

- The decoding matrix shows that, beginning around 100 ms post-stimulus and lasting for about 200 ms, there is reliable information for decoding stimulus category. This window corresponds to peak neural responsiveness.
- Notably, several off-diagonal regions are statistically significant. This implies that information present at one time point generalizes to different time points, indicating sustained or recurrent processing, rather than purely feedforward dynamics.
- These findings suggest that neural representations of visual categories are not strictly transient but are maintained over time, supporting theories of dynamic and interactive cortical processing in the visual system.

3.5 Mutual Information Analysis Across Time

To quantify the amount of stimulus-related information present in neuronal activity over time, I computed the Mutual Information (MI) between the spike counts and stimulus categories at each time point. To assess statistical significance, a permutation test was conducted: the stimulus labels were randomly shuffled to create a null distribution, and the significance threshold was defined as the 95th percentile of the MI values across permutations at each time point. The result is shown in Figure 31.

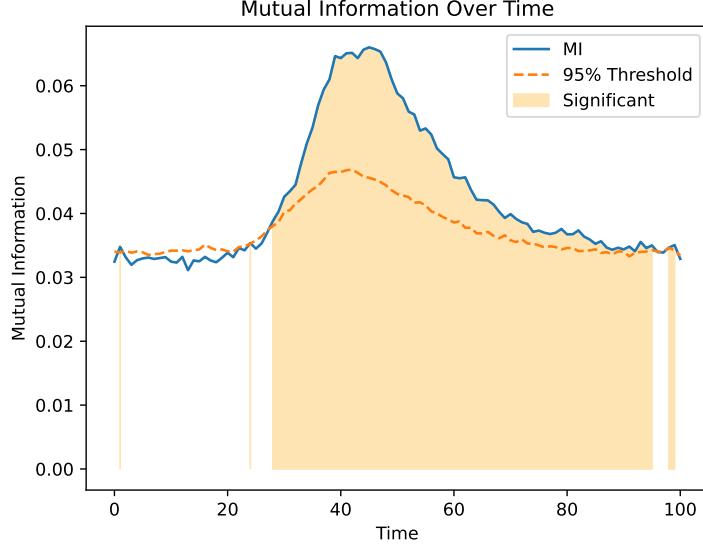


Figure 31: Mutual Information (MI) between neuronal activity and stimulus category as a function of time. The shaded region indicates the time period with statistically significant MI values (above the 95th percentile of the permutation-based null distribution).

Questions

- The MI curve peaks between 100 ms and 200 ms post-stimulus, indicating this is the period where neurons convey the most information about stimulus identity.
- This result is highly consistent with findings from the SVM and time-time decoding analyses, which also identified the 100–200 ms window as most informative.
- These findings suggest that during this period, the firing patterns of neurons in the inferior temporal cortex are highly informative and can be effectively used to decode

the category of the presented stimulus.

3.6 Category Discriminability Using d-prime

To further evaluate category selectivity at the single-neuron level, I computed pairwise d-prime (d') values for neuron 62. The d' metric quantifies how well the neuron distinguishes between pairs of stimulus categories based on its spike count distribution. The results are presented in Figure 32.

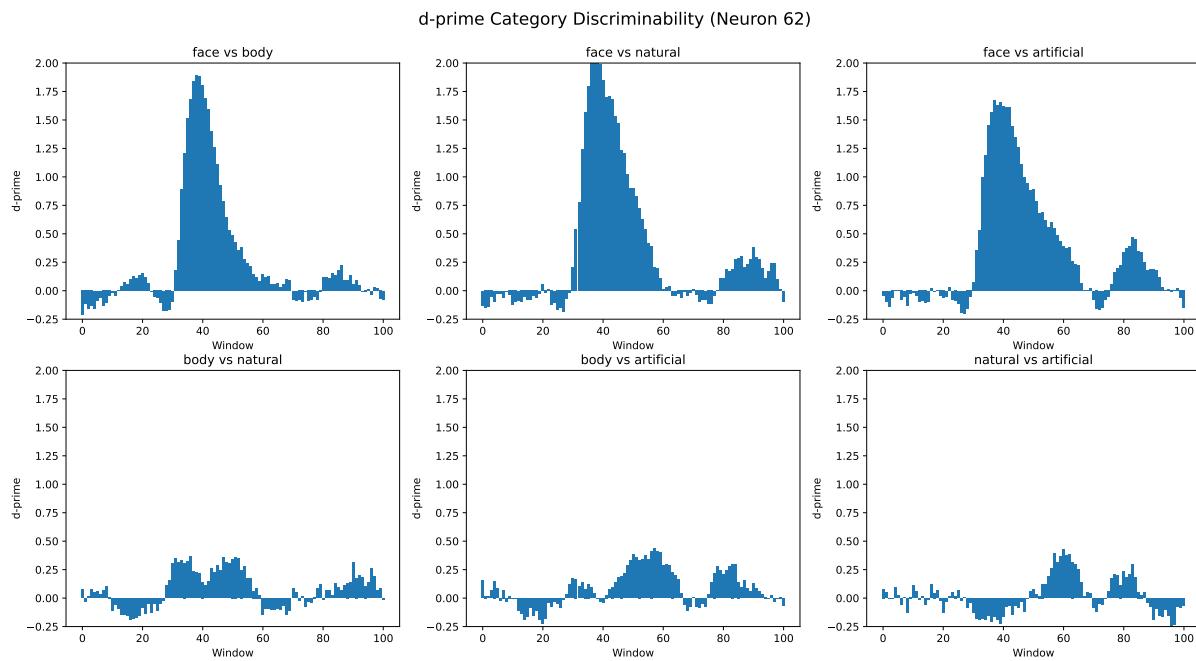


Figure 32: Pairwise d-prime values for neuron 62. Each bar indicates the discriminability between a pair of stimulus categories. Higher values reflect better separability.

Questions

- As illustrated, the neuron exhibits strong discriminability for the face category when compared to all others, while showing less distinction between the body, natural, and artificial categories.
- These results are fully consistent with the previous analyses (PSTH, SVM, MI), confirming that this neuron is highly selective for faces.

- The analysis indicates that face-related stimuli evoke unique and distinguishable responses in certain neurons, enabling binary decisions (e.g., face vs. non-face) based on neuronal activity.

Analysis of Population Activity

4.1 Representational Dissimilarity Matrix (RDM) and Kendall's Tau Correlation

To investigate the alignment between neural population responses and stimulus category structure, I computed the Representational Dissimilarity Matrix (RDM) across time using spike count vectors for each stimulus. I also constructed a ground truth RDM based on categorical similarity (same vs. different categories). Kendall's tau correlation was then calculated between the neural RDM and ground truth RDM at each time slice. The resulting time-resolved alignment is shown in Figure 33.

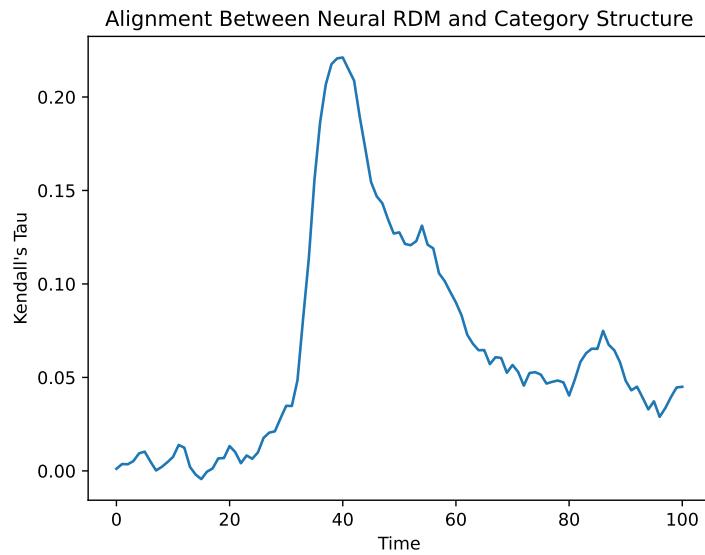


Figure 33: Kendall's tau correlation between the neural RDM and category-based ground truth RDM as a function of time. The peak alignment occurs around 100 ms post-stimulus.

Questions

- The highest alignment between neural and category RDMs occurs at approximately 100 ms after stimulus onset. Figure 34 displays the RDM at time slice 41, where same-category stimuli exhibit low dissimilarity (dark blocks), and different-category stimuli exhibit higher dissimilarity, forming visible category boundaries.

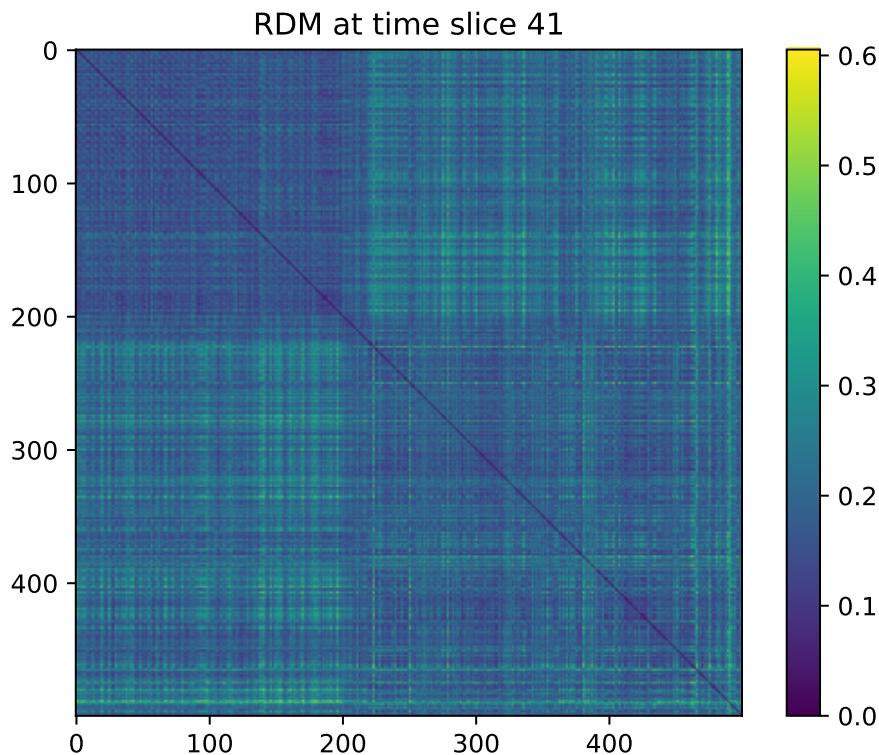


Figure 34: Neural Representational Dissimilarity Matrix (RDM) at the time of peak category alignment (time slice 41, ≈ 100 ms). Lower dissimilarities (darker colors) within categories and higher dissimilarities (lighter colors) across categories indicate structured population encoding.

- This analysis reinforces the conclusion that around 100 ms post-stimulus is a critical window for decoding categorical information from population neural activity.
- The structure of the RDM suggests that the neuronal population exhibits distinct and category-specific patterns of activity in response to different stimulus groups.

4.2 Generalized Linear Model (GLM) to Predict Ground Truth RDM

To evaluate whether the neural representations reflect higher-level categorical information beyond low-level visual features, I extracted visual features using a pretrained VGG16 convolutional neural network. Then, I trained a Generalized Linear Model (GLM) to predict the ground truth RDM using both the neural RDM and the visual feature RDM as predictors. The R^2 score across time is shown in Figure 35.

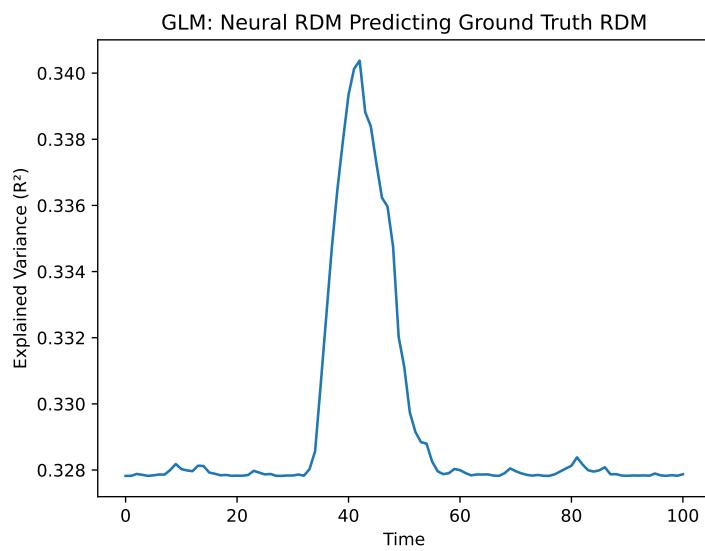


Figure 35: Variance explained (R^2 score) by a GLM trained on neural and visual feature RDMs to predict the ground truth RDM. The peak model performance occurs approximately 100–150 ms post-stimulus, highlighting the contribution of both neural and visual representations to category encoding.

Phase-Amplitude Coupling (PAC) and Spectrum Analysis

To explore frequency-based neural dynamics, I analyzed Phase-Amplitude Coupling (PAC) for each stimulus category using two methods: Modulation Index (MI) and Canolty's method. For each method, PAC was computed across combinations of low- and high-frequency bands for all four stimulus categories (Face, Body, Natural, Artificial).

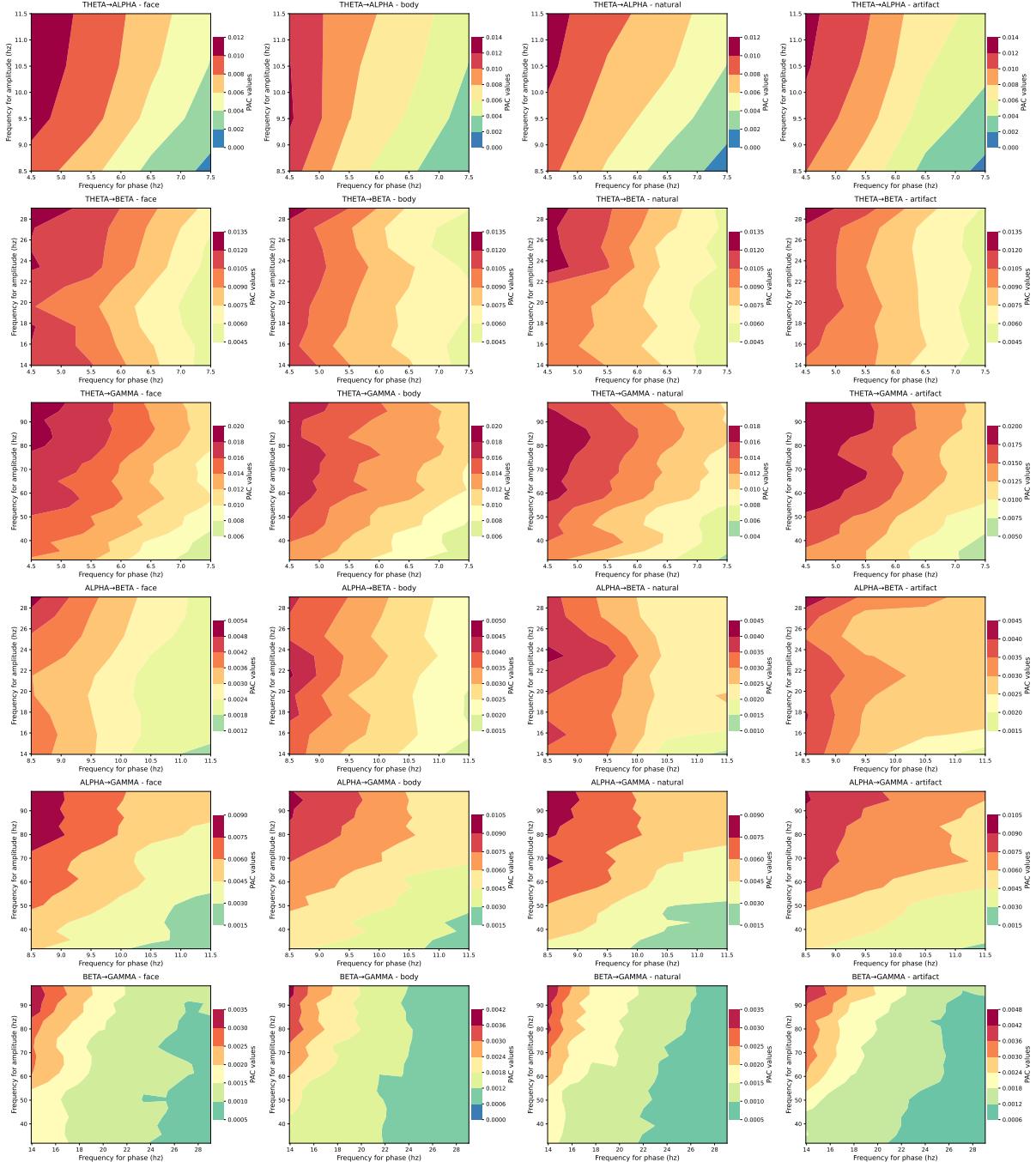


Figure 36: PAC plots using the Modulation Index (MI) method for each stimulus category. The strength of coupling between low-frequency phase and high-frequency amplitude is visualized across frequency pairs.

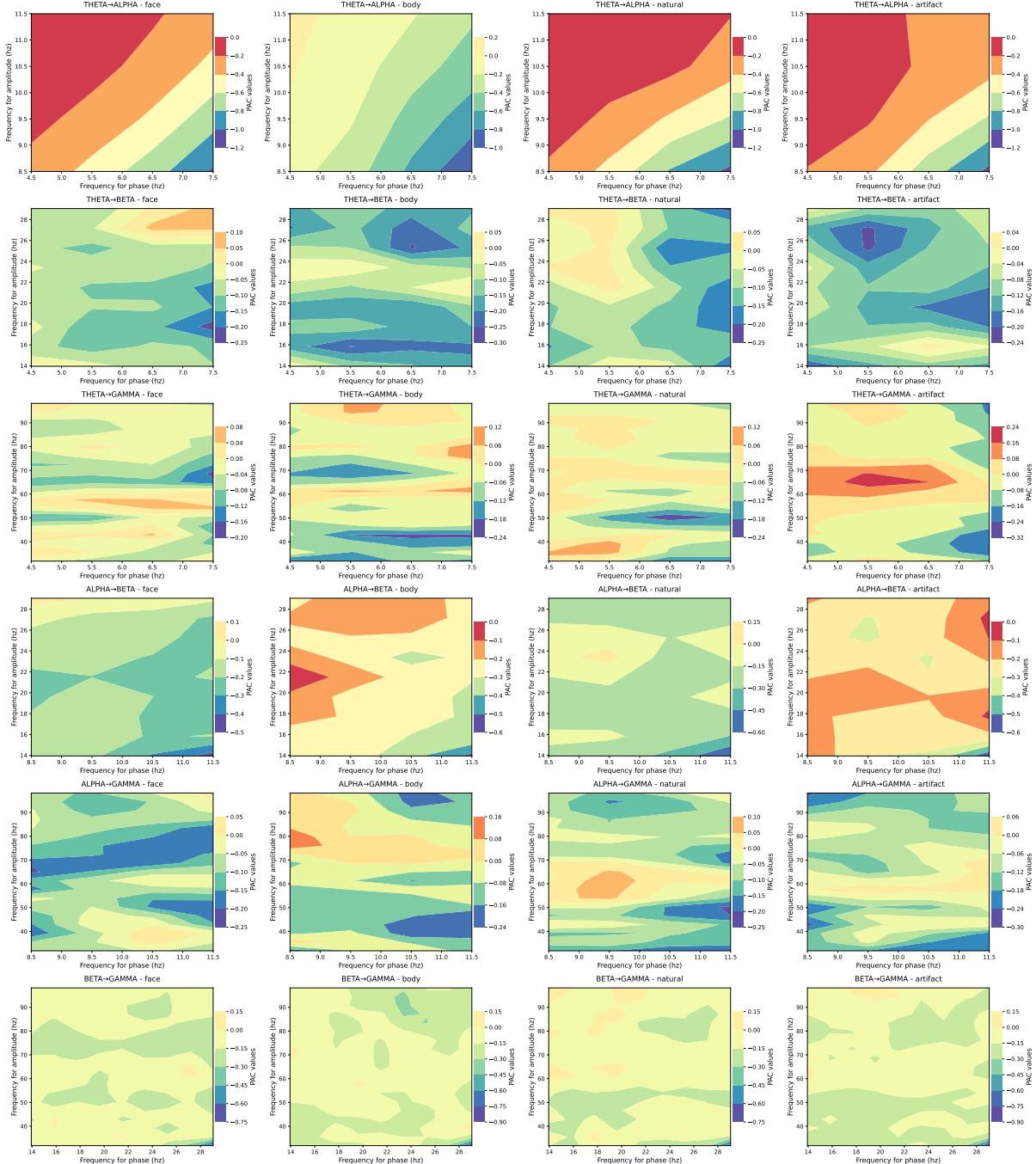


Figure 37: PAC plots using the Canolty method. Compared to MI, this approach shows subtle differences across categories, with "Body" exhibiting stronger coupling in specific frequency regions.

In addition to PAC, I computed the power spectrum for each stimulus category to analyze frequency-specific activation.

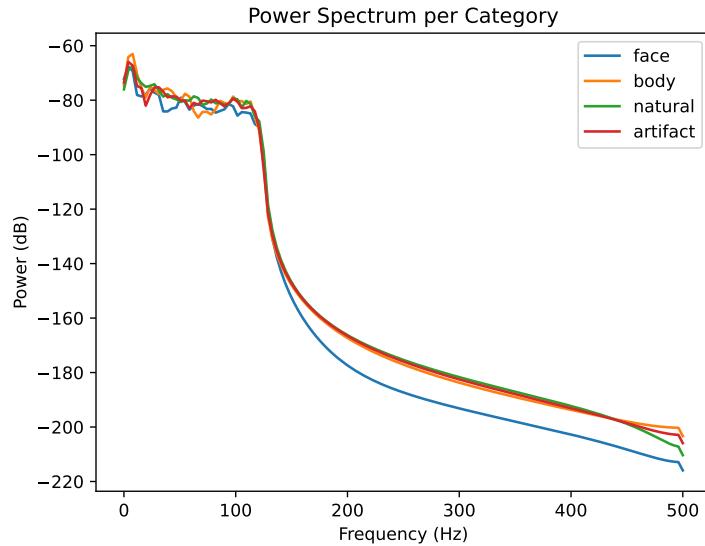


Figure 38: Power spectrum across frequency bands for each stimulus category. All categories show dominance in low-frequency ranges, but the "Face" category exhibits reduced power in the high-frequency range compared to others.

Questions

- For this particular neuron, there is no substantial difference in overall PAC strength across categories, suggesting uniform phase-amplitude interaction.
- However, using the Canolty method, the "Body" category exhibits stronger PAC compared to the others, indicating this method may be more sensitive to subtle differences in coupling dynamics.
- All categories show dominant activity in low-frequency bands. Notably, the "Face" category displays markedly lower power in the high-frequency range, whereas the "Body," "Natural," and "Artificial" categories have relatively similar high-frequency profiles. This indicates a unique spectral signature for face-related stimuli.
- Together, these findings suggest that while traditional PAC measures may not always reveal categorical distinctions, alternative coupling metrics (e.g., Canolty) and spectral

analysis can uncover nuanced differences in how different stimulus types are represented in neural activity.