Understanding Neuronal Diversity: Role of Input Dynamics and Filtering

Nishant Joshi[1], Sven van der Burg[2], Tansu Celikel[3,6,**], and Fleur Zeldenrust[1,4,6,8,*]

[1]Donders Institute for Brain, Cognition and Behaviour, Radboud University, The Netherland
[2]Netherlands eScience Center, Amsterdam, The Netherlands
[3]Department of Psychology, Georgia Institute of Technology, The United States of America
[4]Present address: Heyendaalseweg 135, postvak 33, kamer 00.111 6525 AJ Nijmegen, the Netherlands
[5]These authors contributed equally
[6]Senior author
[8]Lead contact
[*]Correspondence: f.zeldenrust@neurophysiology.nl
[**]Correspondence: celikel@gatech.edu

SUMMARY

Neurons exhibit diverse morphological, electrophysiological, and transcriptomic properties. Their role as spatiotemporal filters suggests that the interaction between input dynamics and postsynaptic neuronal properties should be crucial for shaping this diversity. To test this, barrel cortex neurons were studied using two input types: a ”step-and-hold” stimulus and a dynamic input mimicking presynaptic activity. It was found that stimulus type strongly influenced neuronal classification. Neuronal heterogeneity under a dynamic stimulus was further analyzed using four attribute sets: action potential attributes, passive biophysical attributes, adaptation current, and linear input filters via the Spike Triggered Average (STA). Multi-set factor analysis revealed that the linear input filter explained the most private variance, making it the most informative attribute for neuronal heterogeneity. These findings highlight the interplay between input dynamics and neuron’s linear filtering properties as key to understanding functional neuronal diversity, emphasizing the need to go beyond static intrinsic features for understanding neuronal heterogeneity.

KEYWORDS

Frozen Noise, Clustering, Heterogeneity, Input dynamics, Variance Explained, Input Filters

INTRODUCTION

Neural circuits are made up of neurons with varying degrees of specificity to presynaptic input patterns. These neurons are organized into circuits, giving rise to sensory processing, cognition, and behavior ([1]). The neural circuits that perform these complex computations consist of neurons with heterogeneous properties, whether viewed from morphological, transcriptomic or electrophysiological properties. Grouping neurons with a similar set of functional attributes in classes can simplify the assessment of this heterogeneity for further exploration. However, a clear consensus on a set of functional attributes that captures functional heterogeneity is still lacking. This lack of a clear consensus comes from the fact that neuronal classification is a difficult problem ([2–9]). Two key issues in neuronal classification have been highlighted by[8]. The first problem is of under-sampling; single-cell electrophysiology is laborious and is prone to errors, and collecting morphological and transcriptomic labels slows down the process even more and causes scalability issues. Problems with imprecise targeting and lack of ground truth further

exacerbate the issue[10]. The second problem is deciding on the features that are conceptually useful for distinguishing neurons and grouping them into clusters. All neurons are essentially different from each other in terms of morphology, molecular makeup, and function, which makes each neuron a class of its own. However, this level of separation is not useful for understanding functional differences within neuronal populations. On the other hand, neurons can be grouped into broad classes, such as excitatory or inhibitory based on the neurotransmitters they release ([1]). Although these broad categories appear to have different activity profiles and functional properties, this broad classification is not sufficient to explain the network computation properties ([11,12]). There exists a wide range of classifications from coarse to fine, from molecular to physiological, a clear consensus on how to divide neurons into cell types, and how to measure neuronal diversity, is still lacking. Given the fact that neuronal properties mentioned above are high dimensional in nature, the problem of deciding upon the features that are the most informative about heterogeneity becomes more complex. Here, we we add an additional dimension to the problem of classification, that of input dependence, and assess the electrophysiological classification of neurons as a function of the input.

Recent advances in data collection techniques such as patch sequencing ([13–15]) have made a significant improvement in the problems of neuronal classification possible, as it allows for the simultaneous extraction of transcriptomic, morphological, and physiological properties of neurons. These datasets provide a multi-modal perspective on cell diversity and promise a high-fidelity classification of neurons. The Allen Brain Institute dataset ([16–18]) has presented the most generalized attempt so far: an unsupervised classification of neurons from the visual cortex and motor cortex based on electrophysiological, morphological, and transcriptomic features (MET types). However, even with these high-dimensional datasets, classification based on electrophysiological properties was still shown to be problematic, due to several reasons. Firstly, the electrophysiological properties of neurons presented in these studies were shown to lie on a continuum ([3]) and were found to have unclear boundaries between different classes. Secondly, it was shown that individual attribute sets are weak predictors of each other. For example, Cells in types defined on the basis of molecular fingerprints were found to have overlapping electrophysiological properties ([16,18]. Thirdly, these studies hypothesize that intrinsic electrophysiological differences can vary depending on the stimulation protocol ([16,18]). Although multi-modal patch-seq-like techniques are informative about neuronal functional classification, an important aspect the molecular and morphological markers fail to consider is that under natural conditions, a given neuron receives input from tens of thousands of other neurons in a circuit. Its firing properties and therefore its functional identity are shaped dynamically by the input that it receives from its presynaptic neurons. Therefore, it is important to understand the effect of input dynamics on functional properties and ultimately on functional classification. Hence, we need to consider experimental conditions where neurons receive physiologically realistic input ([19,20]).

In an in-vivo situation, neurons receive high-dimensional time-varying signals from presynaptic neurons. As described above, neuronal dynamics interact with the dynamic properties of the input a neuron receives, thereby influencing the input features to which a neuron is responsive. To get a true picture of a neuron's functional identity, examining its response to an in-vivo like input is important as it gives a complete picture of the intrinsic functional properties under naturalistic conditions. This is currently missing from MET-based classification studies. Experimentally tracing the synaptic input that a given neuron receives requires observing and modulating the time-varying activity of every pre-synaptic neuron to which it connects. This could uncover the full extent of a given neuron's response properties. However, so far such experiments are unfeasible. Without such an experiment, an alternative is to computationally simulate the confluence of synaptic input that a neuron receives. So when aiming to assess the functional heterogeneity of a neuronal population, it is important to probe neurons with a physiologically realistic input.

In this study, we aim to uncover the physiological diversity of layer 2/3 barrel cortical neurons

when they receive a physiologically realistic dynamic input, and assess its dependence on the input dynamics, i.e. in the stimulation protocol. A recently proposed experimental setup provides an alternative to the traditionally used static step-and-hold (SH) protocol: the frozen noise (FN) protocol ([21]), which was designed to estimate the mutual information between a neuron's input and output (i.e. spike train) in an in-vitro setting. This method simulates the activity of the pre-synaptic neurons using an artificial neural network, as it responds to a preferred stimulus (i.e. hidden state) that randomly appears and disappears. The presynaptic responses are linearly combined to produce the synaptic input that the post-synaptic neuron receives. Therefore, the Frozen Noise patch clamp (FN) protocol provides a basis for understanding the functional heterogeneity of a neuron population under realistic stimuli.

As explained above, the neuronal classification literature so far focused on individual MET type identity of neurons, but did not focus on the importance of input dynamics in revealing functional heterogeneity. In this work, we look at stimulus-dependent classification across neuron populations in the somatosensory cortex, with the help of single-unit patch-clamp recordings[22,23]). We assess the cell classification of the same cells stimulated by a static SH and a physiologically realistic FN input. We contrast high dimensional features extracted from neurons when injected with this static or the dynamic FN input, to observe the contribution of the input type in explaining functional classification, we showed that functional heterogeneity of the neuronal population depends strongly on the input it receives. To further explore what a physiologically realistic stimulus reveals about the functional heterogeneity of a neuronal population, we extract 4 attribute sets that are commonly studied: 1) action potential attributes, 2) passive biophysical attributes, 3) adaptation current, and 4) linear input filter. We study the relative contributions of these 4 attribute sets in explaining the neural heterogeneity across (putative) inhibitory and excitatory neuron populations in the form of shared and private variance explained by each attribute. We find that the linear input filter (attribute 4) is the most informative about neural heterogeneity across the excitatory and inhibitory populations. This work challenges the paradigm of studying neuronal functional heterogeneity without considering their characteristic input. This work will nudge single-cell electrophysiologists and neuron taxonomists to not overlook input dynamics when defining neuron types. It will hopefully open new avenues for modelers and experimentalists alike when designing experiments for studying neuronal dynamics and functional heterogeneity across populations.

RESULTS

We aimed to observe the contribution of the input type in explaining functional classification and to explore what a physiologically realistic stimulus reveals about the functional classification of a neuron. For the first aim, we researched the stimulus dependence of neural classification using two different sets of classification features: 1) action potential waveforms and 2) other action potential attributes, such as the spike threshold and spiking dynamics. Next, to research the second aim, we investigated which attributes are the most informative about neuronal heterogeneity under a physiologically realistic stimulation protocol. Therefore, We performed classifications based on four different attribute sets that capture 1) action potential attributes, 2) Passive biophysical attributes, 3) Adaptation current, and 4) linear input filter through a Spike Triggered Average (STA). Finally, we used a method known as Multi-set correlation and Factor Analysis (MCFA) to compare the variance explained by the shared structure across these four attribute sets.

0.1 Stimulus dependence of neural classification

We first aim to understand the role of a neuron's input in the functional classification of neural populations. For this, we analyze single-cell patch-clamp recordings ([22,23]) recorded with two different input conditions: 1) a Step and Hold (SH) and 2) a Frozen Noise (FN) protocol (see Methods). We want to understand the influence of a physiologically realistic FN input on classification using commonly used features in contrast to SH input. To contrast the heterogeneity between SH and FN input conditions, we compare compare classification in these two input conditions using the neuronal spike waveforms and action potential attribute sets and also measure the similarity in the waveforms and action potential attributes between each cluster across the two input protocols.

0.1.1 Stimulus dependence of neural classification using intracellular action potential waveforms

To understand if waveform-based classification differs under FN and SH input conditions, we analyzed control (aCSF) trials from a total of 186 in-vitro whole-cell patch-clamp neural recordings that consist of the same cell recorded under 2 different input protocols, namely Step and Hold (SH) and Frozen Noise (FN) (Fig. 1 **c-d**, see methods). Using extracted waveforms of the same length (5 ms) from both FN and SH trails, we standardize (subtracting the mean and scaling to unit variance) the waveforms and then normalize the data using an L2 norm. Next, we apply an unsupervised high dimensional clustering algorithm that combines UMAP and Louvain Community detection (see methods) and found 7 clusters for SH trials as well as 7 clusters for FN trials (Fig. 1 **a-b**) respectively. We also measured the stability of the clusters (see Methods) against changing the hyperparameter (resolution parameter) for the unsupervised method. We found the clusters to be stable (low standard deviation) for the chosen resolution parameter (chosen at a value of 1.0) (Fig.1**g**).

From inspecting the two-dimensional UMAP projection, we found that the manifold representing the waveforms in the FN protocol (Fig. 1**a**) consists of two broad sub-divisions, whereas the SH protocol manifold seems much more connected and spread out (Fig. 1**b**). We compared waveform shapes across clusters by creating a pairwise cosine similarity matrix between FN and SH protocols and then taking the average of the submatrix for each SH and FN cluster pair. The average similarity across clusters is summarized using a heatmap (Fig. 2**e-f**). We first perform this comparison for each protocol separately and find that some cluster waveform shapes are similar (cosine similarity > 0.9) to their immediate neighbors (e.g. cluster 7, 3, and 4 in SH and cluster 6, 3, and 4 in FN in (Fig. 1**e**) for both FN and SH protocols. Similarly, comparing the waveform shapes between SH and FN, we find that narrow-width clusters for the SH protocol (4, 3, and 7) are highly similar to narrow-width clusters for the FN protocol (6, 2, and 3). We also project all the SH and FN waveforms together on a single embedding space and observed they were distinct (Appendix. 0.7**a**), suggesting that waveform shapes are not fixed across protocols. We then test if there exists a drift in the recordings by dividing the 360-second trial into two parts and comparing the waveforms between the first and the second trial by projecting the waveforms on the same 2D UMAP space. Both halves are overlapping suggesting that waveform shapes are similar between the two halves of the trail.

Finally, we calculate the likelihood for neurons in one of the SH clusters to be clustered together in one of the FN clusters (see methods and Fig. 2**h**). We find that for each SH cluster, the likelihood of clustering in one of the FN clusters is spread without a strong majority, suggesting that neurons grouping in the SH protocol do not group in the FN protocol. We further quantify the clustering agreement between the SH and FN clusters using an adjusted random score (ARI) and adjusted mutual information score (AMI) between FN and SH clusters. Both

measures show low values ($ARI = 0.085$, $AMI = 0.133$). This confirms that neurons cluster differently between SH and FN protocols. To verify if the found clusters don't have a bias as a result of the method that we use, we repeat the analysis using the Ensemble clustering method (Appendix. 8) to cluster the SH and FN waveforms. Taking the average of the highest values for each row in the co-clustering matrix, we find 84% correspondence for FN waveforms between wave map and ensemble clustering and 96% correspondence for SH waveforms between wave map and ensemble clustering. Comparing the ensemble clustering with Louvain clustering using an adjusted mutual information score (AMI Louvain vs Ensemble (FN) = 0.765, AMI Louvain vs Ensemble (SH) = 0.736), we find a high level of similarity for both SH and FN protocols. In conclusion, we find that clustering neurons into cell classes based on their waveforms results in different cell classes in the SH and FN input protocols, resulting from differences in the waveform shapes due to the stimulation protocol. This shows that waveform-based neuronal identity is stimulus-dependent.

0.1.2 Stimulus dependence of neural classification using action potential attributes

It has been found that excitability measures such as total spike count and AMPA conductance threshold (dynamic) vs rheobase (static) have a low correlation between static and dynamic stimulus conditions ([24],[25]). As we found in the previous section that cell clustering based on waveforms depend critically on the stimulation protocol. This intrigued us to investigate if neuronal classification based on action potential attributes (see Methods), that incorporate among others spiking dynamics, spike threshold and action potential height and width (a total of 22 attributes), is also input-specific. We designed a set of features to allow for comparisons across input protocols. This choice of attributes is based on previous literature that clusters neurons based on electrophysiological attributes. The attribute list is not exhaustive but rather confined as features can be input specific. The properties included in the spiking dynamics set incorporate the firing statistics such as inter-spike interval (ISI), firing rate, and action potential count among others that are commonly studied. Next, we design a spike threshold attribute set to capture spike threshold-related parameters for the trial. The spike thresholds were calculated for each action potential in each trial for the analysis (see Methods). Finally, we measure the action potential height and width parameters, even though we have clustered based on average waveforms in the previous section. We do this for two reasons, first, height and width parameters are electrophysiological parameters commonly used in the clustering literature and second, the average waveforms studied previously do not capture the change in the height and width of the waveform within each trial. We used descriptive statistics such as mean, median, minimum and maximum values for some of the features to capture the distribution of these features in each trial.

We use the unsupervised UMAP+Louvain clustering method (explained in the previous section) on the 22-dimensional feature set that we extracted from the SH and FN trials for 186 neurons. We find 7 classes for SH and 7 classes for FN protocol respectively (Fig. 4 **a-b**). Unlike in the waveform-based clustering, the UMAP representation is more continuous for the FN trials than for the SH protocols (Fig. 2 **a and b**). We visualize the differences attributes between the classes using a radar plot to investigate this further. We consider features based on spiking dynamics (such as firing rate, ISI, etc.), spike threshold, and action potential height and width separately for both SH and FN protocols (Fig. 2 **d-e** and methods for a list of all features and their sub-classification). We plot these features for each trial (thin line) along with the mean for each class (thick line). On visual inspection, we observe that the means for each cluster (thick line) in both SH and FN protocols are non-overlapping for all the 3 attribute sets (spiking dynamics, spike threshold, and AP height and width) respectively. Suggesting that each cluster has distinct action potential attributes. We measure the stability of the clusters (see methods) as a

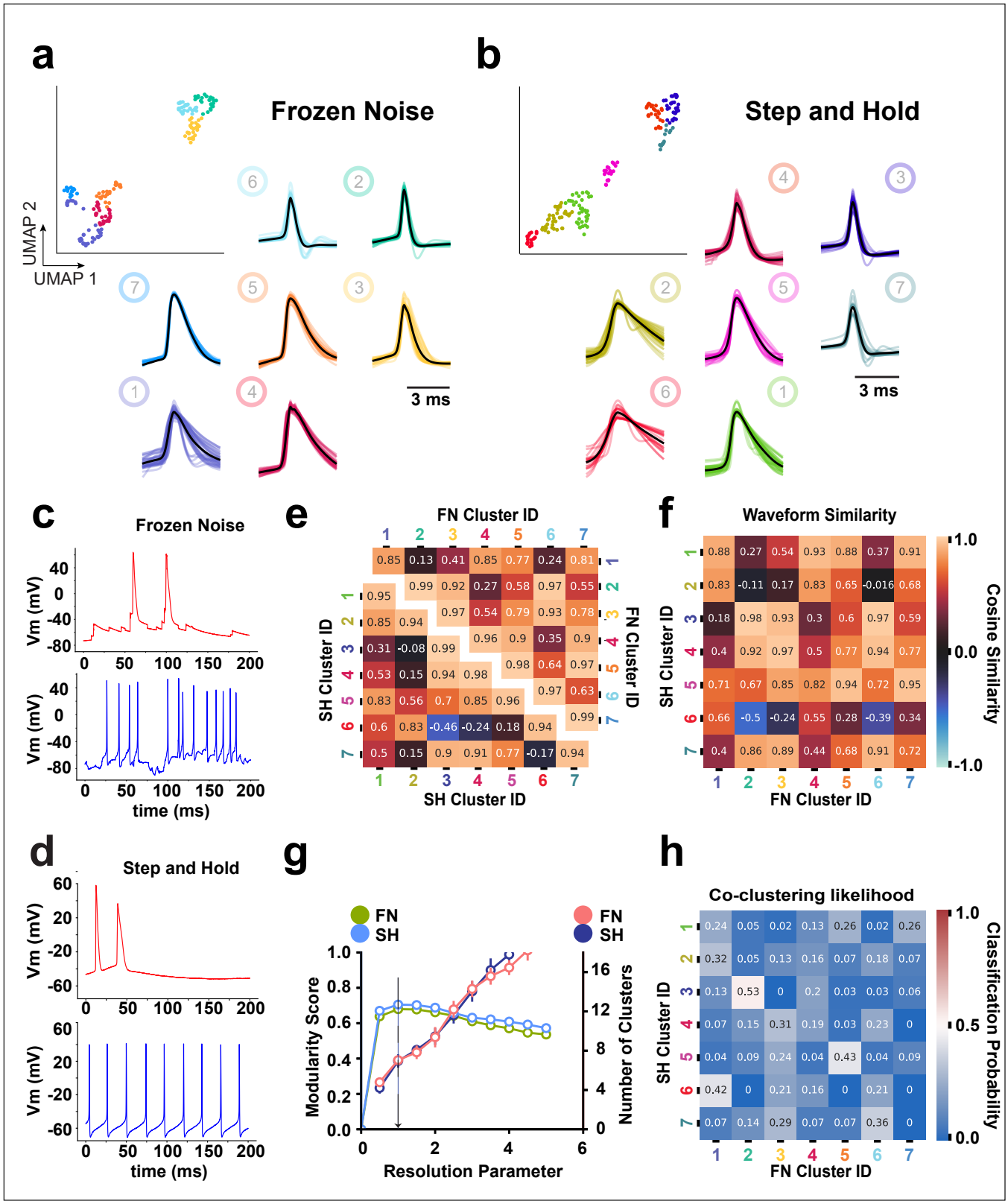


Figure 1: **Intracellular waveform-based neural identity (cluster membership) depends on the stimulus type.** Each of the 186 cortical neurons was studied using two distinct stimulation protocols: a step and hold (SH) protocol and a frozen noise (FN) protocol. (**continued on the next page**)

(a-b) UMAP representation of spike waveforms from the **FN** and **SH** protocols, colored with their respective cluster identities (top) and their corresponding waveform shapes (bottom). **(c)**200 ms of a FN recording of a neuron with a broad-width spike shape (red) and narrow-width spike shape (blue). **(d)**200 ms SH trial of a neuron with a broad-width spike shape (red) and a narrow-width spike shape (blue). **(e)** The confusion matrix shows an inter-class waveform comparison based on the cosine similarity across SH and FN classes. **(f)** The confusion matrix shows the cosine similarity-based comparison between waveform shapes for each SH class with all FN classes (row), and each SH class shows a similarity with at least 2 FN classes (> 0.8). **(g)** The mean and standard deviation of the modularity score for SH and FN protocols (left axis) and the number of clusters (right axis) as a result of running the clustering algorithm 25 times over 90% of the data sampled randomly each time, with changes in the resolution parameter ranging from 0 to 5 with increments of 0.5. The resolution parameter that maximizes the modularity score (black arrow) was chosen. (h) The heat map shows the likelihood of waveforms for each SH class being classified in one of the FN classes. Class identification between FN and SH protocols is inconsistent, as they show low values for the Adjusted Random Score (0.085) and Adjusted Mutual Information score (0.133).

result of changing the hyperparameters (resolution parameter) for the unsupervised clustering method. We find the clusters to be stable (low standard deviation) for the chosen hyperparameter (resolution parameter = 1.0) (Fig.2**c**).

We want to compare action potential attribute sets across SH and FN trials, so we quantify the differences across clusters for the FN and SH protocols using the cosine similarity between the feature vectors. We calculate the pairwise cosine similarity matrix between all SH and FN trials and take the means of the sub-matrix for neurons in each SH and FN cluster pair. This gives a cosine similarity matrix comparing each SH cluster with each FN cluster. We observe that none of the SH clusters show a high similarity (> 0.9) with FN clusters, suggesting that action potential attributes differ drastically as a result of input protocol (Fig. 2**f**). We also overlay the UMAP embedding for each feature vector for the SH and FN protocol (Appendix. 0.7**c**) and found that the SH and FN feature manifolds are entirely different.

To understand the role of input in overall action potential attributes based cluster assignment, we calculate the likelihood (see Methods) for a neuron in one of the SH clusters to be clustered in one of the FN clusters (Fig. 2**g**). We find that for each SH cluster, the likelihood for clustering in one of the FN clusters is spread over the FN clusters without a strong majority for a specific FN cluster. We further quantify the correspondence between SH and FN clusters using an adjusted random index score (ARI) and adjusted mutual information score (AMI), both values were found to be low (ARI = 0.149 and AMI = 0.206). This shows neurons cluster differently between SH and FN protocols based on Spike Event-based features. These results suggest that action potential attributes based clustering of neuronal populations is input dependent.

0.2 Functional classification of neurons stimulated by the FN protocol based on different feature sets

We have shown in the previous section that neuronal classification based on waveforms and action potential attributes is input-dependent. We want to expand our understanding of which features result in distinctive functional classifications within the FN-stimulated neurons. To understand the variance in the neuronal population captured by different attribute sets, we perform classifications based on four different attribute sets that capture 1) the commonly used action potential attributes, 2) passive biophysical attributes, 3) adaptation attributes, and 4) linear input

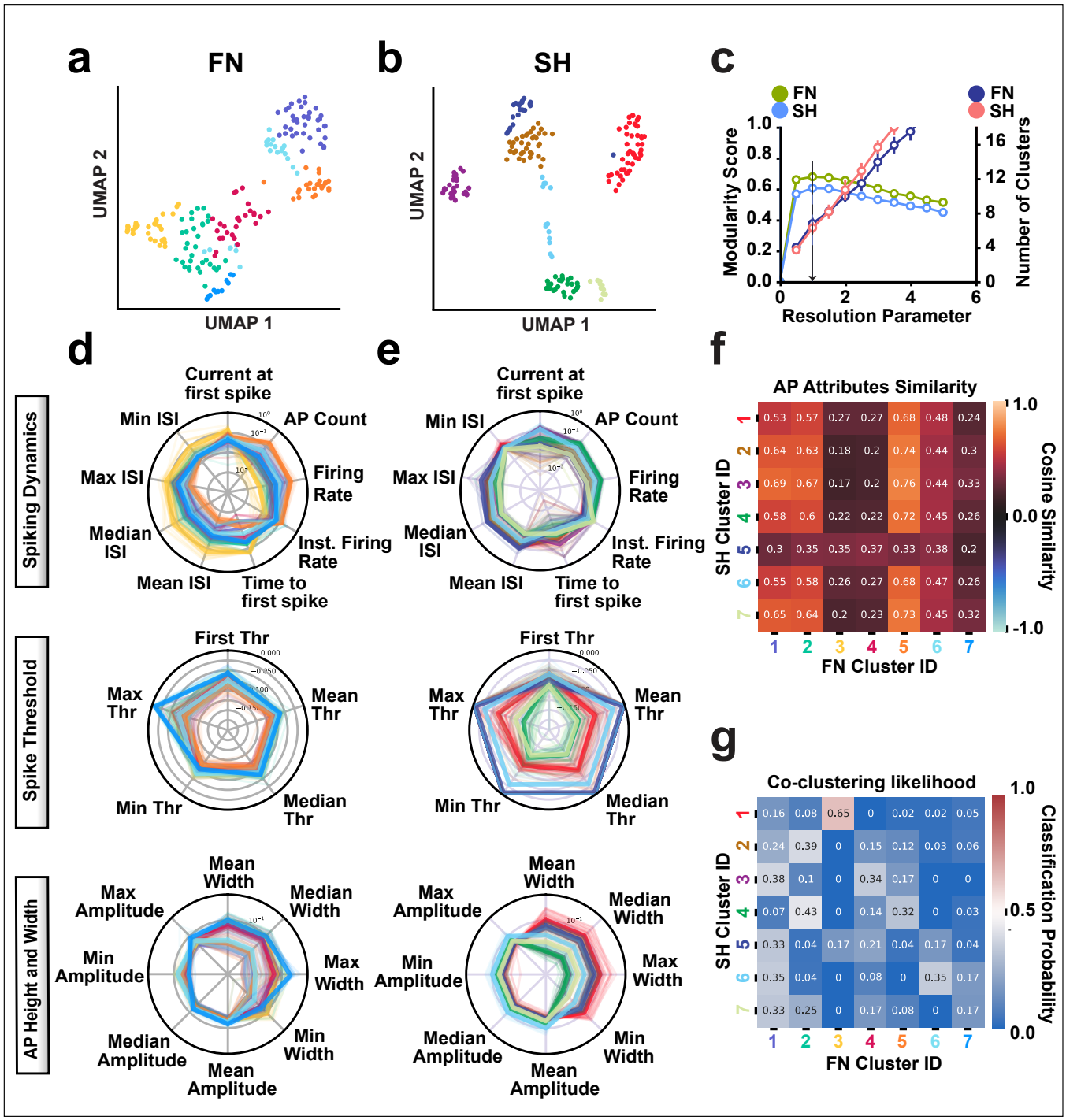


Figure 2: **Neural identity based on Action Potential attributes depends on the stimulus type.** (**a-b**) UMAP representation of Action Potential attributes for neurons was recorded with FN and SH protocols, and each color corresponded to a cluster identity. (**c**) shows the mean and standard deviation of the Modularity Score for SH and FN protocols (left axis) and the number of clusters (right axis) as a result of running the clustering algorithm 25 times over 80% of the data with changes in the resolution parameter ranging from 0 to 5 with an increment of 0.5, the resolution parameter is chosen that maximizes the Modularity score (black arrow). (**d-e**) The radar plot shows action potential attributes divided into spiking dynamics, spike threshold, and AP height and width features for each neuron, recorded with SH and FN protocols with each cluster in their respective color. Thick lines represent the mean over trials for each cluster. (**f**) Heatmap comparing action potential attributes using cosine similarity between SH and FN clusters. It can be seen that none of the values in the heatmap is greater than 0.8 suggesting that SH and FN action potential attributes are dissimilar. (**Continued on the next page**)

(**g**) Heatmap showing the likelihood of neurons to classify in one of the FN classes for each SH class. Note that the class identification between FN and SH protocols is inconsistent. The adjusted Random Score was 0.149 and the Adjusted Mutual Information score was 0.206, suggesting inconsistency between SH and FN clusters.

filters approximated using the Spike Triggered Average (STA), to assess input feature selectivity.

Ample experimental evidence suggests that cortical neurons can be divided into two broad functional categories ([26]), namely excitatory (glutamatergic) and inhibitory (GABAergic), based on the type of effect (either excitation or inhibition) they have on their post-synaptic neurons ([8]). Excitatory and inhibitory neurons have also been found to have distinct electrophysiological properties and thus are known to perform different functions. We therefore subdivide our data into excitatory and inhibitory groups to study the diversity within and across populations. Previous studies have associated neuronal waveform shapes with functional identity ([27,28]). The broad and narrow spike-width neurons have also been found to have a characteristic firing statistic, i.e. neurons with narrow-width waveforms were found to have a high firing rate (putatively inhibitory), and neurons with broad-width waveforms were found to have lower firing rates (putatively excitatory) ([19,29,30]) in the barrel cortex. This suggests that inhibitory neurons can be putatively characterized by narrow width and high firing rate and that excitatory neurons can be putatively characterized by broad width and low firing rate. Also, barrel cortical excitatory neurons are more adaptive compared to inhibitory neurons ([31]). Based on this reasoning, we partitioned our data into a putative excitatory and an inhibitory population.

We extract the average intracellular waveforms from the entire dataset, 312 cells in total. Note that the number of neurons under analysis is much larger than in the previous sections because more FN than SH experiments were performed. We then apply the UMAP+Louvain algorithm to classify intracellular waveforms and find 8 clusters. (Fig. 3**a**) shows the UMAP projection of all waveforms with their corresponding cluster label colors. Next, we plot the distribution of the firing rates and half-widths for each cluster with matching colors using a violin plot (Fig. 3**b-c**). We observe clusters 1,5 and 6 to have a narrow width and high firing rate relative to the rest of the clusters; therefore, we categorize these clusters as putatively inhibitory (I) and the rest as putatively excitatory (E). We compare the average cosine similarity between E/I populations. We find that the excitatory and inhibitory populations are significantly different (Welch's ANOVA, (F(2,49135)=19148.00, ***p= 0.0; Post-hoc Games-Howell test, E vs I (**p=0.001), I vs ExI (**p=0.001), and E vs ExI (**p=0.001), Fig. 3**d**), but the excitatory population is more heterogeneous than the inhibitory one, based on the average similarity of the waveforms within the E/I population (Fig. 3**d**).

0.2.1 Action potential attributes based neuron profiles using the FN protocol

As pointed out in the previous section, we aimed to understand and compare the usefulness of commonly used physiological attributes in uncovering functional classification in neuronal populations when neurons receive a physiologically realistic FN input. For that aim, we study commonly used action potential attributes for excitatory and inhibitory populations separately to discern if action potential attributes sufficiently capture within-population heterogeneity in excitatory and inhibitory groups. We also aim to unravel the differences between excitatory and inhibitory populations regarding their action potential attributes. For this, we extract 22 attributes (see Methods) subdivided into spiking dynamics, Spike threshold, and action potential height and width attributes with their descriptive statistics incorporating the mean, median, minimum, and maximum values.

We cluster the E/I populations separately based on the features and find 7 clusters for excita-

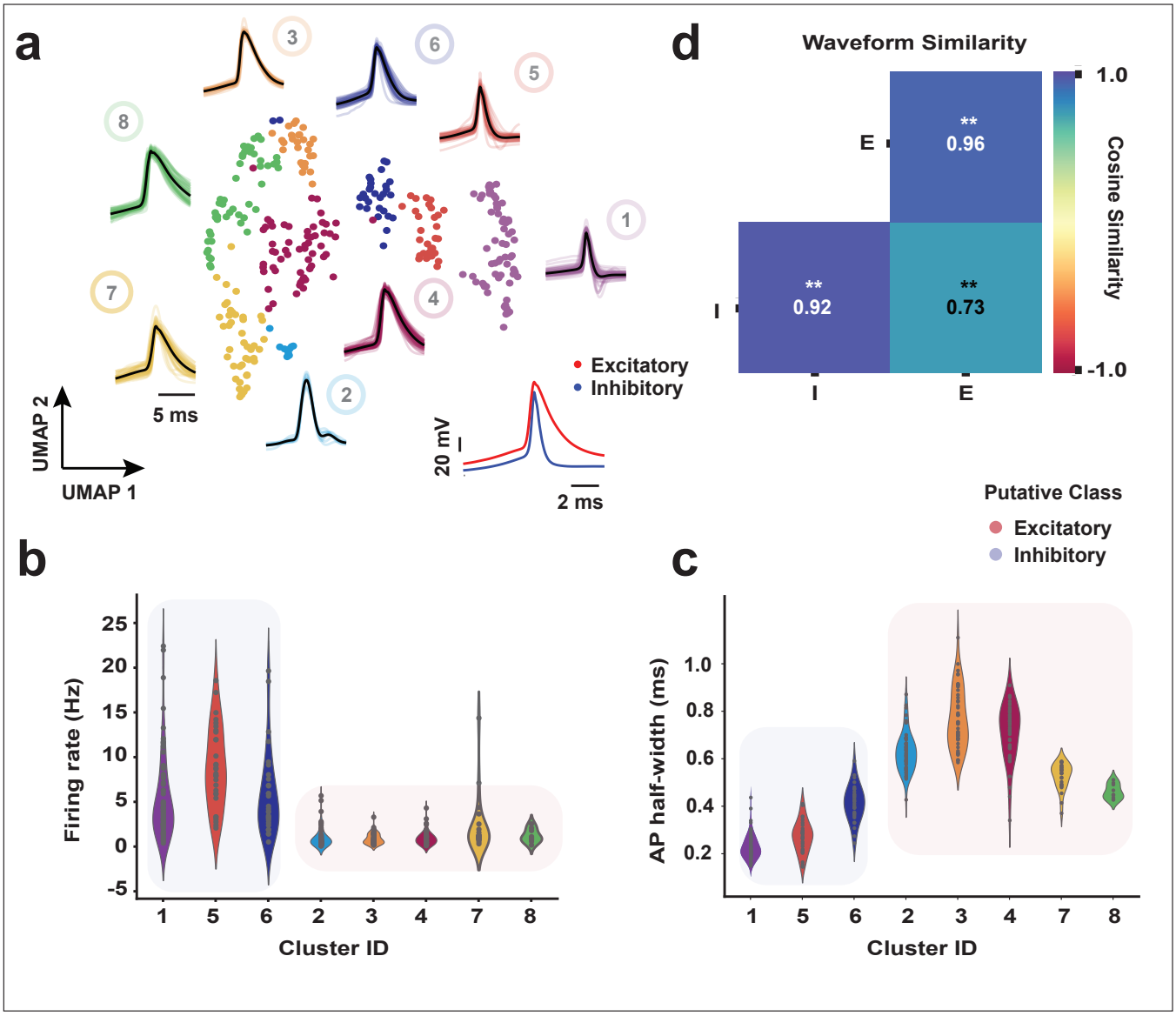


Figure 3: **Putative classification of excitatory and inhibitory neurons based on waveform and firing characteristics.** 312 cortical neurons were recorded using the FN protocol. Before studying neuronal classification based on different functional attributes, a putative excitatory and inhibitory classification of neurons was performed based on firing rate and AP half-width attributes. (**a**) The UMAP representation of the clustered average waveform along with their corresponding shape represented with the same color from 312 neurons recorded from layer 2/3 in the somatosensory cortex. (inset) average waveforms for putative excitatory and inhibitor populations. (**b**) Violin plot showing the distribution of firing rate across clusters identified in (**a.**), cluster IDs [1, 5, 6; n= 66, 30, 30] were found to have higher firing rates and are labeled as putatively inhibitory (n=126) and putatively excitatory (n=186). (**c**) shows the AP-half width across clusters identified in (a), cluster IDs [2, 3, 4, 7, 8; n= 58, 48, 44, 27, 9] were found to have lower half-width. (**d**) The average cosine similarity between putative excitatory and inhibitory classes. On average the putative Excitatory waveform group shows higher heterogeneity and excitatory and inhibitory waveforms differ. Cosine similarity measures between in-population excitatory (E), in-population inhibitory (I), and across excitatory-inhibitory populations (ExI) show significant differences. Welch's ANOVA, $F(2, 49135) = 19148.00$, ***p= 0.0); Post-hoc Games-Howell test, E vs I (**p= 0.001), I vs ExI (**p= 0.001), and E vs ExI (**p= 0.001).

tory cells (shaded in red) and 6 clusters for inhibitory cells (shaded in blue). The UMAP representation of the spike-based properties for E/I populations with unique colors for each cluster for the chosen cluster parameter is shown in (Fig. 4**a-b**). We show the cluster stability (see Methods) in (Fig. 4**c**), the number of clusters is stable (low standard deviation) for the chosen resolution parameter (black arrow). We also test the stability of the clusters by excluding one attribute at a time and repeating the stability analysis (see Methods) and find the inhibitory clusters to be more stable to attribute exclusion (Appendix. 9**a-b**). We compare the means of action potential attributes simultaneously between excitatory and inhibitory populations and find them to be significantly different (one-sided MANOVA (see Methods), Wilks' lambda; $F(21, 290.0) = 35.1841$; $p = 0.000^{***}$, Pillai's trace; $F(21, 290.0) = 35.1841$; $p = 0.000^{***}$, Hotelling-Lawley trace; $F(21, 290.0) = 35.1841$; $p = 0.000^{***}$, Roy's greatest root; $F(21, 290.0) = 35.1841$; $p = 0.000^{***}$). To identify the relative importance of each attribute in the separation of excitatory and inhibitory populations, we perform a canonical correlation analysis (CCA) (see Methods) between the excitatory and inhibitory action potential attributes and calculate the loading (structure correlation) for each attribute (Fig. 4**d**). We find that spiking dynamics attributes have the most influence on the inhibitory canonical variates suggesting that these variables have the most influence on separating the inhibitory population apart from the excitatory population. Alternatively, AP height and width attributes strongly influence the excitatory population canonical variate, suggesting that AP height and width are the most discriminatory. These results also show that individual action potential attributes contribute differently to the canonical variate (i.e., the latent structure) for excitatory and inhibitory populations respectively.

Next, we want to compare the level of heterogeneity within excitatory (E-E) and inhibitory (I-I) populations as well as across (I-E) populations using the action potential attributes. We visualize the action potential attributes using a radar plot shown in (Fig. 4**e-f**). It shows the diversity of spiking dynamics, spike threshold, and AP height and width for each neuron in each cluster (with the same color as in the UMAP projection (Fig. 4**a-b**) along with the mean for the cluster (thick line). We observe that spiking dynamics attributes differ between excitatory and inhibitory clusters. The spiking threshold profiles were also found to be different across E/I clusters. Comparing the AP height and width attributes, however, show a similar profile across E/I clusters. We quantify the heterogeneity in the E/I clusters using a cosine similarity measure. The excitatory population had a significantly higher cosine similarity measure within the population than the inhibitory population (Welch's ANOVA, F(2,49135)=3626998.65, p= 0.0; Post-hoc Games-Howell test, E-E vs I-I (**p=0.001), I-I vs I-E (**p=0.001), and E-E vs I-E (**p=0.001), (Fig. 4**g**), suggesting that inhibitory action potential attributes are more heterogeneous than their excitatory counterparts. The mean cosine similarity score between excitatory and inhibitory populations are significantly lower than the within-population cosine similarity score for both excitatory and inhibitory populations respectively, suggesting that excitatory and inhibitory feature action potential attribute vectors are different from each other, reiterating the MANOVA results. The results above demonstrate that action potential attributes are different between excitatory and inhibitory populations. The inhibitory population is more heterogeneous in its action potential attributes than the excitatory population.

0.2.2 Passive biophysical feature and adaptation-based profile using the FN protocol

As we describe in the last two sections, we aim to compare neuronal attribute sets to find the attributes(s) most informative about heterogeneity in excitatory and inhibitory populations and try to understand how neurons cluster based on these properties. The passive biophysical attributes (i.e., membrane resistance, capacitance, etc.) shape the response properties of a neuron. Moreover, it has been reported that adaptation current which captures the adaptation properties of the membrane is another important neuron property that can determine cell classes with noisy input

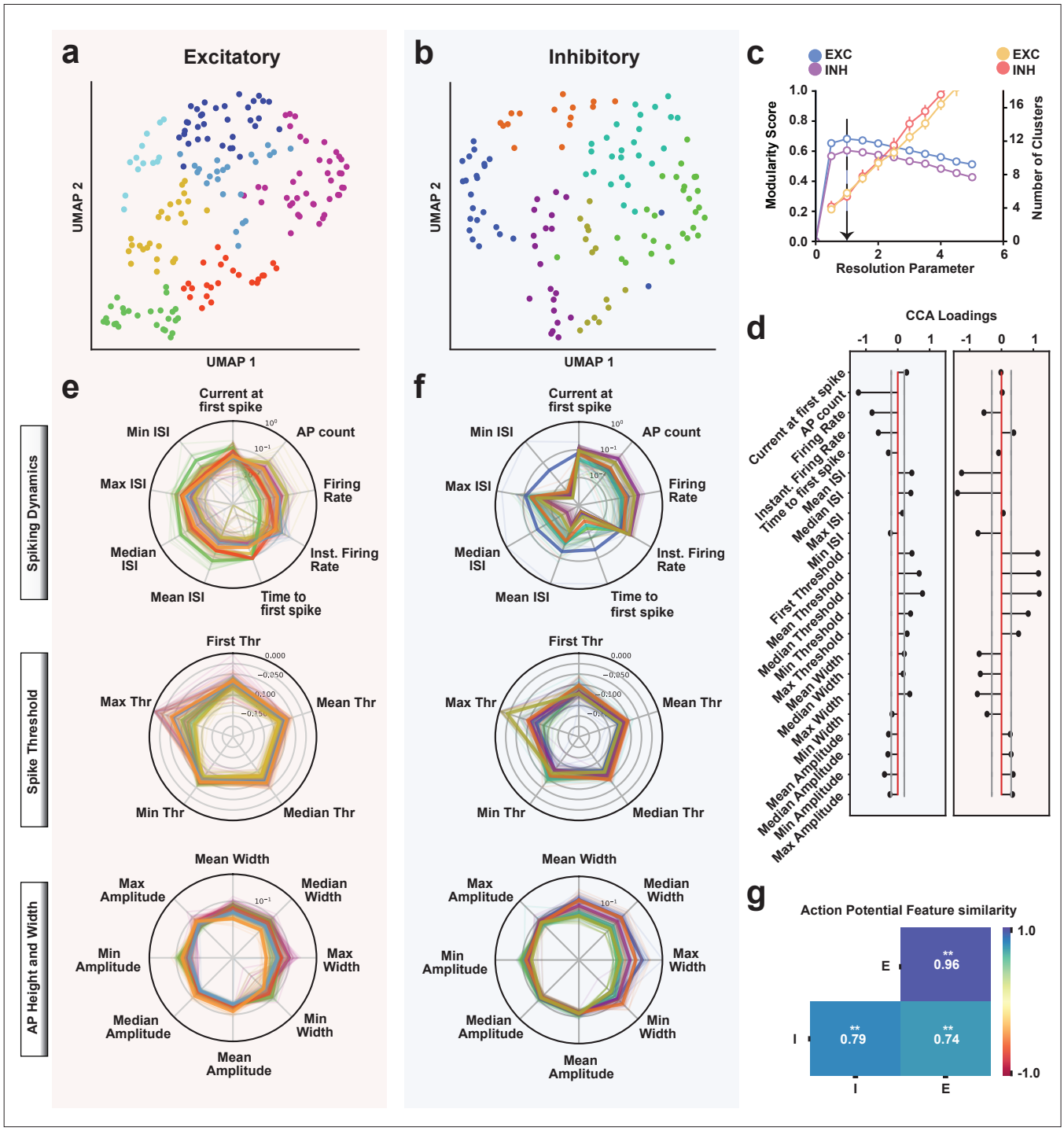


Figure 4: **Cell type-specific heterogeneity and the principal discriminatory attributes of the action potential.** NB Excitatory/Inhibitory results are grouped with red and blue backgrounds respectively. (**a-b**) The clustering procedure groups E/I neuron populations into 7 and 6 clusters respectively. (**c**) shows the chosen resolution parameters corresponding to the highest modularity and the corresponding number of clusters. (**d**) The means of excitatory-inhibitory action potential attributes were found to be significantly different (one-sided MANOVA, Wilks' lambda; $F(21, 290.0) = 35.1841$; $p = 0.000^{***}$, Pillai's trace; $F(21, 290.0) = 35.1841$; $p = 0.000^{***}$, Hotelling-Lawley trace; $F(21, 290.0) = 35.1841$; $p = 0.000^{***}$, Roy's greatest root; $F(21, 290.0) = 35.1841$; $p = 0.000^{***}$). We performed a post-hoc CCA analysis between the excitatory and inhibitory populations and showed the loadings for each attribute with a 95% confidence interval (grey line). (**e-f**) action potential attributes profile broken down into spiking dynamics, spike threshold, and AP height and width parameters plotted with the average for each cluster (thick line) and individual neurons (thin lines) and their corresponding cluster colors. (**Continued on the next page**)

Spiking dynamics parameters show the most variability across clusters for both excitatory and inhibitory groups. (**g**) The cosine similarity between E/I feature vectors. Inhibitory neurons are more heterogeneous in their action potential attributes than excitatory neurons. Welch's ANOVA shows a significant difference between cosine similarity measures between within-population excitatory (E-E), in-population inhibitory(I-I), and across excitatory-inhibitory populations (I-E) ($F(2, 49135) = 3626998.65$; $p = 0.0$); Post-hoc Games-Howell test, E-E vs I-I (**p= 0.001), I-I vs I-E ($**p = 0.001$), and E-E vs I-E ($**p = 0.001$).

([32]). We, therefore, consider the passive biophysical attributes and the adaptation current as potential candidates that capture heterogeneity within the neuronal population. To understand the heterogeneity of passive biophysical properties and adaptation parameters across the excitatory and inhibitory populations, we extract both a set of passive parameters (see methods) and an adaptation current by fitting a Generalized Leaky Integrate and Fire (GLIF) model on the first 100 seconds of each FN trial, using the automated method described by ([33]) (see methods). (Fig. 5**a**) shows an example of a 10s instance of a GLIF-fitted model and one of the original recordings. We characterized the goodness of fit by measuring the explained variance (EV) between the subthreshold membrane potential of the original data and the model. (Fig. 5**b**) shows the distributions of the EV and Γ for the entire dataset. We eliminate the models with an EV value below 0.7 and obtain a set of 307 samples (out of 312 samples).

We cluster the recordings into cell classes based on a set of passive biophysical parameters along with threshold adaptation constants (in total 6 features, see Fig.5**a-b**) using the UMAP+Louvain clustering method explained in the previous sections, we find 7 clusters for the excitatory population and 6 clusters for the inhibitory population. We examine the stability of the clusters (Fig. 5**c**) with different hyperparameters (resolution parameters, see Methods) and choose (black arrow) the resolution parameter with the maximum modularity score. The clusters for the chosen hyperparameter for both excitatory and inhibitory populations were stable (low standard deviation). We also test the stability of the clusters by excluding one attribute at a time and repeating the stability analysis (see Methods) and found the inhibitory clusters to be more stable to attribute exclusion (Appendix. 9**c-d**). (Fig. 5 **d-e**) shows the UMAP representation of the feature set (inset) for excitatory (red background) and inhibitory (blue background) populations, the radar plot shows the individual biophysical parameter values for each neuron and each thin line in the radar plot represents a single neuron and are color matched to respective clusters, the mean for each cluster is plotted with a thick line. We compare the means of all the passive biophysical attributes simultaneously between excitatory and inhibitory populations and find them to be significantly different (one-sided MANOVA, Wilks' lambda; $F(6, 301) = 23.2629$; $p = 0.000^{***}$, Pillai's trace; $F(6, 301) = 23.2629$; $p = 0.000^{***}$, Hotelling-Lawley trace; $F(6, 301) = 23.2629$; $p = 0.000^{***}$, Roy's greatest root; $F(6, 301) = 23.2629$; $p = 0.000^{***}$). To identify the relative importance of each passive biophysical attribute in the separation between excitatory and inhibitory populations, we perform a canonical correlation analysis (CCA) (see Methods) between the excitatory and inhibitory passive biophysical attributes and calculate the loading (structure correlation) for each attribute (Fig. 5**h**). We find dissimilar contributions from excitatory and inhibitory passive parameters towards their respective canonical variate, suggesting that each passive biophysical attribute contributes differently to the excitatory-inhibitory latent structure and therefore sets the excitatory-inhibitory populations apart. Next, we compare the level of heterogeneity between the passive feature vectors across the E/I populations (Fig. 5**i**) using the cosine similarity measure within and across excitatory populations and find no significant difference between the mean cosine similarity measure within excitatory and inhibitory populations as well as across excitatory-inhibitory population (due to low effect size) (Welch's ANOVA, F(2,66792)=0.16777; p= 0.84, Fig. 5**i**). These results suggest that even though the mean of the excitatory and in-

hibitory passive biophysical attributes are significantly different from each other (measured using MANOVA), the level of heterogeneity within excitatory-inhibitory populations respectively based on biophysical parameters is quite low. Therefore the passive biophysical attributes are not so informative about the neuronal heterogeneity.

As we saw above, passive biophysical features are not informative about population heterogeneity, this result is consistent with ([32]), which finds that passive biophysical features do not distinguish between neuron types. The adaptation current is more useful in putting neuron types apart. Therefore we classify the neurons using the adaptation current (η) extracted from the fitted GLIF model for the E/I population separately. We find 6 classes for the excitatory population (red) and 5 classes for the inhibitory population (blue). In (Fig. 5**f-g**) we show the UMAP representation of the adaptation current and the corresponding average normalized shapes for each cluster (inset) each cluster has its respective color in the UMAP as well as shape plots. We observe that the excitatory adaptation currents have a stronger negative amplitude than their inhibitory counterparts. Moreover, the adaptation currents of the inhibitory population relax back to their resting values at earlier times than for the excitatory population.

We quantify the heterogeneity of adaptation currents within excitatory and inhibitory populations using the cosine similarity measure. The heatmap in (Fig. 5j) shows that the excitatory adaptation currents have a significantly higher average cosine similarity compared to the inhibitory adaptation currents, suggesting that the inhibitory adaptation current is more heterogeneous than the excitatory population (Welch's ANOVA; F(2,46968) = 2796.77, **p = 0.0; Post-hoc Games-Howell test E vs I (**p = 0.001), I vs ExI (p = 0.75), and E vs ExI (**p = 0.001), Fig. 5**j**). The similarity measure is not significantly different between the inhibitory (I) population and across the inhibitory and excitatory (I vs E) population. This suggests that inhibitory adaptation currents are as different from each other as they are different from the excitatory adaptation currents. These results suggest that adaptation current profiles are different between excitatory and inhibitory populations. These results also show that adaptation currents are useful for understanding the heterogeneity within the inhibitory population but not so much for the excitatory populations.

0.2.3 Neuronal classification based on linear input filter approximated using a Spike Triggered Average (STA) in the FN protocol

In last two sections, we explored neuronal heterogeneity based on action potential, passive biophysical, and adaptation attributes in neurons while responding to an FN input. We found that except for passive biophysical attributes, other properties show both within population across population heterogeneity for both excitatory and inhibitory populations. We observe a higher level of heterogeneity for inhibitory populations. Since we aim at understanding the extent to which various neuronal properties help uncover neuronal diversity, in this section we focus on the linear input filter of neurons and aim to understand the neuronal diversity based on this attribute. We consider this attribute important as neurons respond strongly to specific features in the input ([34]), making linear input filter an important property to study the functional diversity of the neuronal population.

The spike-triggered average (STA) method, as described in ([35]), estimates a neuron's linear input filter by identifying the features in the input that trigger the neuron to spike. It does this by averaging the input signal over a specific time window preceding each spike, using data from all observed spikes. This makes the STA a useful method to approximate the linear input filter of a neuron. We want to investigate the STA diversity across the E/I population, for this, we extract the STA from all the neurons (see Methods). For calculating the STA, we use the injected current which is the result of shifting a theoretical dimensionless input with a constant baseline and scaling it by a factor as explained by ([21,22]).

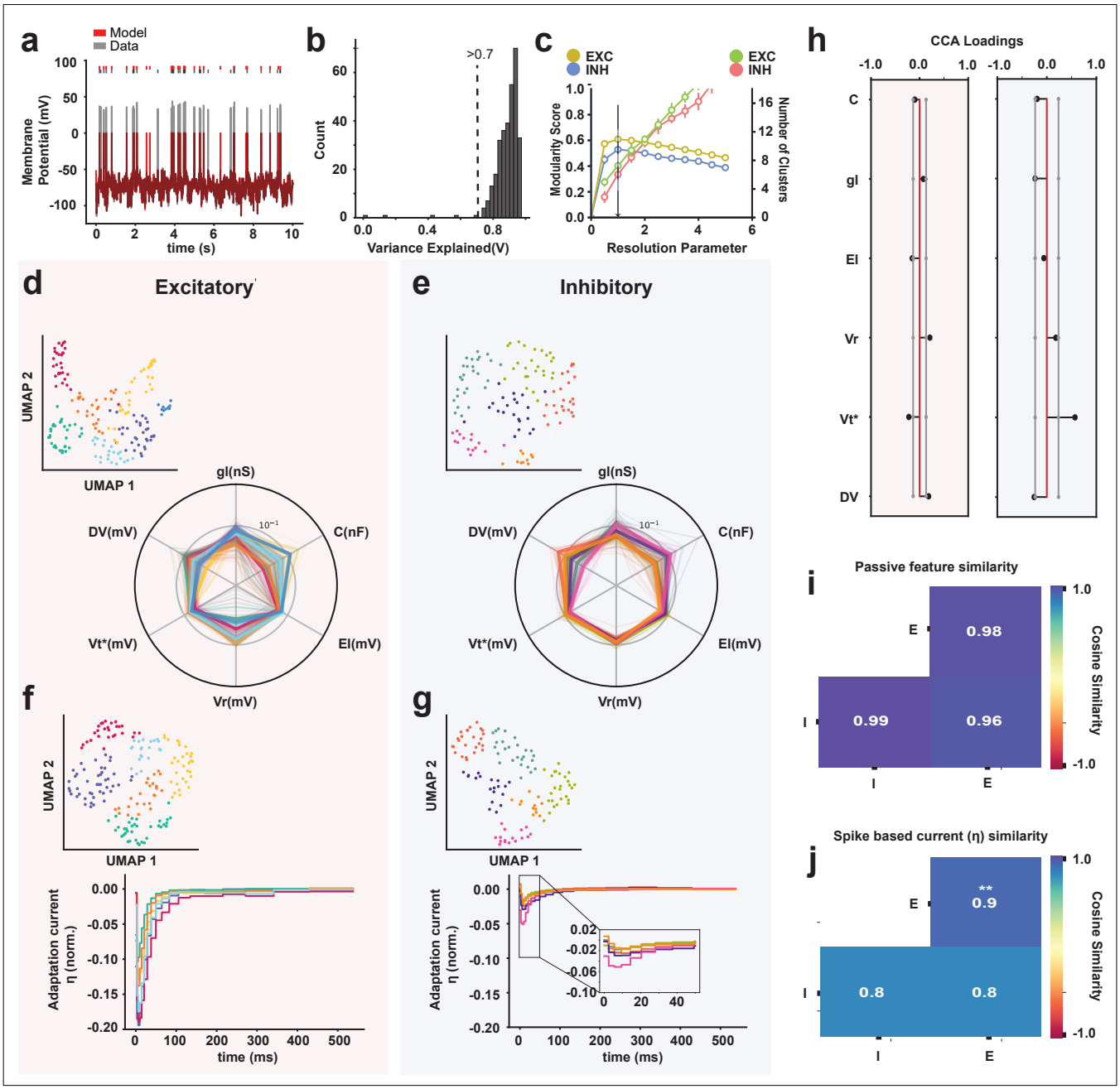


Figure 5: **Discriminatory power of passive biophysical and adaptive membrane properties for neuronal classification.** A GLIF model was fitted to the membrane potential recordings of neurons stimulated using the FN protocol. (**a**) An example of a 10 s long snippet of the GLIF model (red) with the original trial (grey). (**b**) Histogram showing the variance explained by the subthreshold membrane potential of the GLIF model. We included only models that exceeded a subthreshold variance explained higher than 0.7 (dashed line) for clustering purposes. (**c**) The chosen resolution parameters correspond to the highest modularity and the corresponding number of clusters for passive parameters. (**d-e**) UMAP representation of 7 clusters for excitatory (left) and 6 clusters for inhibitory (right) found using the passive biophysical parameters (inset). The radar plot shows the passive biophysical profile of excitatory (left) and inhibitory cells (right), thick lines show the mean for each cluster and thin lines show the passive biophysical properties for each neuron colored in its respective cluster color. (**f-g**) UMAP representation of 6 clusters for excitatory (left) and 6 clusters for inhibitory (right) neurons found using the adaptation current (inset) and the corresponding adaptation current shape. (**h**) Post-hoc CCA between excitatory and inhibitory population and the loadings for each attribute with a 95% confidence (grey line). (**continued on the next page**)

The excitatory-inhibitory passive biophysical attributes means were found to be significantly different (one-sided MANOVA, Wilks' lambda; $F(6, 301) = 23.2629$; $p = 0.000^{***}$, Pillai's trace; $F(6, 301) = 23.2629$; $p = 0.000^{***}$, Hotelling-Lawley trace; $F(6, 301) = 23.2629$; $p = 0.000^{***}$, Roy's greatest root; $F(6, 301) = 23.2629$; $p = 0.000^{***}$). (**i**) Heatmap showing passive biophysical attributes vectors compared within the excitatory (E-E) and inhibitory (I-I) population as well as across excitatory and inhibitory populations (I-E) using the cosine similarity measure, the values are averaged for I-I, E-E, and I-E comparisons. E/I population passive parameters were not significantly different for both within (I-I, E-E) and across (I-E) population comparisons (Welch's ANOVA, F(2, 66792) = 0.16777; p= 0.84). (**j**) Heatmap showing the cosine similarity-based heterogeneity measurement of normalized adaptation currents for E-E, I-I, and I-E populations. The average cosine similarity values of adaptation current currents were significantly higher for the E-E group than the I-I group suggesting that the inhibitory population is more heterogenous. The I-I and I-E mean cosine similarity values were not significantly different. (Welch's ANOVA, F(2,46968)=2796.77, **p= 0.0); Post-hoc Games-Howell test I-I vs E-E (**p=0.001), I-I vs I-E (p=0.75), and E-E vs I-E (**p=0.001)).

We cluster the excitatory and inhibitory population STAs separately using an unsupervised UMAP+Louvain clustering method (see Methods). We find 7 clusters for the excitatory population STAs (red) and 8 for the inhibitory population STAs. The STAs were normalized and standardized before clustering (see Methods). The UMAP representation and the corresponding averaged and normalized STAs for each cluster are shown in (Fig. 6**a-b**). A visual inspection shows that inhibitory and excitatory STA shapes differ in their peak amplitudes and in the maximum slope of their initial rise (Fig. 6**a-b** inset), this difference is shown more clearly in (Fig. 6**d**). The stability of the clustering algorithm for the chosen parameters is shown in (Fig. 6**c**). To quantify the heterogeneity within the excitatory and inhibitory population as a result of that linear input filters, we first calculate the STA shape similarity between clusters, using a cosine similarity between excitatory and inhibitory STA clusters (Fig. 6**e-f**). The values shown in the heatmap in Fig. 6**e-f** were calculated by averaging the cosine similarity matrices between STAs of each cluster pair and within each cluster (for the within-cluster comparison, only the upper triangular values were included in the average). The average cosine similarity value between excitatory STA clusters shows higher values (> 0.9) with each other except for clusters IDs 2 and 6, which show a low similarity with every other cluster except for itself. On average, each excitatory cluster shows a high similarity (¿0.9) with more than 2 other clusters. On the other hand, the average cosine similarity values between inhibitory clusters is low (¡0.9), with some exceptions such as cluster IDs (2, 8), (3, 6), (5, 4), (6, 5) and (8, 6). On average, the STA of each inhibitory cluster is highly similar to approximately 1 other cluster. The average cosine similarity value is high between STAs within each inhibitory cluster. We summarize the STA similarities within the excitatory (E-E) and inhibitory populations (I-I) as well as across the two populations (I-E) in the heatmap in (Fig. 6**g**). This is done by averaging the cosine similarity matrices between STAs within excitatory (E-E) and inhibitory (I-I) populations as well as across excitatory and inhibitory (I-E) populations, these matrices are shown in (Appendix. 11**b**). For within-population averages (E-E and I-I), we only take the upper-triangular portion of the matrix. The average cosine similarity value within the inhibitory population is significantly lower than that of within the excitatory population (Fig. 6**g** Welch's ANOVA, F(2,48825)=1005.90, **p= 0.0; Post-hoc Games-Howell test, E vs I (**p=0.001), I vs ExI (**p=0.001), and E vs ExI (**p=0.001)), suggesting that inhibitory STAs are more heterogeneous than excitatory ones. Also, the STA shape similarity between excitatory and inhibitory populations is significantly lower than the within excitatory and inhibitory similarity, suggesting that excitatory and inhibitory STAs are different. We can see that on average inhibitory neurons respond to more diverse features in the input than excitatory

neurons.

0.2.4 Comparing Physiological Heterogeneity across attribute sets using Multi Correlation and Factor Analysis

We aim to explore what a physiologically realistic stimulus reveals about the functional heterogeneity of a neural population. For this aim, we need to compare FN-based neuronal attribute sets and find out which attribute set is the most informative about neuronal heterogeneity under the dynamic FN stimulation protocol. This is a complicated problem, as the four attribute sets are of varying dimensions. To perform this comparison, we use a method known as Multi Correlation and Factor Analysis (MCFA,[36]), which is an unsupervised multi-set integration method based on probabilistic Principal Component Analysis (pPCA) and Factor Analysis (FA), that can help to understand the common and shared factors across multi-modal data (see methods &[36]). We use this method to compare the private and shared variance explained by the 4 attribute sets we examined in the previous sections (i.e., action potential attributes, passive biophysical attributes, adaptation currents, and linear input filters (STA)). We use the pPCA space to model the shared structure across attributes (see Methods), and the residuals are modeled as a private structure for each attribute set using factor analysis (see methods).

Comparing the variance explained by the shared structure across attributes, we found that the excitatory population which is inferred by a 4-dimensional shared structure, explains almost 50% of the variance in action potential and passive biophysical attributes (Fig. 7**a** top), this is higher than linear input filter (STA) (32.9 %) and adaptation current attribute (19.7 %). Most importantly, we see that the linear input filters (STAs) explain the highest private variance (57.2 %), followed by adaptation current (41.2 %). The action potential and passive biophysical attributes have a relatively lower private variance (35.7 %). We can further investigate the contribution to the most important shared factor by each attribute, through which we find that action potential and passive biophysical features explain the most variance for the most important shared dimension (Fig. 7**a**(middle, bottom)). Similarly, for the inhibitory population, explained by a 3 dimensional shared structure, we find that the linear input filter (STA) explains the most amount of private variance (80.6 %), much higher than the excitatory population (Fig. 7**b** (top)), followed by adaptation current (29.4 %), action potential features (35.7 %), and passive biophysical features (17.1 %) respectively. The action potential features and Passive biophysical features explain the most and almost equal amount of shared variance (33.8 % and 33. 85 % respectively) (Fig. 7b (middle)), followed by adaptation Current (29.4 %) and linear input filter (9.9 %) (Fig. 7**b** (top)). We find that the adaptation current and passive features explain the largest amount of variance for the most important shared dimension (Fig. 7**b** (bottom)) for the inhibitory population. It is important to observe that the linear input filter explains more than 85% of the total variances (shared+private) for both excitatory and inhibitory populations, higher than the other 3 attribute sets.

The high values of private variance explained by the linear input filters (STAs) in both excitatory and inhibitory cases show that passive biophysical, action potential, and adaptation currents are not highly correlated with the linear input filter of a neuron. Therefore, the linear input filter contains unique information about the excitatory and inhibitory population that is not shared with the other 3 attribute sets. Most importantly, the high value of private variance for the linear input filter, along with low residual values compared to other attribute sets indicates that linear input filters are likely to be the most informative attributes for explaining neuronal heterogeneity compared to passive biophysical, action potential, and adaptation current attributes. On the contrary, high values of shared variance for passive biophysical attributes and action potential attributes show that these attributes share a common structure, and can be a good predictor of one another. It is also important to observe that the total variance explained (shared + private) by adaptation currents is lower than action potential attributes for both excitatory and inhibitory

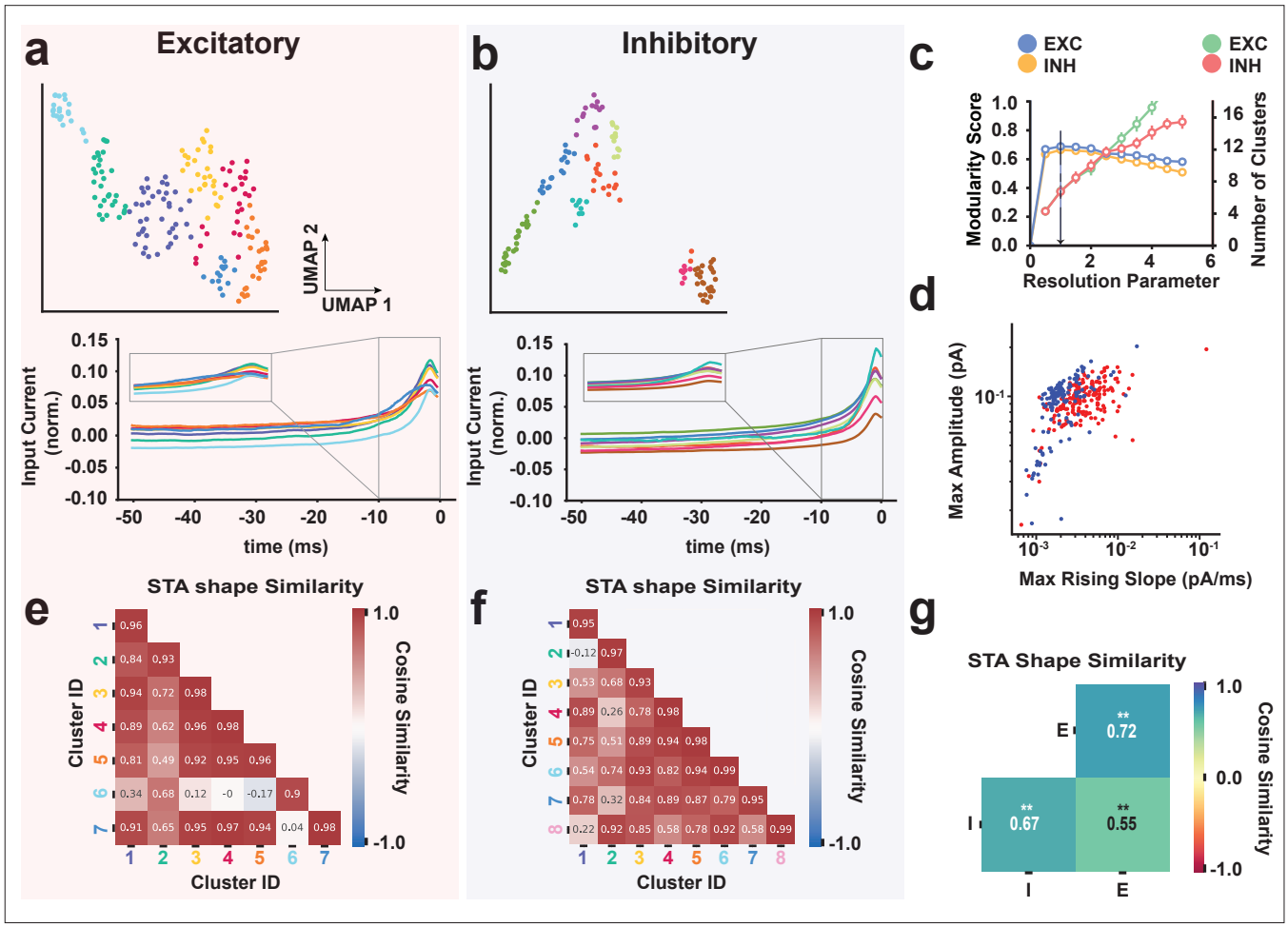


Figure 6: **Linear Input Filter measured using a Spike Triggered Average (STA)** (**a-b**) UMAP cluster representation for excitatory (left) and inhibitory (right) populations and the corresponding normalized average STA shape for each cluster in the 50 ms before a spike. Excitatory and inhibitory neuron populations were found to have 7 and 8 clusters respectively. (**c**) The chosen resolution parameters correspond to the highest modularity and the corresponding number of clusters. (**d**) Scatter plot showing the maximum slope vs maximum amplitude extracted from each STA (excitatory neurons depicted in red and inhibitory in blue). (**e-f**) Average cosine similarity across excitatory (e) and inhibitory (f) STA clusters. The excitatory clusters show a higher similarity with each other compared to inhibitory clusters. Each excitatory cluster on average shows high similarity ($>$0.9) with ≈ 2 other clusters, the inhibitory clusters on the other hand show a high similarity with ≈ 1 other cluster on average. (**g**) Heatmap showing the average cosine similarity for both within E/I and across E/I population STAs. The inhibitory population (I-I) shows a significantly higher heterogeneity (lower similarity) compared to the excitatory population (E-E). Welch's ANOVA shows a significant difference between the cosine similarity measures between excitatory (E), inhibitory (I), and across excitatory-inhibitory populations (ExI) F(2,48825)=1005.90, ***p= 0.0); Post-hoc Games-Howell test, E vs I (**p=0.001), I vs ExI (**p=0.0.001), and E vs ExI (**p=0.001).

populations, suggesting that action potential attributes are more informative about neuronal heterogeneity than adaptation currents.

DISCUSSION

In this work, we aimed to study the effect of input protocol (SH vs FN) on neuronal functional classification as well as to draw consensus on the attribute that are most informative about functional heterogeneity. For this, we first looked at the effect of the input protocol on neuronal classification based on two feature sets: waveforms and action potential features, while neurons were stimulated using two stimulus protocols: a static Step and Hold (SH) and a physiologically realistic Frozen Noise (FN) input. We found neuronal classification based on FN and SH protocols to be inconsistent when comparing the similarity in the cluster assignment in both input conditions using the co-cluster likelihood and cluster similarity indices (ARI and AMI). This highlights the importance of using a physiologically realistic input for studying functional heterogeneity of a neuronal population. We then aimed to determine which attribute(s) is/are the most informative about the neuronal functional heterogeneity when neurons are stimulated with a physiologically realistic FN input. We first explored the functional diversity of cells using 4 attribute sets (action potential, passive biophysical, adaptation current, and linear input filter (STA) attributes) for putative excitatory and inhibitory populations. To infer which attribute(set) is most informative about neuronal heterogeneity, we compared the private and shared variance across the four attribute sets. We found that the linear input filters (STAs) explain the most variance, especially private variance across the excitatory and inhibitory neurons, and thus, contains unique information about neuronal functional heterogeneity compared to other attributes and is the most informative about neuronal functional heterogeneity.

0.3 The role of input in neuronal functional diversity

In the present work, we found that neuronal functional classification changes as a function of the input the neurons receive, this point is still undiscussed in neuronal clustering literature. Our results fill an important gap in the literature, the effect of input on physiological classification, by showing that waveforms and electrophysiological attributes-based classifications are different when neurons are stimulared with static versus dynamic inputs. Previous studies that combined electrophysiological, morphological, and molecular attributes for neuronal classification ([15,16,18]), hypothesized that electrophysiological classification of a neuron can potentially change when using a different stimulus protocol. Our classification results confirm this hypothesis, by showing that clustering based on intracellular waveforms and action potential attributes are inconsistent across SH and FN-based inputs.

Our results showcase the importance of considering the functional dimension of neuronal identity that emerges from the interaction between input characteristics and intrinsic neuronal properties. These results do not only showcase how neurons cluster differently as a result of changing input protocol, which is an important and novel finding in the context of neuronal classification, but they also show that spiking dynamics, spike threshold, and AP height and width attributes were distinct between FN and SH protocols, as shown by the consistently low cosine similarities between FN and SH clusters and the radar plots comparing the spiking dynamics, spike threshold, and AP height and width attributes for each neuron. This is in agreement with earlier results of ([24,25,37]), which have already shown that firing intensity measures (AP count and AMPA conductance) are weakly correlated between static and dynamic stimuli (dynamic clamp) conditions. We extended these results by showing that the commonly studied action potential

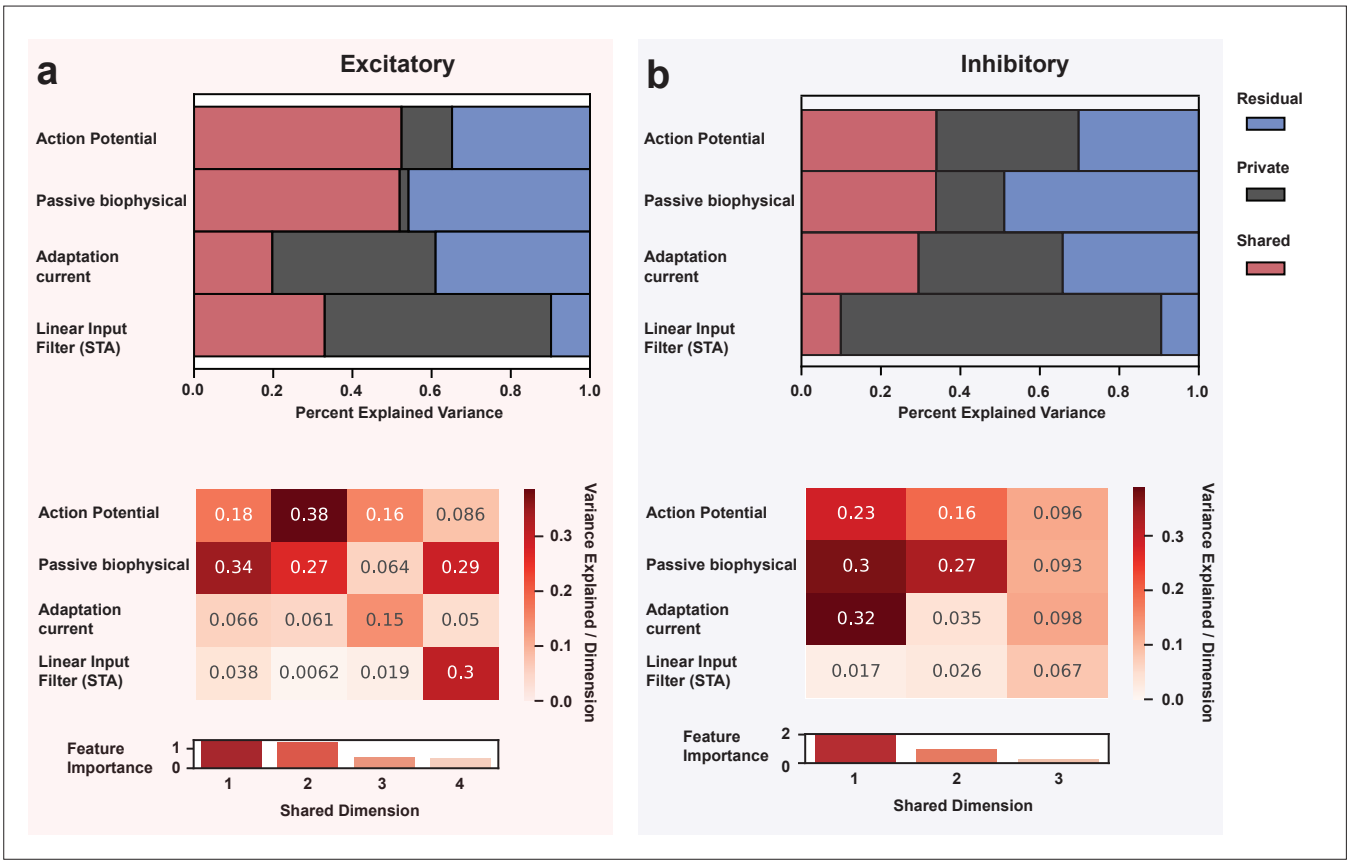


Figure 7: **MCFA analysis comparing shared and private variance explained by Action potential, Passive biophysical, Adaptation current, and Linear input filter (STA)**: (**a**) (top) Histogram showing the percentage of shared and private variance explained by each attribute for the excitatory population. action potential parameters (shared = 52.3%, private = 12.7%, residual = 34.8%) and passive biophysical parameters (shared = 51.9%, private = 2.2%, residual = 45.8%) explain most of the shared variance between attributes suggesting that these properties are not uniquely informative about population heterogeneity, on the other hand, linear input filter (STA) contains most of the unique information (shared = 32.9%, private = 57.2%, residual = 9.8%) about the excitatory population and the least amount of un-modelled residual, adaptation current explains more private variance (shared = 19.7%, private = 41.2%, residual = 38.9%) than action potential and passive biophysical features but less than linear input filter (STA). (middle) Heatmap showing the contribution of each shared dimension (total 4 based on selection criteria (methods)) in explaining variance across each attribute and their relative importance. (bottom) passive biophysical and action potential parameters have contributed to the most important shared dimensions. (**b**) (top) Histogram showing the percentage of shared and private variance explained by each attribute for the Inhibitory population. Action potential parameters (shared = 33.9%, private = 35.7 %, residual = 30.2%) and passive biophysical parameters (shared = 33.85%, private = 17.1%, residual = 48.9%) explain the highest amount of shared variance between attributes suggesting that these properties are not uniquely informative about population heterogeneity on the other hand, passive biophysical parameters have the most amount of un-modeled (residual) variance. STA contains most of the unique information (private variance) about the inhibitory population (shared = 9.9%, private = 80.6%, residual = 9.4%) and the least amount of un-modeled (residual) variance followed by the adaptation current (private = 29.4%, shared = 36.29%, residual = 34.2%). Heatmap (middle) showing the contribution of each shared dimension (total 3, based on selection criteria (methods)) in explaining variance across each attribute and their relative importance. (bottom) Most important shared dimensions have contributions from passive biophysical, action potential parameters, and adaptation current.

Based on this, a strong suggestion can be made for using linear input filter(STA) as the marker for heterogeneity in electrophysiological studies for both, excitatory and inhibitory populations. This demonstrates the importance of input in determining the difference across neural populations.

attributes span different latent manifolds as a result of input protocol (Appendix 0.7). Therefore, these classical action potential features do not encode a signature of the neurons that is invariant to the input, but are susceptible to the input the neurons receive. Since the conventionally used step and hold protocol does not represent an input a neuron receives in vivo, our findings advocate for using a physiologically realistic dynamic inputs such as the FN input for studying functional diversity. Our results therefore establish that input dynamics and its effects on functional classification and need to be considered even before molecular and morphological markers.

We used an unsupervised UMAP+Louvain clustering method in this study which has already been shown to distinguish neuronal classes based on extracellular waveforms ([28]) and also has been shown to improve upon classifications based on a low number of features extracted from waveforms. We capitalized on this idea by using UMAP+Louvain clustering to cluster SH and FN intracellular waveforms and action potential attributes and then compared the cluster assignments across the two protocols. Since the Louvain community method clusters on the high-dimensional graph structure provided by the UMAP algorithm, rather than projecting the data on a low dimensional space ([16,38]), which often leads to a loss of information about the latent structure ([38]), the clusters using the UMAP graphs were more robust than found either using a dimensional feature set or a low dimensional projection of the original high dimensional feature set. A second important advantage of using a non-linear dimensionality reduction technique like UMAP is that it allows for a manifold comparison between the two input conditions (see Appendix 0.7**(a-b)**).

0.4 Neuronal classification using frozen noise input-based attributes

We have established that neuronal functional classification is a result of the type of input that neurons receive, therefore a physiologically realistic input is important to understand neuronal functional diversity. A new question emerges: which attribute sets are the most informative about neuronal functional heterogeneity when neurons are presented with a dynamic stimuli? A similar question has been raised by ([8]). We attempted to provide a schema for answering this question. We did not intend to suggest a definitive number of classes of neurons, nor did we want to match our findings with previously established MET type classification, but we rather aimed to delineate a framework for drawing a consensus about which attribute(s) are the most informative about neuronal functional heterogeneity. An array of previous classification studies have relied upon variants of action potential and passive biophysical features ([39–42]), without any consensus about the informativeness of the attributes selected. Also, these studies have conventionally been reliant on feature extraction based on step stimulus protocols, which we demonstrate to produce cell classes that are different from when the neurons are stimulated with a dynamic input.

Although ([24,25]) provide a comparison between neurons responding to static and dynamic inputs, an active recommendation for which parameters are the most discriminatory was still missing. We supersede this by offering an alternative classification paradigm, based on the physiologically realistic frozen noise input. We tried to fill this gap by dividing the neuronal population into putative excitatory and inhibitory groups, then clustering neurons using the unsupervised UMAP+Louvain method, using four different sets of attributes separately. We found that neuronal clusters are subjective to feature selection. There was little consistency in comparing

cluster assignments across the four attribute sets (see Appendix 12**a-b**): the number of classes as well as the cluster assignments were found to be inconsistent across the four attribute sets.

We divided the neuronal population into putative excitatory and inhibitory classes based on the waveform shapes and firing rates, which is rather different from using molecular and morphological labels for E/I classification. Still, the goal of this study is not to align the neuronal identities to their ME-type markers, but rather to highlight the importance of input stimuli in functional classification. We show this by demonstrating that the firing properties of the narrow-width and broad-width neurons, which are conventionally categorized into inhibitory and excitatory classes respectively, change as a result of changing input type.

Parallel to investigating how putative E/I populations cluster based on different features, we also estimated the within-population variance for each feature set within E/I populations, which is representative of population heterogeneity based on a said feature. For action potential and passive biophysical attribute sets, we calculated the differences in the means of cosine similarity between E/I populations. We found that action potential attributes significantly differed between excitatory and inhibitory populations, consistent with previous findings ([19]). We also found that the action potential attributes for the inhibitory population was more heterogeneous than the excitatory population, which is also consistent with previous findings ([16,18]). It is important to highlight that the number of clusters found in our analysis is similar across excitatory and inhibitory populations, which might result from our dataset containing more excitatory than inhibitory neurons. The clustering based on the second feature set, that of passive biophysical properties extracted by fitting a GLIF model to the neural recordings, showed that passive parameters were also significantly different between excitatory and inhibitory neurons, but within-population heterogeneity was not significantly different between excitatory and inhibitory neurons. This suggests that passive parameters do not drive functional heterogeneity. This result is consistent with ([32]), which suggests that passive properties are not sufficiently discriminatory within E and I populations. Moreover, the authors claim that adaptive properties extracted using the GLIF model, such as the adaptation current, are more discriminatory than passive parameters. Our findings confirm this: adaptation currents show low similarity between excitatory and inhibitory populations, we found that the absolute maximum amplitude of all the adaptation current of the inhibitory neurons is smaller than that of the excitatory ones. We also found that inhibitory neurons have a significantly more heterogeneous adaptation profile than the excitatory population. This result provides data for designing and studying heterogeneous adaptive network models to further enhance our understanding of neural circuits' functional underpinnings.

Neurons can be functionally classified based on their input response features ([34,43,44]). We analyzed the linear input filters of neurons using the Spike Triggered Averages (STAs). This technique has been extensively used in studying the stimulus preference of neurons in the visual cortex ([45,46]). Our results show that the STA features are effective for estimating neuronal functional heterogeneity. We found that on average the STA shapes between the excitatory and inhibitory populations were different, based on comparing the average cosine similarity values between E and I populations, showing the difference in linear input filter between the two populations. We also found that the STA heterogeneity within the inhibitory population is significantly higher than the excitatory population. It has been shown ([47]) that visual cortical fast-spiking (putatively inhibitory) and regular-spiking (putatively excitatory) neurons have distinct levels of feature selectivity due to differences in passive biophysical attributes, such as the membrane time constant and input resistance. Since fast-spiking neurons have higher membrane leak conductance, therefore lower resistance it results in sharpening of neuronal selectivity to its preferred input as shown in previous studies such as ([48]). Conversely, regular spiking neurons were found to have lower conductance and thus, lower sensitivity to preferred stimuli. Previous computational studies have found that physiological and passive biophysical have degenerative relationships with STA kernels ([49]), this study has shown that the same STA shape can be achieved by multiple

difference values of the passive biophysical properties. This kind of degeneracy has been observed at the ion channel level as well ([50,51]). On the contrary, we observed a limited variability in biophysical features giving rise to a higher heterogeneity (comparing the within-population similarity matrices) in the STAs, i.e. in the functional linear input filter. We expect that the measured passive biophysical parameters from our data would be a good starting point to study the relationship between STA shape differences and the the range of passive biophysical attributes, this needs further computational modeling efforts. Since the diversity of feature preference by single neurons in the barrel cortex is not completely understood, a quantification of the functional heterogeneity observed in the linear input filter provided by our results is important for creating biophysically realistic models of cortical circuits and for a better understanding of circuit characteristics.

0.5 MCFA-based variance comparison across attribute sets

To determine which attribute set is most informative about neuronal functional heterogeneity, we compared the amount of neuronal population heterogeneity showcased by each attribute set by comparing the amount of private and shared variance explained by each attribute set. To our knowledge, the multi-attribute set comparison has not been done with physiological attributes before. This method provides a structured pathway to understand the limitations of commonly used electrophysiological features in cluster studies and helps to reach a consensus about the choice of attributes to be used for functional classification. We found that linear input filter explained the highest amount of private variance of all the attribute sets, for both excitatory and inhibitory neuronal populations. This is a clear indication of the usefulness of a linear input filter (STA) as an attribute to explore functional heterogeneity. Contrarily, we found that passive biophysical attributes and action potential attributes explain the most shared variance for excitatory and inhibitory populations, suggesting that these attributes are correlated and contain similar latent structures. This is an important result that can aid in the debate around selecting a feature that is the most informative about neuronal functional heterogeneity. We expect our approach to provide a framework for comparing heterogeneity across other brain regions as well.

0.6 Limitations

There are several limitations to our study. Firstly, the sample size for the shared FN and SH comparison set was 186 cells, which might not be enough to capture all variability across the somatosensory cortex layer 2/3. For the second part of the study, the sample size was 312 neurons. We expect an increased sample size would increase confidence in functional neuronal clusters. It would be insightful to sample from all layers of the barrel cortex and compare the linear input filter across layers, to gain a complete insight into the barrel cortical functional diversity. It has been argued that the activity of a large population of neurons that captures a certain behavior can be approximated by a low-dimensional representation from a few neurons ([52]) and therefore that the number of recorded neurons should depend on the neural task complexity. Even though the Frozen Noise is a better approximation of the synaptic input than a Step-and-Hold stimulus, it might not be representative of the full dimensionality of the input that a neuron receives in vivo, as the FN input is based on somatic current injection. We understand that a somatic current injection doesn't represent the full range of non-linear dendritic integration but it is quite helpful in estimating the linear input filter of a neuron which is found to be the most varying across neural populations and is sufficient for understanding the functional heterogeneity. We also lack the morphological and transcriptomic labels of the neurons, which makes the clusters found in this study incomparable to the commonly known types ([10]). We expect a dynamic clamp-like setup

using the frozen noise input along with morphological and transcriptomic labels to provide more clarity. Our study is also limited in providing a mechanistic relationship between the attribute sets we use, such as how the input changes the passive biophysical features and ultimately the linear input filter. Although the point GLIF model is quite helpful for extracting passive features, it does not provide a mechanistic description at ion channel level resolution for the variability in adaptation current and the feature selectivity. As explained above, we expect that a more detailed single neuron model, when studied with a physiologically realistic input, would provide a more elaborate picture of how the passive biophysical properties give rise to action potential and adaptation properties, which eventually result in the linear input filter. However, more detailed neuron models are difficult to fit and often do not produce unambiguous model properties ([53,54]).

0.7 Conclusion

In conclusion, we show that neuronal functional classification is a function of the input protocol and therefore, a physiologically realistic input should be preferred for functional classification. We also established that linear input filters are the most distinguishing property, compared to action potential, passive biophysical, and adaptation current attributes, for understanding the functional diversity of neurons when stimulated with a physiologically realistic input. These results provide an important recommendation for neural taxonomists and electrophysiologists: to consider a neuron's physiological input when defining neural identity. We expect computational single neuron as well as cortical network modeling efforts to discover the implications of the heterogeneity found in the 4 attribute sets of the excitatory and inhibitory neuronal populations we studied.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Fleur Zeldenrust (f.zeldenrust@neurphysiology.nl).

Materials availability

This study did not generate new materials.

Data and code availability

All analyzed current clamp and simulation data and the code to analyze and simulate them can be found in this repository: https://doi.org/10.34973/4f3k-1s63. The database is an extended version of the dataset ([22,55]). The code for all figures and analyses is available here: https://github.com/Nishant-codex/single_cell_analysis.git.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 860949.

AUTHOR CONTRIBUTIONS

Conceptualization, F.Z., T.C. and N.J.; methodology, N.J., T.C., and F.Z. ; analysis, N.J. ; writing–original draft, N.J. ; writing-–review & editing, N.J., S.V.D.B, T.C., and F.Z.; funding acquisition, F.Z. and T.C.; resources, T.C. and F.Z.; supervision, S.V.D.B, T.C., and F.Z.

DECLARATION OF INTERESTS

”The authors declare no competing interests.”

SUPPLEMENTAL INFORMATION INDEX

Figure S1. Waveform and Action Potential properties manifold comparison

Figure S2. Louvain vs Ensemble Clustering for Graphs (ECG) algorithm comparison

Figure S3. Cluster stability leaving one attribute out

Figure S4. Diversity of firing rate and AP half-width

Figure S5. STA Heterogeneity

Figure S6. Cluster assignment likelihood across attributes

A Waveform and Action Potential properties manifold comparison

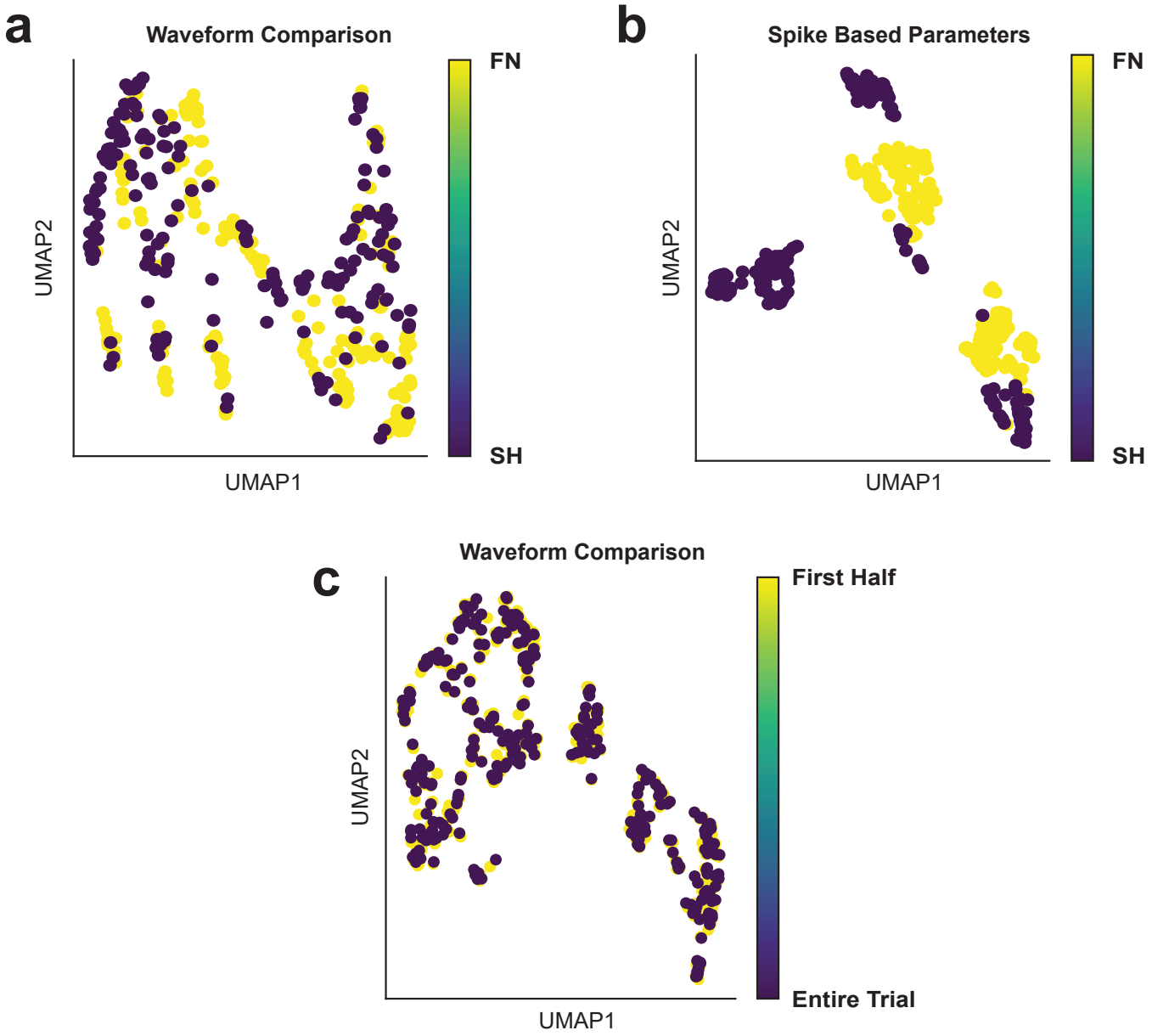


Figure 8: (**a**) Overlaid UMAP representation FN and SH waveforms from 186 neuron used in classification. The waveform shapes are different between SH and FN protocols. (**b**) FN and SH Action potential parameters used for classification projected together on the same space. The Action potential properties are different between FN and SH protocols. The SH Action potential properties show a bigger spread than FN properties. (**c**) UMAP projection of averaged Waveform comparison between the first half and the entire trial. The average waveform shape doesn't change as a result of trial length.

B Louvain vs Ensemble Clustering for Graphs (ECG) algorithm comparison

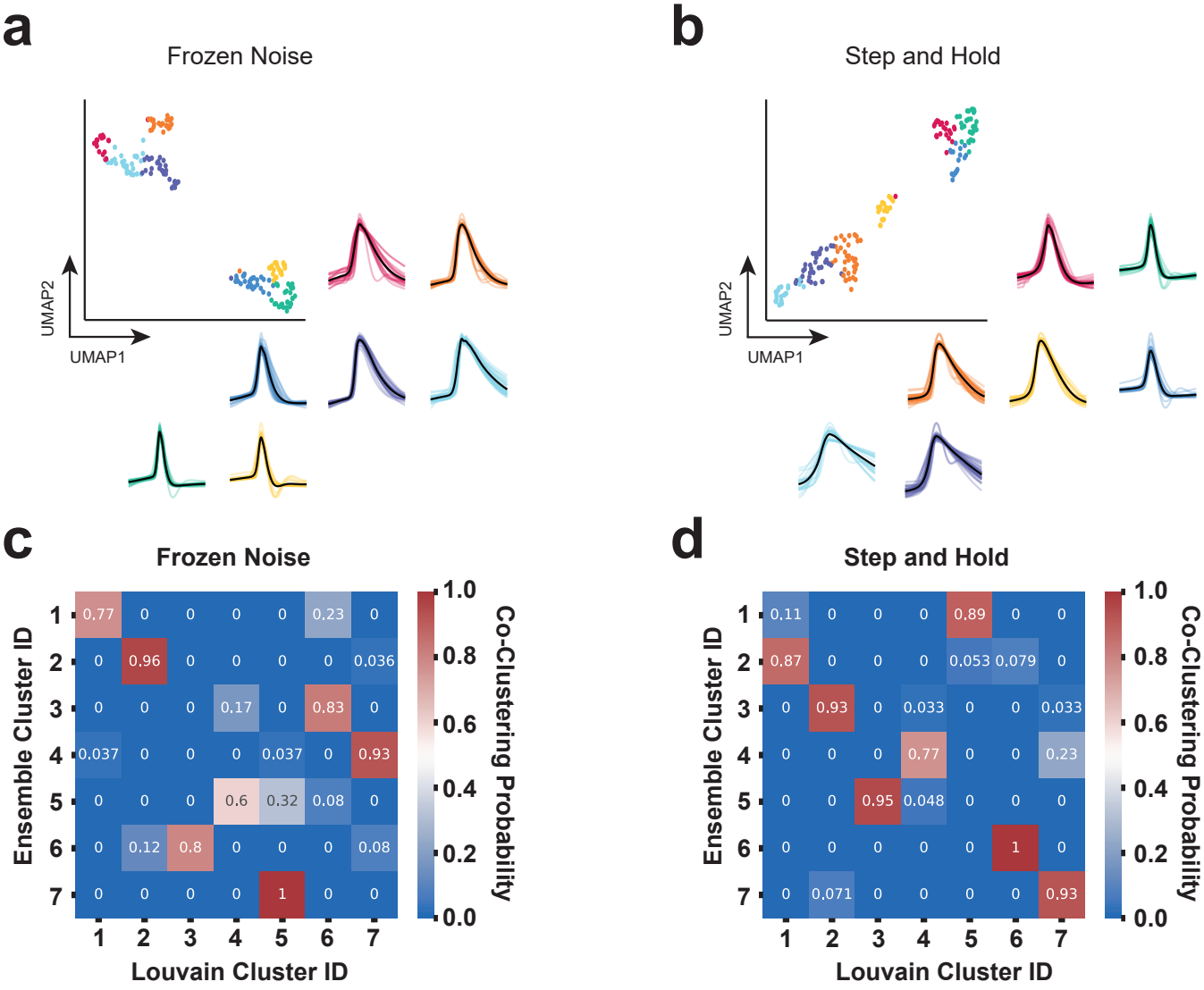


Figure 9: (**a**) UMAP embedding of FN waveforms colored with cluster labels found using Ensemble clustering, with original waveforms in the same color as the respective color. 7 clusters were observed, the same number as in the case of the Louvain community method. (**b**) UMAP embedding of SH waveforms colored with cluster labels found using Ensemble clustering, with original waveforms in the same color as the respective color. 7 clusters were observed, the same number as in the case of the Louvain community method. (**c**) Heatmap showing the correspondence between Louvain and Ensemble clustering for graph on FN waveforms. Clusters using the Louvain community detection algorithm show a high correspondence with the clusters obtained using Ensemble clustering for the graph method. (**d**) Heatmap showing the correspondence between Louvain and Ensemble clustering for graph on SH waveforms. Clusters using the Louvain community detection algorithm show a high correspondence with the clusters obtained using Ensemble clustering for the graph method.

C Cluster stability leaving one attribute out

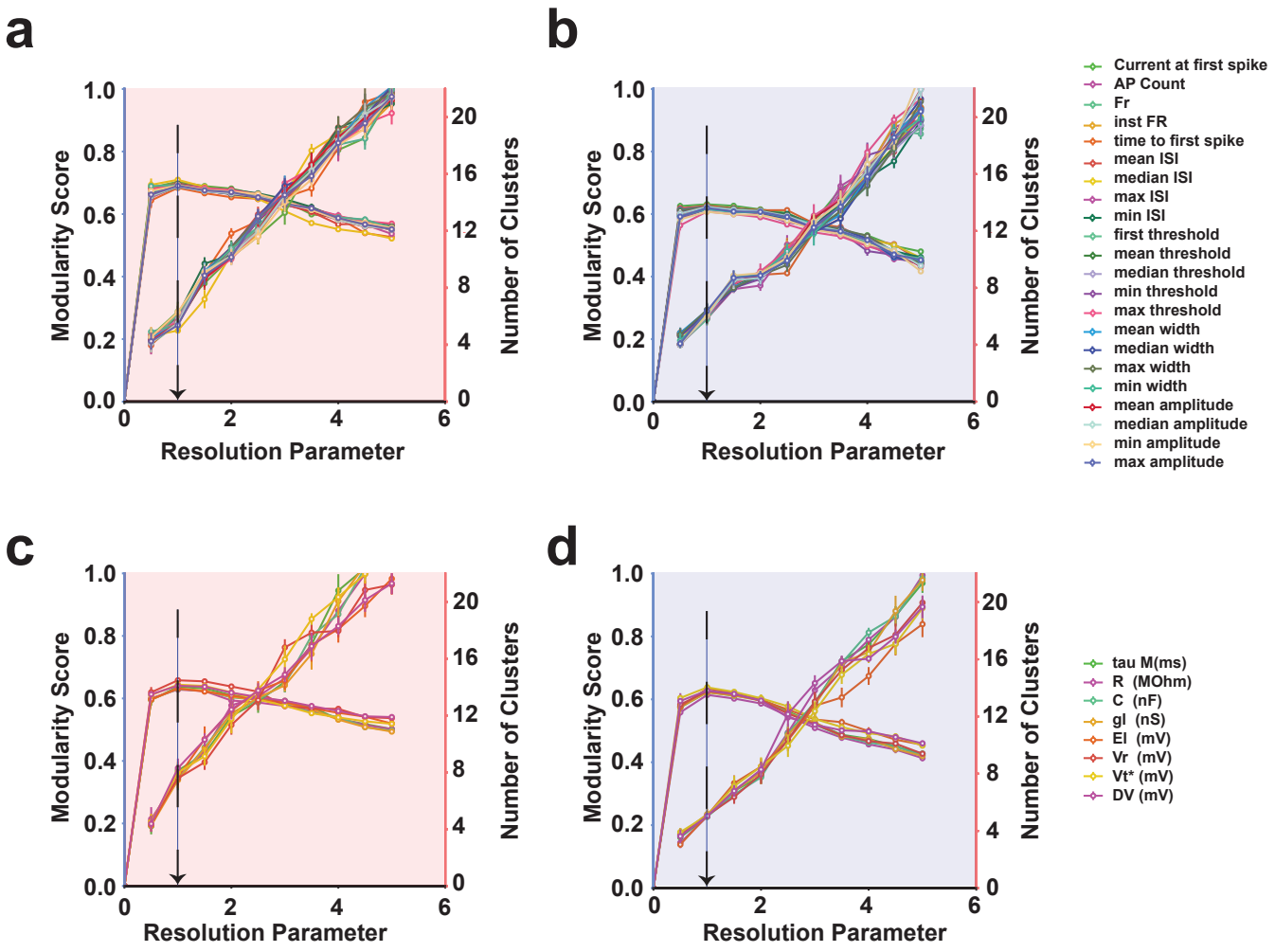


Figure 10: (**a-b**) Stability of clusters for action potential attributes leaving one attribute out for excitatory and Inhibitory sets. The clustering was performed 25 times with random 80% samples for resolution attributes ranging from 0.0 to 1, the mean and standard deviation of the resulting number of clusters and modularity score are plotted. For the chosen resolution parameter (1.0), the cluster numbers fluctuate between 5-7. The cluster number fluctuates more for the excitatory set (left) than the Inhibitory set (right) (**c-d**) Stability of clusters for biophysical attributes leaving one attribute out for excitatory and Inhibitory sets. The clustering was performed 25 times with random 90% samples for resolution parameters ranging from 0.0 to 1, the mean and standard deviation of the resulting number of clusters and modularity score are plotted. For the chosen resolution parameter (1.0), the excitatory population fluctuates between 6-8. On the contrary, the Inhibitory population doesn't show any fluctuation.

D Diversity of firing rate and AP half-width

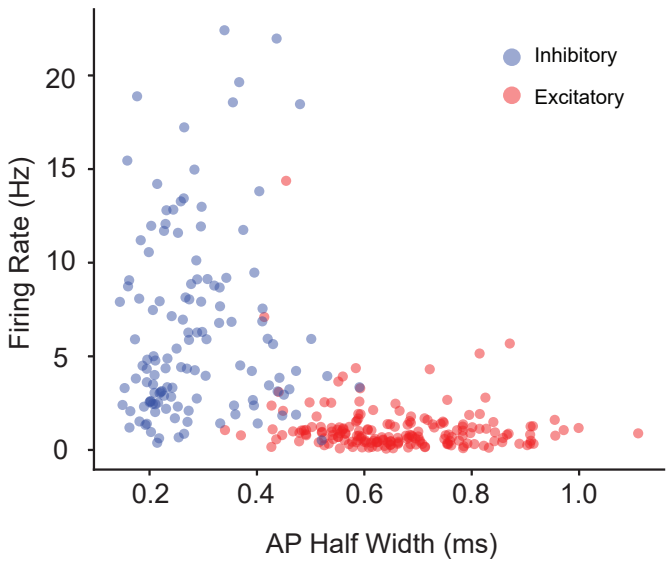


Figure 11: Firing rate vs AP half-width between excitatory (red) and inhibitory (blue) populations. The excitatory population has a lower firing rate and higher AP width. The Inhibitory population has a lower AP width and higher firing rate.

E STA Heterogeneity

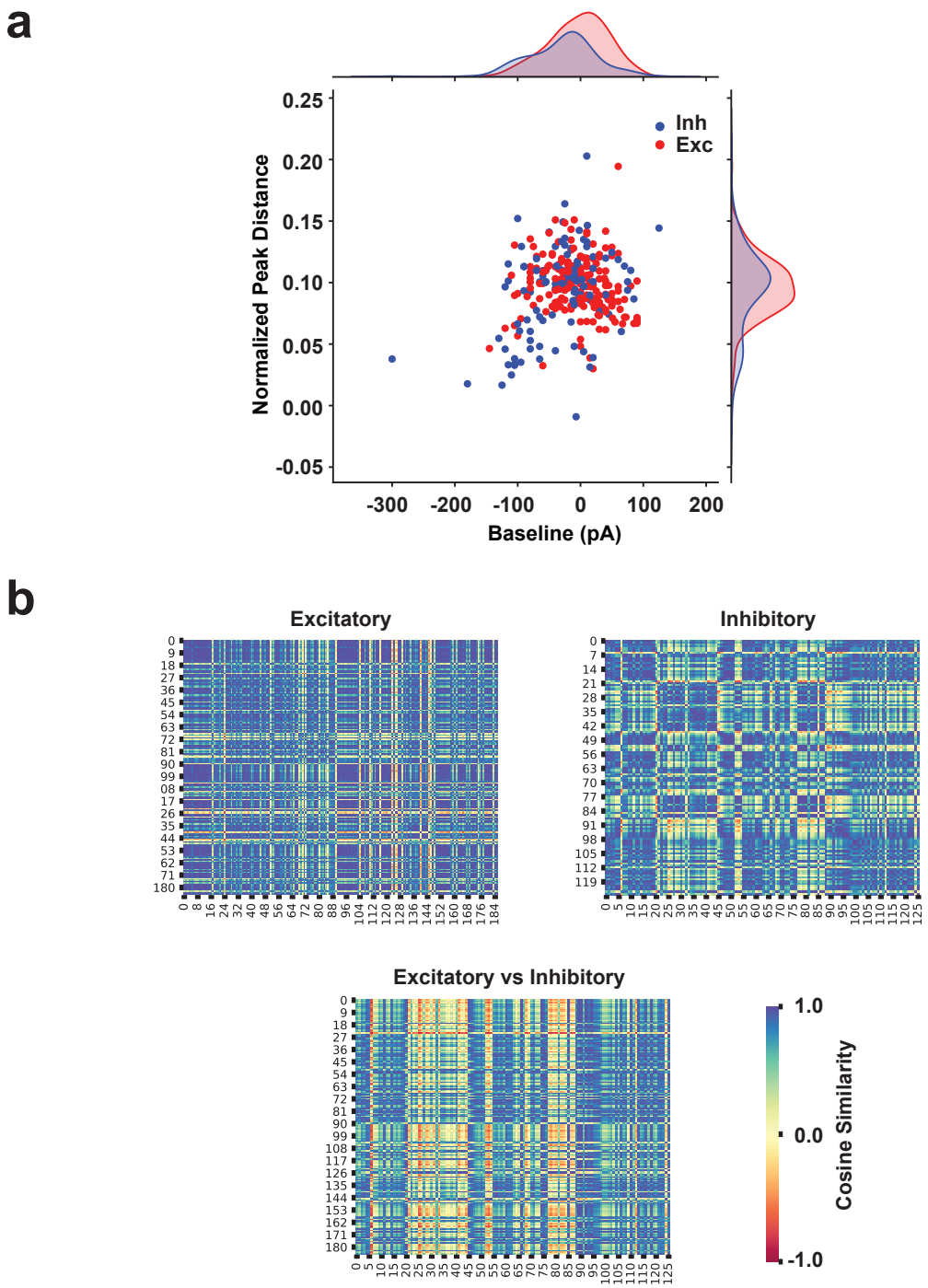


Figure 12: (**a**) The scatter plot shows the baseline added to the theoretical input vs the Peak distance of the STA. The excitatory population has a higher baseline and higher peak distance. The variance for both baseline and peak distance seems higher for the excitatory population than inhibitory population. (**b**) The STA cosine similarity is for the excitatory population (top left), the STA cosine similarity is for the inhibitory population (top right), and the STA cosine similarity is between the excitatory and inhibitory populations. The excitatory population has higher similarity than the inhibitory population.

F Cluster assignment likelihood across attributes

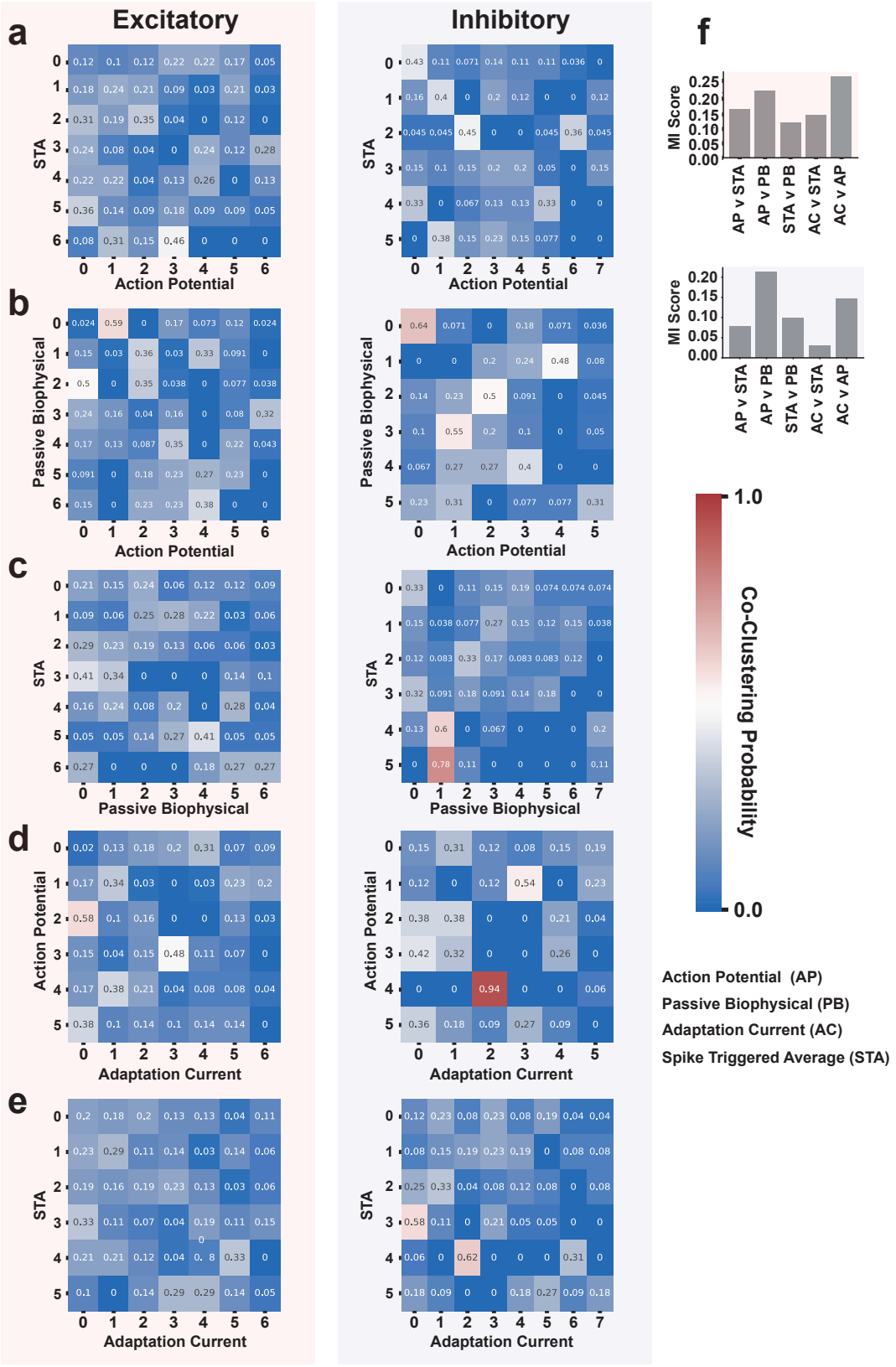


Figure 13: (**a-e**) Heatmap showing the likelihood for neuron clustering in attribute 1 clustering in one of the clusters in attribute 2 for excitatory (left) and inhibitory (right) population. For each attribute (Action potential, Passive biophysical, STA, Adaptation current (η), none of the clusters show a high likelihood to be clustering in another attribute. (f) Cluster label comparison between modularity pairs using Mutual information score for excitatory and inhibitory population. None of the pairs reach the modularity of 0.25 showing that neurons cluster differently across attribute sets.

References

1. Harris, K.D., and Shepherd, G.M. (2015). The neocortical circuit: themes and variations. Nature neuroscience *18*, 170–181.

2. Huang, Z.J., and Paul, A. (2019). The diversity of gabaergic neurons and neural communication elements. Nature Reviews Neuroscience *20*, 563–572.

3. Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. Nature reviews neuroscience *5*, 793–807.

4. Mukamel, E.A., and Ngai, J. (2019). Perspectives on defining cell types in the brain. Current opinion in neurobiology *56*, 61–68.

5. Fishell, G., and Heintz, N. (2013). The neuron identity problem: form meets function. Neuron *80*, 602–612.

6. Masland, R.H. (2012). The neuronal organization of the retina. Neuron *76*, 266–280.

7. Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S. et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature *563*, 72–78.

8. Zeng, H., and Sanes, J.R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. Nature Reviews Neuroscience *18*, 530–546.

9. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods *17*, 261–272. doi: `10.1038/s41592-019-0686-2`.

10. Zeng, H. (2022). What is a cell type and how to define it? Cell *185*, 2739–2755.

11. Selten, M., van Bokhoven, H., and Kasri, N.N. (2018). Inhibitory control of the excitatory/inhibitory balance in psychiatric disorders. F1000Research *7*.

12. Gast, R., Solla, S.A., and Kennedy, A. (2024). Neural heterogeneity controls computations in spiking neural networks. Proceedings of the National Academy of Sciences *121*, e2311885121.

13. Fuzik, J., Zeisel, A., Máté, Z., Calvigioni, D., Yanagawa, Y., Szabó, G., Linnarsson, S., and Harkany, T. (2016). Integration of electrophysiological recordings with single-cell rna-seq data identifies neuronal subtypes. Nature biotechnology *34*, 175–183.

14. Cadwell, C.R., Palasantza, A., Jiang, X., Berens, P., Deng, Q., Yilmaz, M., Reimer, J., Shen, S., Bethge, M., Tolias, K.F. et al. (2016). Electrophysiological, transcriptomic and morphologic profiling of single neurons using patch-seq. Nature biotechnology *34*, 199–203.

15. Scala, F., Kobak, D., Shan, S., Bernaerts, Y., Laturnus, S., Cadwell, C.R., Hartmanis, L., Froudarakis, E., Castro, J.R., Tan, Z.H. et al. (2019). Layer 4 of mouse neocortex differs in cell types and circuit organization between sensory areas. Nature communications *10*, 4174.

16. Gouwens, N.W., Sorensen, S.A., Berg, J., Lee, C., Jarsky, T., Ting, J., Sunkin, S.M., Feng, D., Anastassiou, C.A., Barkan, E. et al. (2019). Classification of electrophysiological and morphological neuron types in the mouse visual cortex. Nature neuroscience *22*, 1182–1195.

17. Scala, F., Kobak, D., Bernabucci, M., Bernaerts, Y., Cadwell, C.R., Castro, J.R., Hartmanis, L., Jiang, X., Laturnus, S., Miranda, E. et al. (2021). Phenotypic variation of transcriptomic cell types in mouse motor cortex. Nature *598*, 144–150.

18. Gouwens, N.W., Sorensen, S.A., Baftizadeh, F., Budzillo, A., Lee, B.R., Jarsky, T., Alfiler, L., Baker, K., Barkan, E., Berry, K. et al. (2020). Integrated morphoelectric and transcriptomic classification of cortical gabaergic cells. Cell *183*, 935–953.

19. Connors, B.W., and Gutnick, M.J. (1990). Intrinsic firing patterns of diverse neocortical neurons. Trends in neurosciences *13*, 99–104.

20. Steriade, M. (2000). Corticothalamic resonance, states of vigilance and mentation. Neuroscience *101*, 243–276.

21. Zeldenrust, F., de Knecht, S., Wadman, W.J., Denève, S., and Gutkin, B. (2017). Estimating the information extracted by a single spiking neuron from a continuous input time series. Frontiers in computational neuroscience *11*, 49.

22. da Silva Lantyer, A., Calcini, N., Bijlsma, A., Kole, K., Emmelkamp, M., Peeters, M., Scheenen, W.J., Zeldenrust, F., and Celikel, T. (2018). A databank for intracellular electrophysiological mapping of the adult somatosensory cortex. GigaScience *7*, giy147.

23. Yan, X., Calcini, N., Safavi, P., Ak, A., Kole, K., Zeldenrust, F., and Celikel, T. (2022). A whole-cell recording database of neuromodulatory action in the adult neocortex. bioRxiv pp. 2022–01.

24. Hernáth, F., Schlett, K., and Szücs, A. (2019). Alternative classifications of neurons based on physiological properties and synaptic responses, a computational study. Scientific reports *9*, 13096.

25. Szabó, A., Schlett, K., and Szücs, A. (2021). Conventional measures of intrinsic excitability are poor estimators of neuronal activity under realistic synaptic inputs. PLOS Computational Biology *17*, e1009378.

26. Avermann, M., Tomm, C., Mateo, C., Gerstner, W., and Petersen, C.C. (2012). Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. Journal of neurophysiology *107*, 3116–3134.

27. Trainito, C., von Nicolai, C., Miller, E.K., and Siegel, M. (2019). Extracellular spike waveform dissociates four functionally distinct cell classes in primate cortex. Current Biology *29*, 2973–2982.

28. Lee, E.K., Balasubramanian, H., Tsolias, A., Anakwe, S.U., Medalla, M., Shenoy, K.V., and Chandrasekaran, C. (2021). Non-linear dimensionality reduction on extracellular waveforms reveals cell type diversity in premotor cortex. Elife *10*, e67490.

29. Bean, B.P. (2007). The action potential in mammalian central neurons. Nature Reviews Neuroscience *8*, 451–465.

30. Kiritani, T., Pala, A., Gasselin, C., Crochet, S., and Petersen, C.C. (2023). Membrane potential dynamics of excitatory and inhibitory neurons in mouse barrel cortex during active whisker sensing. Plos one *18*, e0287174.

31. Heiss, J.E., Katz, Y., Ganmor, E., and Lampl, I. (2008). Shift in the balance between excitation and inhibition during sensory adaptation of s1 neurons. Journal of Neuroscience *28*, 13320–13330.

32. Mensi, S., Naud, R., Pozzorini, C., Avermann, M., Petersen, C.C., and Gerstner, W. (2012). Parameter extraction and classification of three cortical neuron types reveals two distinct adaptation mechanisms. Journal of neurophysiology *107*, 1756–1775.

33. Pozzorini, C., Mensi, S., Hagens, O., Naud, R., Koch, C., and Gerstner, W. (2015). Automated high-throughput characterization of single neurons by means of simplified spiking models. PLoS computational biology *11*, e1004275.

34. Sharpee, T.O. (2014). Toward functional classification of neuronal types. Neuron *83*, 1329–1334.

35. Schwartz, O., Pillow, J.W., Rust, N.C., and Simoncelli, E.P. (2006). Spike-triggered neural characterization. Journal of vision *6*, 13–13.

36. Brown, B.C., Wang, C., Kasela, S., Aguet, F., Nachun, D.C., Taylor, K.D., Tracy, R.P., Durda, P., Liu, Y., Johnson, W.C. et al. (2023). Multiset correlation and factor analysis enables exploration of multi-omics data. Cell Genomics *3*.

37. Szücs, A., and Huerta, R. (2015). Differential effects of static and dynamic inputs on neuronal excitability. Journal of Neurophysiology *113*, 232–243.

38. Druckmann, S., Hill, S., Schürmann, F., Markram, H., and Segev, I. (2013). A hierarchical structure of cortical interneuron electrical diversity revealed by automated statistical analysis. Cerebral Cortex *23*, 2994–3006.

39. Contreras, D. (2004). Electrophysiological classes of neocortical neurons. Neural Networks *17*, 633–646.

40. Halabisky, B., Shen, F., Huguenard, J.R., and Prince, D.A. (2006). Electrophysiological classification of somatostatin-positive interneurons in mouse sensorimotor cortex. Journal of neurophysiology *96*, 834–845.

41. McGarry, L.M., Packer, A.M., Fino, E., Nikolenko, V., Sippy, T., and Yuste, R. (2010). Quantitative classification of somatostatin-positive neocortical interneurons identifies three interneuron subtypes. Frontiers in neural circuits *4*, 1284.

42. Casale, A.E., Foust, A.J., Bal, T., and McCormick, D.A. (2015). Cortical interneuron subtypes vary in their axonal action potential properties. Journal of Neuroscience *35*, 15555–15567.

43. Famulare, M., and Fairhall, A. (2010). Feature selection in simple neurons: how coding depends on spiking dynamics. Neural computation *22*, 581–598.

44. Chéreau, R., Bawa, T., Fodoulian, L., Carleton, A., Pagès, S., and Holtmaat, A. (2020). Dynamic perceptual feature selectivity in primary somatosensory cortex upon reversal learning. Nature communications *11*, 3245.

45. Chichilnisky, E. (2001). A simple white noise analysis of neuronal lightresponses. Network: computation in neural systems *12*, 199.

46. Rathbun, D., Ghorbani, N., Shabani, H., Zrenner, E., and Hosseinzadeh, Z. (2018). Spike-triggered average electrical stimuli as input filters for bionic vision—a perspective. Journal of neural engineering *15*, 063002.

47. Cardin, J.A., Palmer, L.A., and Contreras, D. (2007). Stimulus feature selectivity in excitatory and inhibitory neurons in primary visual cortex. Journal of Neuroscience *27*, 10333–10344.

48. Li, B., Routh, B.N., Johnston, D., Seidemann, E., and Priebe, N.J. (2020). Voltage-gated intrinsic conductances shape the input-output relationship of cortical neurons in behaving primate v1. Neuron *107*, 185–196.

49. Jain, A., and Narayanan, R. (2020). Degeneracy in the emergence of spike-triggered average of hippocampal pyramidal neurons. Scientific reports *10*, 374.

50. Prinz, A.A., Bucher, D., and Marder, E. (2004). Similar network activity from disparate circuit parameters. Nature neuroscience *7*, 1345–1352.

51. Marder, E., and Taylor, A.L. (2011). Multiple models to capture the variability in biological neurons and networks. Nature neuroscience *14*, 133–138.

52. Gao, P., and Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. Current opinion in neurobiology *32*, 148–155.

53. Gonçalves, P.J., Lueckmann, J.M., Deistler, M., Nonnenmacher, M., Öcal, K., Bassetto, G., Chintaluri, C., Podlaski, W.F., Haddad, S.A., Vogels, T.P. et al. (2020). Training deep neural density estimators to identify mechanistic models of neural dynamics. Elife *9*, e56261.

54. Nandi, A., Chartrand, T., Van Geit, W., Buchin, A., Yao, Z., Lee, S.Y., Wei, Y., Kalmbach, B., Lee, B., Lein, E. et al. (2022). Single-neuron models linking electrophysiology, morphology, and transcriptomics across cortical cell types. Cell reports *40*.

55. Zeldenrust, F.F., Calcini, N., Yan, X., Bijlsma, A., Kole, K., and Celikel, T.T. (2024). Cell type specific information processing in inhibitory and excitatory neurons. Radboud University. URL: `https://doi.org/10.34973/4f3k-1s63`. doi: `10.34973/4f3k-1s63`.

56. Kole, K., Zhang, Y., Jansen, E.J., Brouns, T., Bijlsma, A., Calcini, N., Yan, X., Lantyer, A.d.S., and Celikel, T. (2020). Assessing the utility of magneto to control neuronal excitability in the somatosensory cortex. Nature neuroscience *23*, 1044–1046.

57. Kole, K., and Celikel, T. (2019). Neocortical microdissection at columnar and laminar resolution for molecular interrogation. Current Protocols in Neuroscience *86*, e55.

58. (2017). Proteomic landscape of the primary somatosensory cortex upon sensory deprivation. Gigascience *6*, gix082.

59. Miceli, S., Nadif Kasri, N., Joosten, J., Huang, C., Kepser, L., Proville, R., Selten, M.M., van Eijs, F., Azarfar, A., Homberg, J.R. et al. (2017). Reduced inhibition within layer iv of sert knockout rat barrel cortex is associated with faster sensory integration. Cerebral Cortex *27*, 933–949.

60. Martens, M.B., Celikel, T., and Tiesinga, P.H. (2015). A developmental switch for hebbian plasticity. PLoS Computational Biology *11*, e1004386.

61. Fontaine, B., Peña, J.L., and Brette, R. (2014). Spike-threshold adaptation predicted by membrane potential dynamics in vivo. PLoS computational biology *10*, e1003560.

62. Gerstner, W., Kistler, W.M., Naud, R., and Paninski, L. (2014). Neuronal dynamics: From single neurons to networks and models of cognition. Cambridge University Press.

63. Aggarwal, C.C., Hinneburg, A., and Keim, D.A. (2001). On the surprising behavior of distance metrics in high dimensional space. In Database theory—ICDT 2001: 8th international conference London, UK, January 4–6, 2001 proceedings 8. Springer pp. 420–434.

64. McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.

65. Blondel, V.D., Guillaume, J.L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment *2008*, P10008.

66. Poulin, V., and Théberge, F. (2019). Ensemble clustering for graphs. In Complex Networks and Their Applications VII: Volume 1 Proceedings The 7th International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2018 7. Springer pp. 231–243.

67. McInnes, L., Healy, J., Saul, N., and Grossberger, L. (2018). Umap: Uniform manifold approximation and projection. The Journal of Open Source Software *3*, 861.

68. Bonald, T., de Lara, N., Lutz, Q., and Charpentier, B. (2020). Scikit-network: Graph analysis in python. Journal of Machine Learning Research *21*, 1–6. URL: `http://jmlr.org/papers/v21/20-412.html`.

69. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research *12*, 2825–2830.

70. Perktold, J., Seabold, S., Sheppard, K., ChadFulton, Shedden, K., jbrockmendel, j grana6, Quackenbush, P., Arel-Bundock, V., McKinney, W., Langmore, I., Baker, B., Gommers, R., yogabonito, s scherrer, Zhurko, Y., Brett, M., Giampieri, E., yl565, Millman, J., Hobson, P., Vincent, Roy, P., Augspurger, T., tvanzyl, alexbrc, Hartley, T., Perez, F., Tamiya, Y., and Halchenko, Y. (2024). statsmodels/statsmodels: Release 0.14.2. Zenodo. URL: `https://doi.org/10.5281/zenodo.10984387`. doi: `10.5281/zenodo.10984387`.

71. Brown, B., and collinwa (2023). collinwa/mcfa: First pypi release. Zenodo. URL: `https://doi.org/10.5281/zenodo.8128339`. doi: `10.5281/zenodo.8128339`.

STAR METHODS

Key resources table

Experimental model and study participant details

Ethics statement The data used in this research was previously published and made freely available to the community ([22] and[23]). All the experimental work, as outlined in the cited articles, were carried out in compliance with the European directive 2010/63/EU, the national regulations of the Netherlands, and international standards for animal care and use.

Slice electrophysiology Data acquisition procedures, the details of the in vitro slice preparation, intracellular access to anatomically targeted neurons, data digitization, and preprocessing have been described in detail elsewhere ([22,56–59]). In short, Pvalbtm1(cre)Arbr (RRID:MGI:5315557) or Ssttm2.1(cre)Zjh/J (RRID:IMSR_JAX: 013044) mice, including both females and males, were obtained from local breeding colonies and studied after the maturation of evoked neurotransmitter release in the primary somatosensory cortex ([60]).

Mice were anesthetized with Isoflurane (1.5 mL/mouse) before extracting tissue, and coronal slices of the primary somatosensory cortex (barrel subfield) were prepared. The brain was removed, and 300 µm-thick coronal slices were made. Slices were then incubated in artificial cerebrospinal fluid (aCSF) (120 mM NaCl, 3.5 KCl, 10 glucose, etc.), aerated with 95% O_2/5% CO_2 at 37℃, and then at room temperature after 30 minutes.

Whole-cell electrophysiological recordings were performed with continuously oxygenated aCSF. The barrel cortex was localized, and cells in the supragranular layers were patched under 40x magnification using HEKA EPC 9 and EPC10 amplifiers with Patch Master software. Patch-clamp electrodes were pulled from glass capillaries (1.00 mm external diameter, 0.50 mm internal diameter) and used with 5–10 MOhm resistance, filled with intracellular solution (130 mM K-Gluconate, 5 KCl, 1.5 $MgCl_2$, etc., pH adjusted to 7.22 with KOH). Data were band-pass filtered at 0.1–3000 Hz before storage for offline analysis.

Step and hold (SH) protocol The Step and Hold protocol was set up in a current clamp configuration, the resting membrane potential was set to -70 mV before current injection into the soma of the neuron. In total, 10 current step injections, each 500 ms long, were performed. The steps ranged from 40 pA to 400 pA with an inter-sweep interval of 6.5s. The stimulus was repeated 1 to 3 times for neurons. The drifts encountered were not corrected for.

Frozen Noise (FN) protocol The Frozen Noise input protocol consisted of injecting a somatic current that is the result of an artificial neural network of 1000 neurons responding (firing Poisson spikes) to random stimuli i.e., the hidden state, the membrane potential response to the somatic input is recorded with a sampling rate of 20 kHz for a total length of 360 seconds and saved. Each raw data file consisted of a vehicle control trial (artificial Cerebrospinal fluid i.e. aCSf) and a drug trial (a specific neuromodulatory receptor agonist or antagonist was added to the bath and the recording was repeated). Some files consisted of multiple control and drug trials. See ([21,22]) for more details.

Method details

Feature Extraction

Feature extraction is performed separately for SH and FN protocols to classify neurons using different inputs and to determine cell class with a realistic stimulus. For the first part, which is

comparing classification across protocols, we collect waveforms and Action potential features for comparing SH and FN protocols using a subset of the data, (186 neurons). As not all cells were recorded with both types of input protocols, this subset was chosen to match the cell IDs across the two protocols. For the second part, which compares different physiological attributes for the FN protocol, we collect waveforms, spike event-based features, biophysical features, and the STA for each control (aCSF) trial for each cell in the FN dataset. We discard trials that contain recording artifacts (observed distortions or high-frequency noise), in total 11 neurons. Since there were multiple drug and control trials available for each neuron, we always took the first control trial for this study as this prevented any residual effect from the drug condition after washing. The total number of neurons used for this part of the study was 312.

Spike waveforms Spike waveforms were extracted for each neuron from both SH and FN protocol trials (this could not be done for all the neurons as some neuron IDs could not be matched between SH and FN protocols due to missing metadata). Firstly, we identified the spike times in each trial in each protocol. This was done by identifying peaks in the membrane potential trace using the Scipy Findpeaks function[9] with $height = 20mV$, and $distance = 80ms$ as the chosen hyperparameters. Next, we cut the spike waveform by defining a window of 2 ms before and 3 ms after each spike peak to get a 5 ms long waveform for each spike. We ignored the spikes that had an ISI lower than 3 ms as we were interested in non-bursting type spike shapes. We then average over all the waveforms for each trial to get an average waveform shape in both FN and SH protocols respectively.
Extracted Waveforms for the second part were 10 ms long (5 ms before and after the spike peak), to incorporate the subthreshold dynamic before reaching the threshold of the neuron in the waveform classification as well. We observed some variability in the baseline membrane potential values as well as the slope of membrane potential before the threshold value was reached.

Action potential features Action potential features were extracted to study the variability in spike-related dynamics across individual neurons. These features were designed based on their suitability for comparing neurons in FN and SH protocols. These Action potential features were divided into three categories: Spiking Dynamics, spike threshold, and Action Potential height and width.

Spiking Dynamics

This includes the following features:

- **Current at the first spike**: We take the current amplitude when the membrane potential crosses the threshold for the first time in the trial for FN protocol. For SH protocol, it is the current step that produces the spike for the first time.
- **AP count**: We count the number of spikes over the entire trial length of 360 seconds for the FN trial and 500 milliseconds for the SH trial which is the duration for the current onset.
- Inter spike interval: Inter spike interval was measured as the time interval between two spikes in milliseconds. We calculate mean, median, maximum, and minimum values for each trial in SH and FN protocols. For the SH case, we measure the ISI for the maximum current amplitude (400 pA) trial.
- **Time to first spike**: We measure the time (in milliseconds) it takes for the neuron to fire the first action potential. For the SH case, we take the lowest amplitude step where a spike is observed.

- **Firing rate**: We calculate the firing rate as the number of spikes per second. For the FN case, we take the entire length of 360 seconds of the trial and for the SH case, we take the duration of the current onset which was 500 milliseconds for the highest current step (400 pA).

$$fr = \frac{N_{spikes}}{T} \tag{1}$$

  where N_{spikes} is the total number of spikes in the trial and T is the total length of the trial.

- **Interspike Interval**: Interspike interval is the duration between two spikes $t_{spike_{n+1}} - t_{spike_n}$, where $t_{spike_{t+1}}$ and t_{spike_t} are spike times for spikes n+1 and n. We take the interspike interval for all the spikes in the trial for both FN and SH protocols and calculate the mean, median, minimum, and maximum values.

- **Instantaneous rate**: Instantaneous rate is defined as the reciprocal of the average of the Interspike interval.

$$inst.firing_rate = \frac{1}{t_{spike_{n+1}} - t_{spike_n}} \tag{2}$$

Spike threshold

- We measure the threshold of each spike from the trial as described in[61]. The threshold is defined as the voltage V at the spike onset when the first derivative of the membrane potential dV/dt reaches 25 mV/ms for the first time. We take the first threshold value of the trial as well as the mean, median, maximum, and minimum values of the thresholds from the entire trial length.

Action Potential height and width

- **Width**: Spike width is calculated as the time it takes between when the membrane potential reaches the AP threshold and when the membrane voltage goes below the threshold after the spike peak. We calculate the width for each spike in the trial for both FN and SH protocols. We calculate the mean, median, maximum, and minimum for all the values obtained from a trial.

- **Amplitude**: Spike amplitude is calculated as the difference between AP peak and AP threshold value. We calculate the amplitude for each spike in the trial for both FN and SH protocols. We calculate the mean, median, maximum, and minimum for all the values obtained from a trial.

Biophysical Feature extraction using GLIF model

Since it is not possible to empirically observe the biophysical properties of the cell just using the membrane potential, we fit a Generalized Leaky Integrate and fire model (GLIF) to the recordings, which can capture universal spiking and sub-threshold dynamics ([33]). The following equations define the GLIF model:

$$C\dot{V}(t) = -g_L(V(t) - E_L) - \sum_{\hat{t_j}<t} \eta(t - \hat{t}_j) + I(t), \tag{3}$$

where $V(t)$ is the membrane potential, C is the membrane capacitance, gL is the leak conductance, EL is the resting potential, and $\eta(t)$ is the adaptation current triggered by a spike event. Spikes are stochastically produced by a point process that represents conditional firing intensity $\lambda(t|V, V_t)$ that is dependent on the instantaneous difference between the membrane potential and voltage threshold given by:

$$\lambda(t|V, V_t) = \lambda_0 * exp\left(\frac{V(t) - V_T(t)}{\Delta V}\right), \tag{4}$$

where λ_0 is the base firing rate in Hz, V(t) is the membrane potential, and V_T(t) is the moving spike threshold and ΔV controls the sharpness of the exponential threshold function. The probability of a spike $\hat{t}$ between a time interval t and Δt is given by the following equation (based on[62]):

$$P(\hat{t} \in [t, t + \Delta t]) = 1 - exp(-\int_t^{t+\Delta t} \lambda(s)ds) \approx 1 - exp\Big(- \lambda(t)\Delta t\Big), \tag{5}$$

The dynamics of the firing threshold $V_T(t)$ are given by:

$$V_T(t) = V_T^* + \sum_{\hat{t}_j < t} \gamma(t - \hat{t}_j), \tag{6}$$

where γ is the stereotypical movement of the spike threshold after a spike and V_T^* is the threshold baseline.

The method for fitting this neuron model to a membrane potential recording is divided into the following steps:

Preparation step: A 100-second window from the initial part of the trial is taken as the training set for the fitting, ([33]) shows that a longer trial length doesn't improve the fit. Spike times and waveforms are also extracted for the preparation step of the fitting procedure.

Step 1. Fitting the reset voltage: The waveforms extracted in the preparation step were averaged, and then the Reset voltage V_{reset} was extracted using the averaged waveform by setting an arbitrary refractory period t_{ref} and taking the membrane potential value at $t_{\hat{i}} + t_{ref}$, where $t_{\hat{i}}$ is the spike peak. The refractory period is chosen to be always lower than the minimum inter-spike interval, we chose the refractory period of $t_{ref} = 4ms$ in this case.

Step 2. Fitting sub-threshold dynamics : The voltage dynamics in eq (3) are given by parameter set θ_{sub} ={C, g_L, η, and E_L}, by fitting the temporal derivative of the data $\dot{V}_{data}$ in the model, we can extract the set of passive parameter set θ_{sub} for the data. Firstly, we can write the adaptation current η as a linear sum of rectangular basis functions ([32]).

$$\eta(t) = \sum_{k=1}^{K} a_k f^{(k)}(t), \tag{7}$$

Using the fact that the voltage dynamics are approximately linear in the subthreshold regime, θ_{sub} parameter set can be extracted using a multi-linear regression between $\dot{V}_{data}$ and $\dot{V}_{model}$. For this, we created a training set V_{data}^{sub} where we removed the spike waveforms

from V_{data}, $V_{data}^{sub} = \{V_{data}(t)|t \notin (t_{\hat{i}} - 5ms, t_{\hat{i}} + t_{ref})\}$, where $t_{\hat{i}}$ is the spike times. The regression problem can be stated as:

$$\theta_{sub} = (X^T X)^{-1} X^T \dot{V}_{data}^{sub}, \tag{8}$$

where X^T is a matrix representing parameter values at different time points, its row elements x_t^T are of the following form

$$x_t^T = [V_{data}^{sub}, 1, f^{(1)}(t), f^{(2)}(t), ..., f^{(K)}(t)], \tag{9}$$

Step 3. Fitting the spike probability: For fitting the spiking probability to the data, we need to extract parameters defining the dynamics of the threshold eq (6). The stereotypical shape of the adaptation current threshold movement can be expanded as a sum of rectangular basis functions as follows ([32]):

$$\gamma(t) = \sum_{p=1}^{P} \gamma_p f^{(p)}(t), \tag{10}$$

We use the parameters obtained in the previous steps to compute the subthreshold membrane potential of the model using numerical integration of eq (3). We set $\lambda_0 = 1$ Hz and all the threshold parameters $\theta_{thr} = \{\Delta V$, V_T^* , and $\gamma(t)$ $\}$ are extracted by maximizing the likelihood function of the following form based on the experimental spike train:

$$\hat{\theta}_{thr} = \underset{\theta_{thr}}{\operatorname{argmax}} \left[\sum_{t \in \hat{\{t_j\}}} y_t^T \theta_{thr} - \Delta T \sum_{t \in \Omega} exp(y_t^T \theta_{thr}) \right], \tag{11}$$

Where Ω = {t—t $\notin (\hat{t}_j, \hat{t}_j + t_{ref})$},

The subthreshold fit is examined by comparing the variance explained R^2 of the subthreshold membrane potential trace V between the data and the model. All the models chosen for clustering had an R^2 value $>$ 0.7. The sets $\theta_{clustering}$ = {g_L, ΔV, C, V_T^*, E_L, V_{reset}} are the parameters that are extracted from the model that is used in the clustering procedures.

Spike Triggered Average The spike-triggered average (STA) is the average shape of the stimulus that precedes each spike. We extracted the STA using the following equation given by ([35]):

$$STA = \frac{1}{N} \Sigma_{n=1}^{N} \overrightarrow{s}(t_n), \tag{12}$$

where t_n is the nth spike time, s is the stimulus vector preceding the spike for a fixed time window of 100 ms, and N is the total number of spikes. Before clustering, we standardize (i.e. z score) and then normalize the STA vector with an L_2 norm. We didn't use any kind of whitening or regularization to calculate the STA.

UMAP + Louvain clustering

Conventional clustering algorithms such as K-means do not perform well in high dimensional spaces (p$>>$N, where p is the dimension of data and N is the number of samples) due to the

curse of dimensionality ([63]), and therefore need a pre-processing dimensionality reduction step. Addressing this issue, a non-linear dimensionality reduction algorithm such as UMAP ([64]) creates a high-dimensional graph representation of the data which can be utilized for clustering using a graph-based clustering method such as Louvain clustering ([65]) or ensemble clustering ([66]). This method utilizes the high dimensional space of the data for clustering. As shown by ([28]) using the WaveMAP algorithm, the UMAP+Louvain community detection algorithm has been successful in finding neuron types based on extracellular recordings.

UMAP

Universal Manifold Approximator (UMAP) is a non-linear dimensionality reduction technique that preserves local and global relationships between data in high dimensional space ([64]). It is divided into two steps, the first step is creating a k-nearest neighbor graph and the second step is to generate a low-dimensional representation that is similar to the high-dimensional graph structure.

We used the Scikit-learn UMAP-learn software package ([67]) to extract the embedding and the graphs.

Louvain Community detection

The Louvain community detection algorithm ([65]) maximizes modularity amongst the identified groups in a graph. Modularity can be defined by the following equation:

$$Q = \frac{1}{w} \sum_{i,j} \left[A_{i,j} - \gamma \frac{d_i^+, d_i^-}{w} \right] \delta(c_i, c_j) \tag{13}$$

where $A_{i,j}$ is the adjacency matrix, k_i = $\sum_j A_{ij}$ is the sum of the weights of the edges attached to the vertex i, c_i is the community for vertex i, d_i is the degree of node i, d_i^+ and d_i^- are the in degree and out degree for node i,$\delta(u, v)$, Kronecker symbol, is 1 if u=v and 0 otherwise, w = $\sum_{i,j} A_{i,j}$ and $\gamma > 0$ is the resolution parameter. For Louvain graph-based clustering we used the implementation from the Scikit-network software package ([68]).

The clustering approach can be summarized in the following steps:

1. The high dimensional k-neighbor graph is obtained by the first step of the UMAP algorithm using data vectors that are first standardized and then normalized using the L_2 norm. The nearest neighbor and distance parameters for UMAP were 20 and 0.1 respectively. This is to ensure a compact embedding and a clear clustering.

2. Using the graph obtained in the first step, we perform Lovain community-based clustering, using the resolution parameter γ that maximizes the modularity score, and the corresponding community/partition is chosen as the final cluster labels.

3. Using the cluster labels found in the second step, we color the individual points in distinct colors on the low-dimensional UMAP representation.

This unsupervised clustering approach was effective in capturing the global structure of the high dimensional space across attributes such as waveforms, adaptation current, and Spike Triggered average. The clusters found using this method were robust even when clustering was repeated with a sub-sample of the data and while iteratively removing the features.

Cluster Stability and Parameter Selection Cluster stability is tested by clustering a 90% sub-sample of the data chosen at random and repeating the procedure 25 times for each resolution parameter, varying from 0 to 5 with a step of 0.5. The modularity score is calculated for each resolution parameter and finally, the resolution parameter is chosen for which the modularity score is maximal. Variation in the number of clusters is observed for each resolution parameter and is contrasted with the modularity score. Clustering robustness is also tested by repeating the procedure above while excluding one feature at a time for Action potential feature clustering and passive biophysical clustering.

Cluster Likelihood comparison Cluster likelihood between two sets of labels was calculated in two steps, firstly we created a contingency matrix such that $C_{i,j}$ contained the number of times neurons classified in cluster i in the SH protocol classified as cluster j in the FN protocol, such that each row contains the division of elements of cluster i in the SH protocol into all the clusters of the FN protocol. Secondly, to get the likelihood, we divided each row by the total count of neurons in cluster i in the SH protocol.

$$P(j|i) = \frac{C_{i,j}}{\sum_j C_{i,j}} \tag{14}$$

Here $P(j|i)$ is the probability of the neuron classifying into FN cluster j given it is classified in SH cluster i.

Cluster Similarity Measures We measured the similarity between two given cluster assignments using the Adjusted Random Index (ARI) and Adjusted Mutual Information (AMI) Score. Both of these measures amount to change agreements between two clusters. The ARI (range -1-1) measures the pairwise relationships between clusters and the AMI (range 0-1) measures the overall information shared between the two clusters.
We used the scikit-learn Python package ([69]) to calculate the ARI and AMI measures.

Ensemble Clustering for Graphs (ECG)

We used Ensemble Clustering for Graph algorithm ([66]) to validate the clusters found using the Louvain Community detection algorithm. It is a two-step algorithm, the first step called the generation step consists of producing a $k-$level 1 partition *P* $= \{P_1, P_2, ...P_k\}$ by running the first pass of the Louvain clustering algorithm with random vertices on the initial graph G= (V, E). The second step, also known as the integration step, consists of running the Louvain algorithm on a weighted version of the initial graph. Where the weights are the weight of an edge given by

$$W_p(u,v) = \begin{cases} w_* + (1-w_*) * \left(\frac{\sum_{i=1}^{k} v_{p_i}(u,v)}{k} \right) & \text{, if (u,v) is in 2-core of G} \\ w_* & \text{, otherwise} \end{cases} \tag{15}$$

where $0 < w_* < 1$ is the minimum ECG weight and $v_{p_i}(u,v) = \sum_{j=1}^{l_i} 1_{C_i^j}(u) \cdot 1_{C_i^j(v)}$ shows if the vertices u and v co-oSHur in the same cluster of P_i or not. Thus it takes advantage of multiple instances of the Louvain clustering algorithm to make a clustering based on consensus. We used the implementation provided by ([28]) for comparing the waveform clustering for SH and FN protocols. The graph used for Ensemble clustering was the same as in the original clustering using the UMAP algorithm.

Quantification and statistical analysis

MANOVA We measured the significance between excitatory and inhibitory action potential and passive biophysical feature vectors using a one-sided MANOVA. We present the following statistics: Wilk's lambda, Pillai's trace, Hotelling-Lawley trace, and Roy's greatest root using the Stats Model python package[70].

Canonical Correlation Analysis To perform a post-hoc analysis on Action potential and Passive biophysical features, to find the importance of each feature, we used a Canonical Correlation analysis. Which is a method to find a linear combination of features between two datasets that maximizes the correlation between them. It is a deterministic method that results in canonical variates of the two datasets that are maximally correlated. Since the excitatory/inhibitory populations were different in numbers, we repeated the SHA procedure 10 times with random sampling from the larger group to make the population size equal between excitatory and inhibitory groups. We then obtain loadings for each dataset by averaging over the 10 repetitions, which represents the correlation of a feature with the canonical variate. We used the Scikit learn python package ([69]) for performing the SHA procedure.

Welch's ANOVA To compare the significance of cosine similarity measures for excitatory and inhibitory populations, we first calculated the cosine similarity matrix comparing the excitatory and inhibitory populations separately and a third matrix comparing excitatory with inhibitory populations. We take the upper triangular part of the excitatory and inhibitory within-population comparison matrix and the entire excitatory vs inhibitory matrix for the significance test. The

Welch's ANOVA test was used to compare the three groups of cosine similarity for each attribute with one categorical variable namely, excitatory or inhibitory. This test was chosen because the populations for each attribute were heterodrastic and of unequal size. We performed a post-hoc Games-Howell test to determine which group was significantly different, this was done because the variances across groups were heterodrastic. The significance levels are reported in the figures based on the p-values obtained by the post-hoc test. We used the Stats model python package[70] for calculating Welch's ANOVA.

Multi Correlation and Factor Analysis

The Multi Correlation and Factor Analysis (MCFA) was performed using the procedure as described by ([36]), and the accompanying software was used for the analysis ([71]). The problem can be stated as follows, let $\{Y^m\}_{m=1}^M$ be a set of M attributes extracted from the electrophysiological data, each with dimension NxP_m, where N is the number of samples and P_m is the dimension of each attribute (Action potential parameters (P_1 =22), passive biophysical parameters (P_2 =6), Spike triggered Average (P_3 =2000), and Spike triggered current (η) (P_4 =10000)). Each attribute set can be modeled as having a contribution from two factors, a shared and a private factor respectively, as shown below.

$$z_n \sim N(0, I_d) \tag{16}$$

$$x_n^m \sim N(0, I_{k_m}) \tag{17}$$

$$y_n^m \sim N(W_m z_n + L_m x_n^m, I_{k_m}, \Psi_m) \tag{18}$$

Here z_n is the shared factor of dimension d, and W_m is the shared space loading matrix of shape P_mx d. k_m is the dimension of each private mode. x^m is the private space for each attribute m of dimension k_m, L_m is the private space loading of shape P_mxk_m. $\Psi_m = diag(\psi_m^1, ..., \psi_m^{\rho_m})$ are the diagonal residual covariance matrices. Given Y, d, and k_m the goal is to find hidden factors z_n, x_n^m, and loading matrices W_m, L_m. This is achieved by the Expectation Minimization (EM) method[36].

We center and scale all variables as in[36] and initialize the loading matrices similarly to the original method, using the pairwise correlations with average variance constraint initialization. To model the shared latent space z_n, we chose the most informative PCA components based on Marchenko Pasteur Law to control over-fitting ([36]), which states for any normalized dataset ($\mu = 0$, $\sigma = 1$), the principal components with eigenvalues above $\lambda_m = 1 + \sqrt{p_m/N}$ are considered non-noise. We set the size of the private space $k_m = 1$ for passive biophysical parameters due to its relatively lower dimension and $k_m = 2$ for the other attributes. After initializing all the variables, we run the expectation minimization (EM) algorithm to obtain z_n, x_n^m, W_m, and L_m matrices.

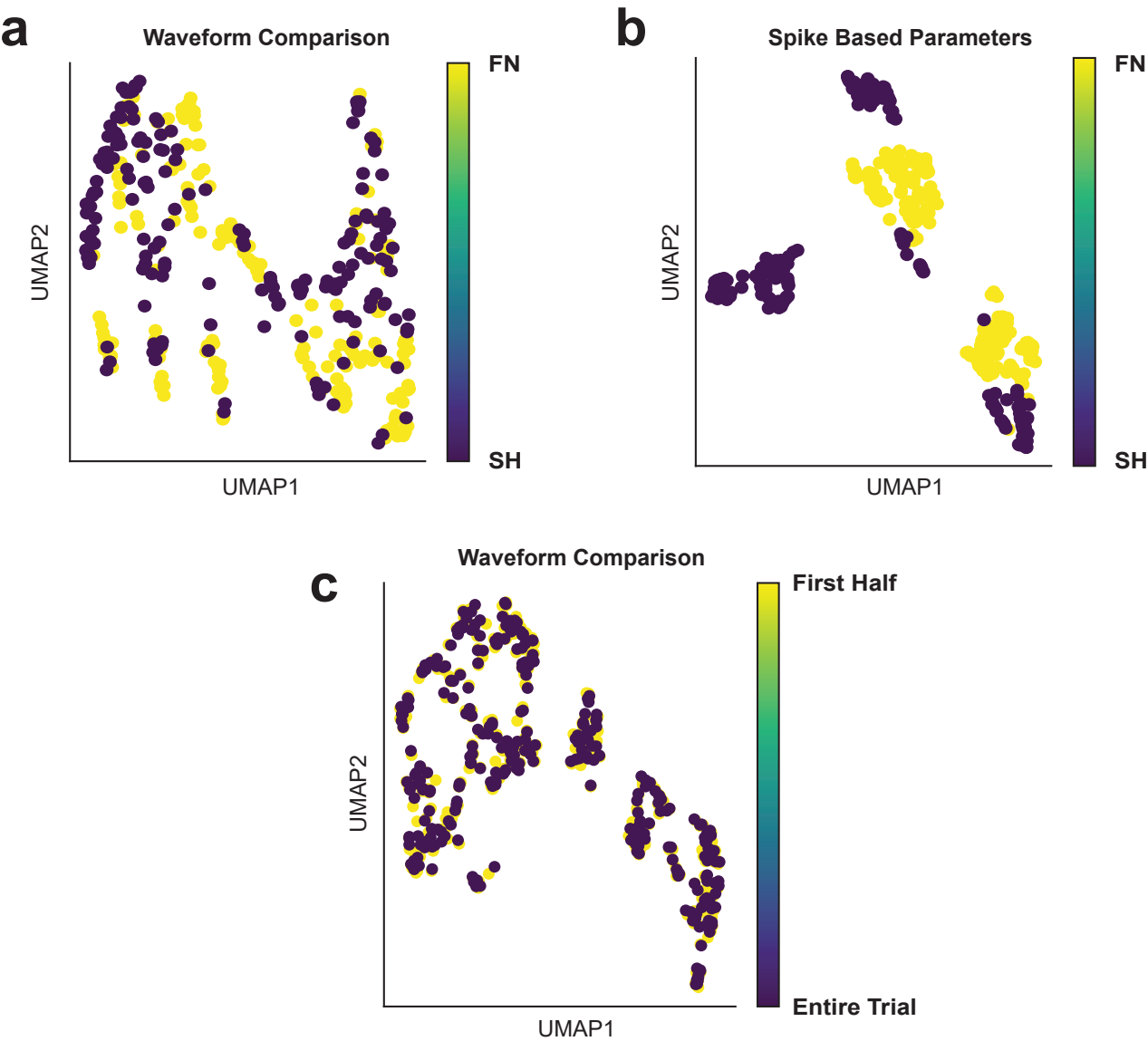
For feature j of mode m, the variance explained by a shared feature d is given by $W_m^{(j,d)2}$. Similarly, the variance explained by the kth private factor of feature j of mode m is given by $L_m^{(j,k_m)2}$. The total variance explained for a mode by a shared factor is given by $\sum_j W_m^{(j,d)2}$, similarly, the total variance explained by the private factor k_m is $\sum_j L_m^{(j,k_m)2}$. Hence, the total variance explained by all shared and all private factors is given by $\sum_{j,d} W_m^{(j,d)2}$ and $\sum_{j,k_m} L_m^{(j,k_m)2}$.

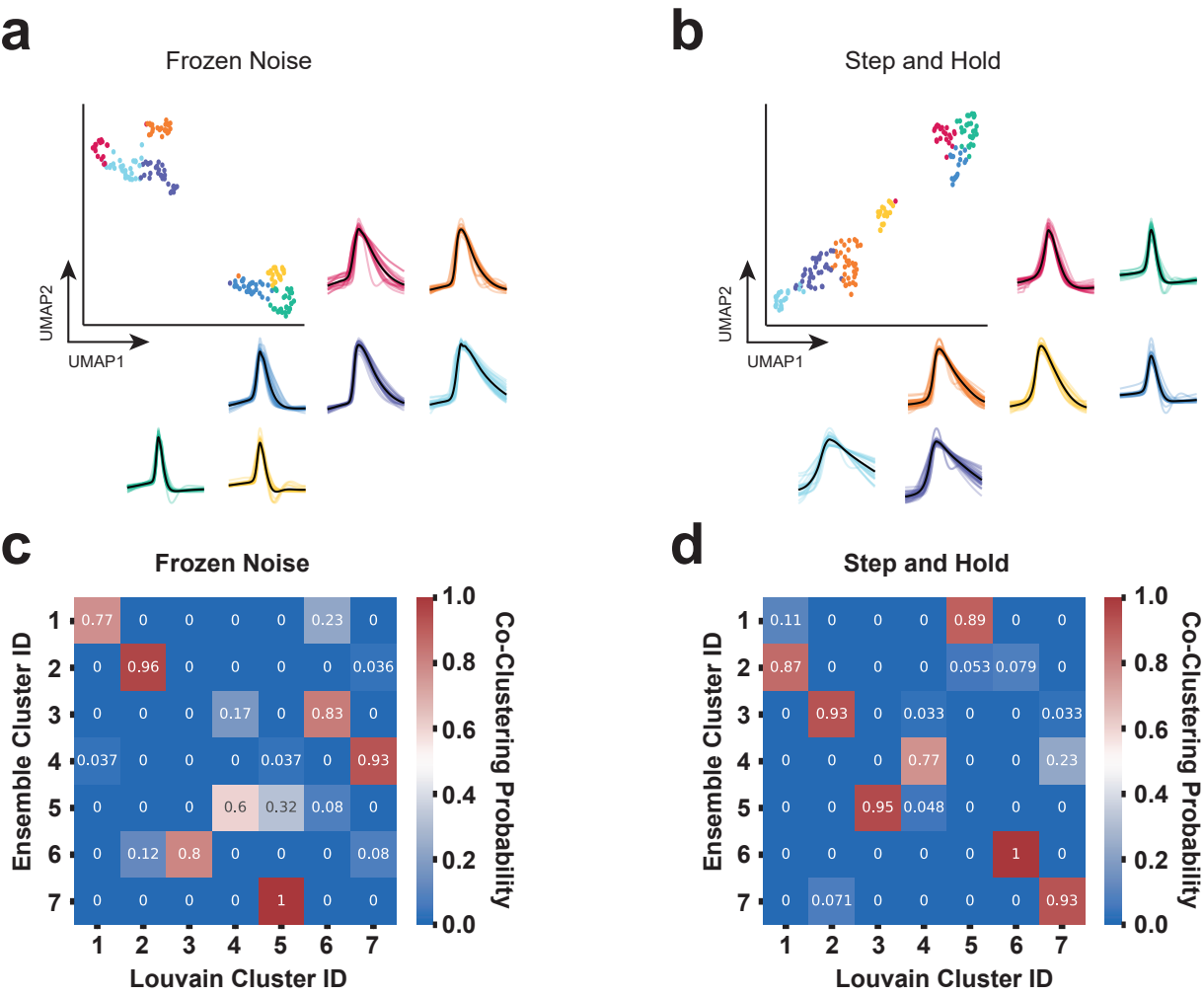
The relative feature importance is given by the cross-correlation of columns of the posterior mean of Z on observing a single mode denoted by $\hat{Z}_m$[36]

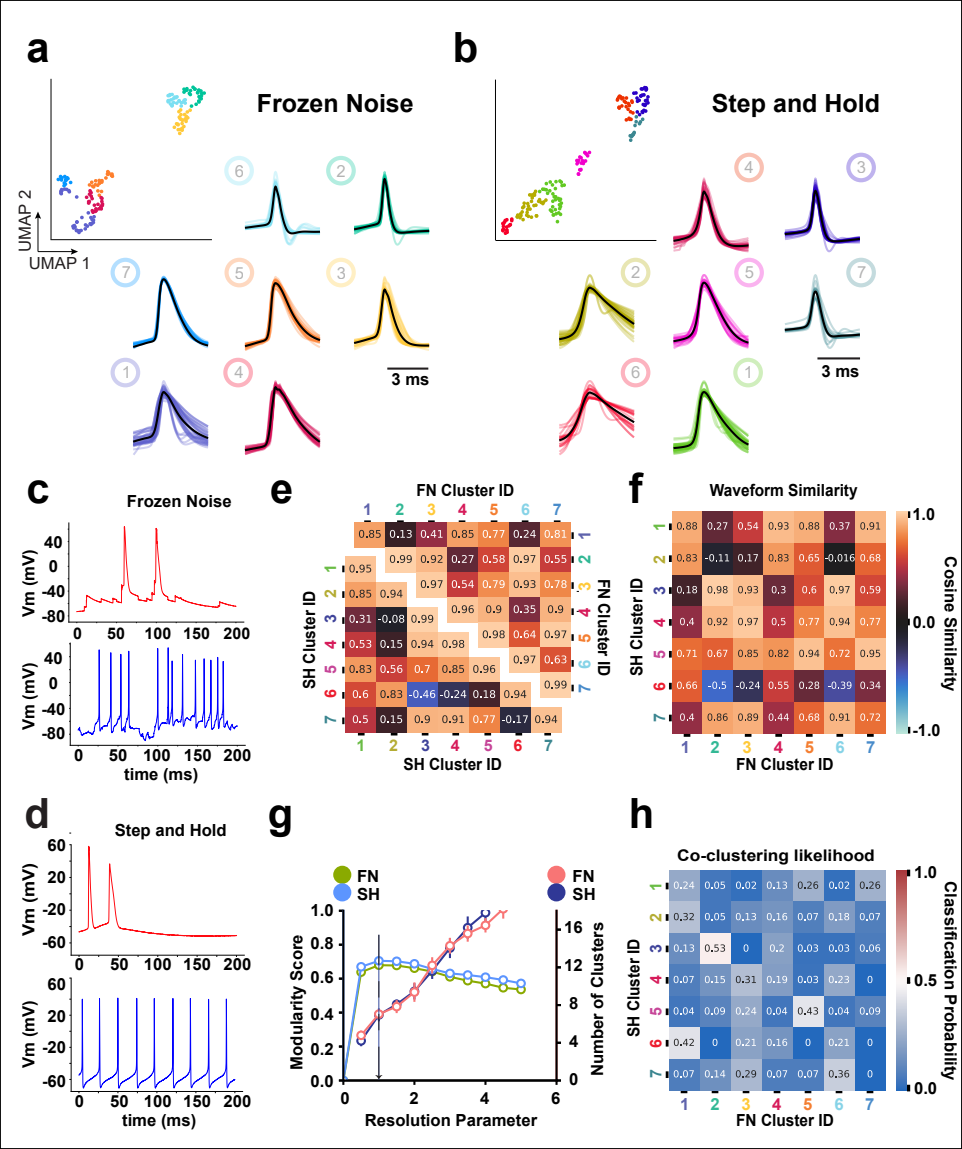
$$\hat{Z}_m = E[Z|W_m, \Psi_m, L_m, Y_m] = Y_m(W_m W_m^T + L_m L_m^T \Psi_m)^{-1} W_m \tag{19}$$

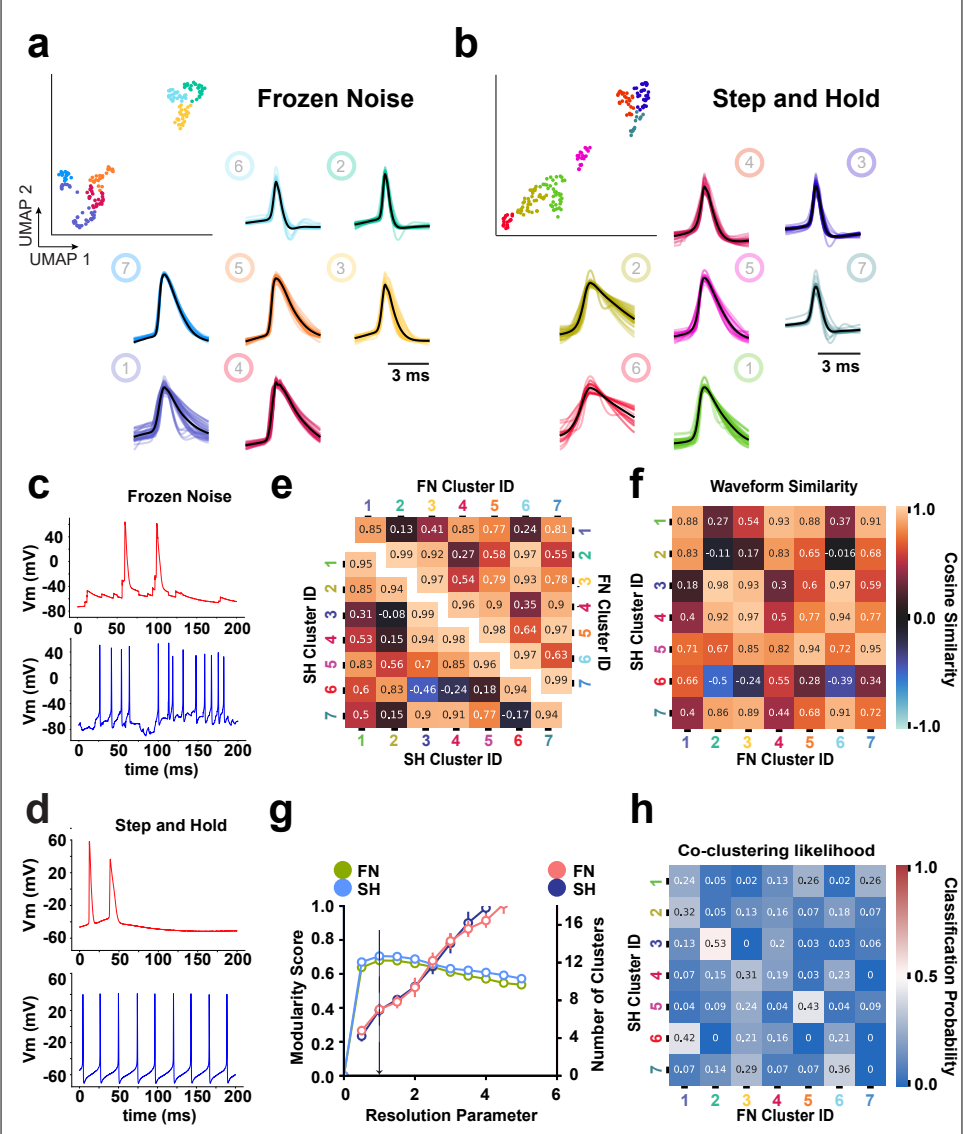
$$S_d = cor(\hat{Z}_1^{(:,d)}, ..., \hat{Z}_m^{(:,d)}) \tag{20}$$

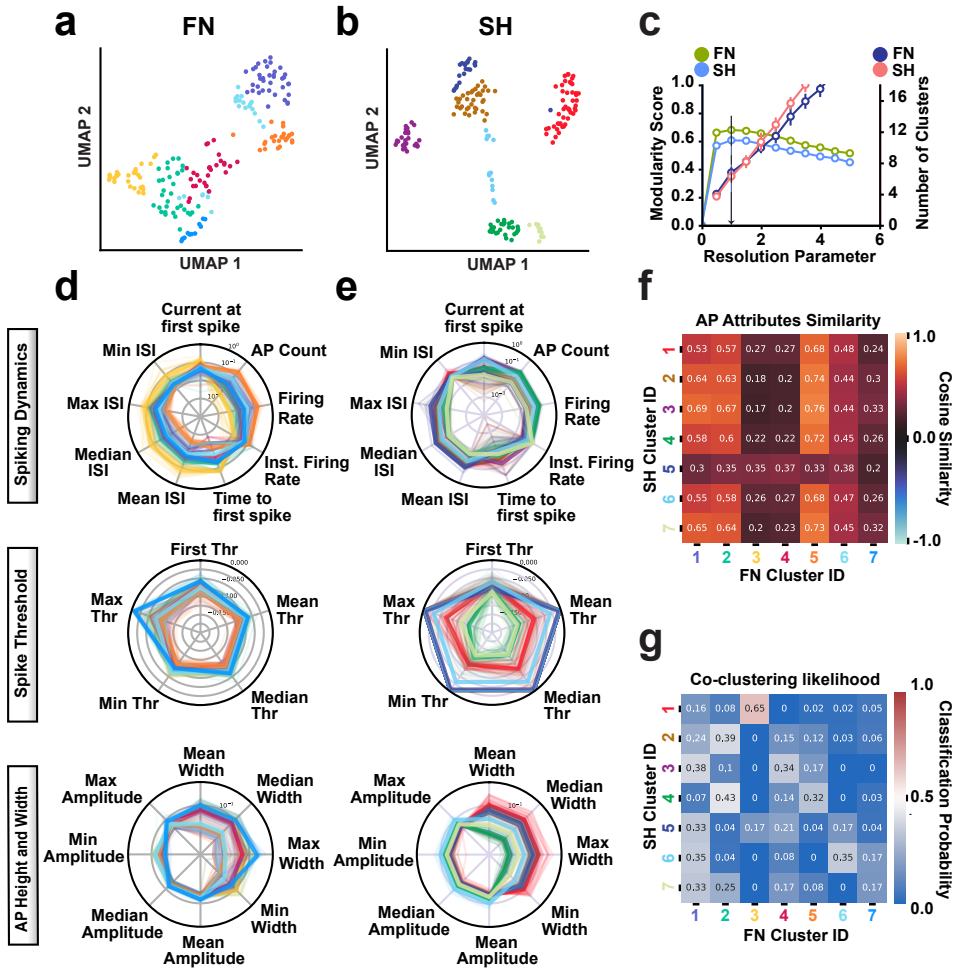
Code and Data Availability All analyzed current clamp and simulation data and the code to analyze and simulate them can be found in this repository: https://doi.org/10.34973/4f3k-1s63. The database is an extended version of the dataset ([22,55]). The code for all figures and analyses is available here: https://github.com/Nishant-codex/single_cell_analysis.git.

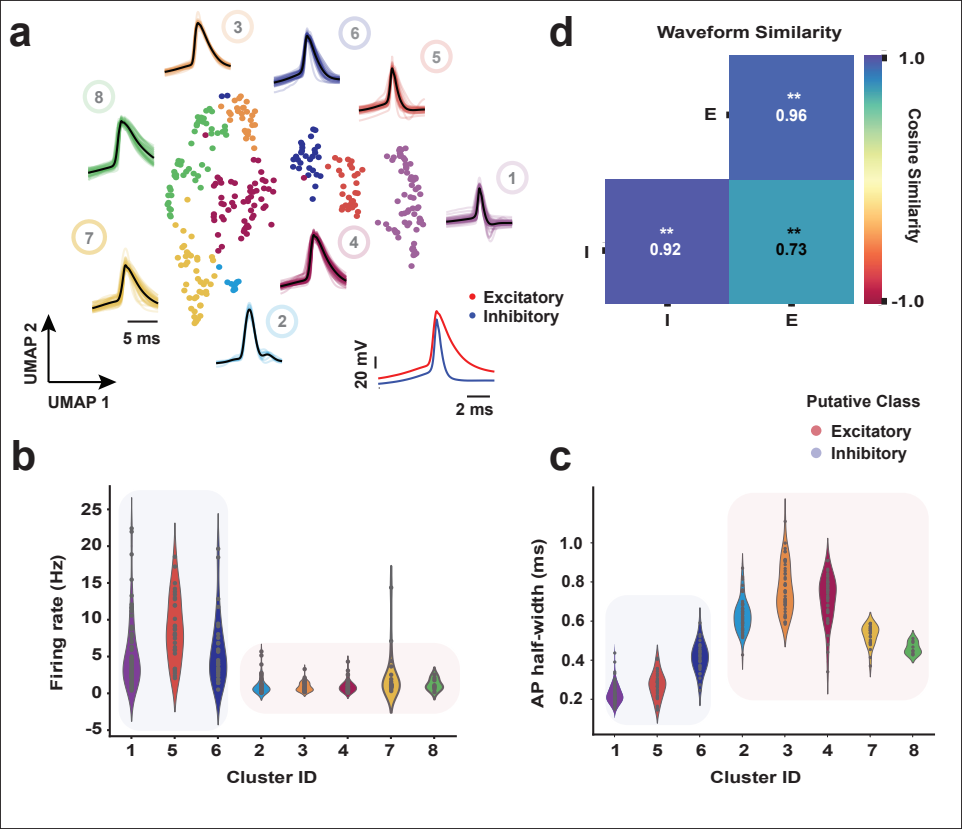

a
Waveform Comparison
FN
SH
UMAP2
UMAP1
b
Spike Based Parameters
FN
SH
UMAP2
UMAP1
c
Waveform Comparison
First Half
Entire Trial
UMAP2
UMAP1

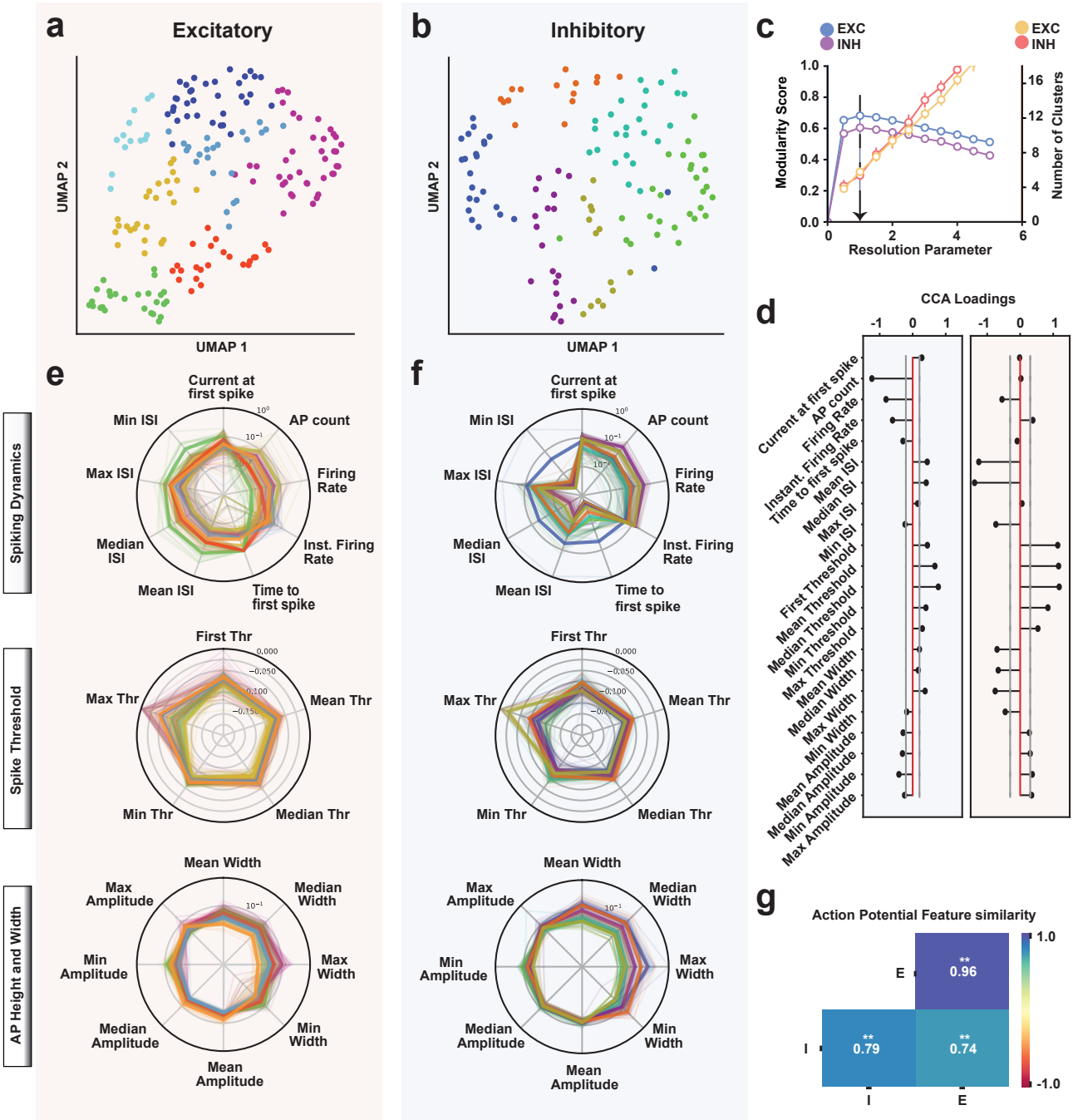

a
Frozen Noise
UMAP2
UMAP1
b
Step and Hold
UMAP2
UMAP1
c
Frozen Noise
Ensemble Cluster ID
Louvain Cluster ID
Co-Clustering Probability
1 2 3 4 5 6 7
0.77 0 0 0 0 0.23 0
0 0.96 0 0 0 0 0.036
0 0 0 0.17 0 0.83 0
0.037 0 0 0 0.037 0 0.93
0 0 0 0.6 0.32 0.08 0
0 0.12 0.8 0 0 0 0.08
0 0 0 0 1 0 0
1.0 0.8 0.6 0.4 0.2 0.0
d
Step and Hold
Ensemble Cluster ID
Louvain Cluster ID
Co-Clustering Probability
1 2 3 4 5 6 7
0.11 0 0 0 0.89 0 0
0.87 0 0 0 0.053 0.079 0
0 0.93 0 0.033 0 0 0.033
0 0 0 0.77 0 0 0.23
0 0 0.95 0.048 0 0 0
0 0 0 0 0 1 0
0 0.071 0 0 0 0 0.93
1.0 0.8 0.6 0.4 0.2 0.0

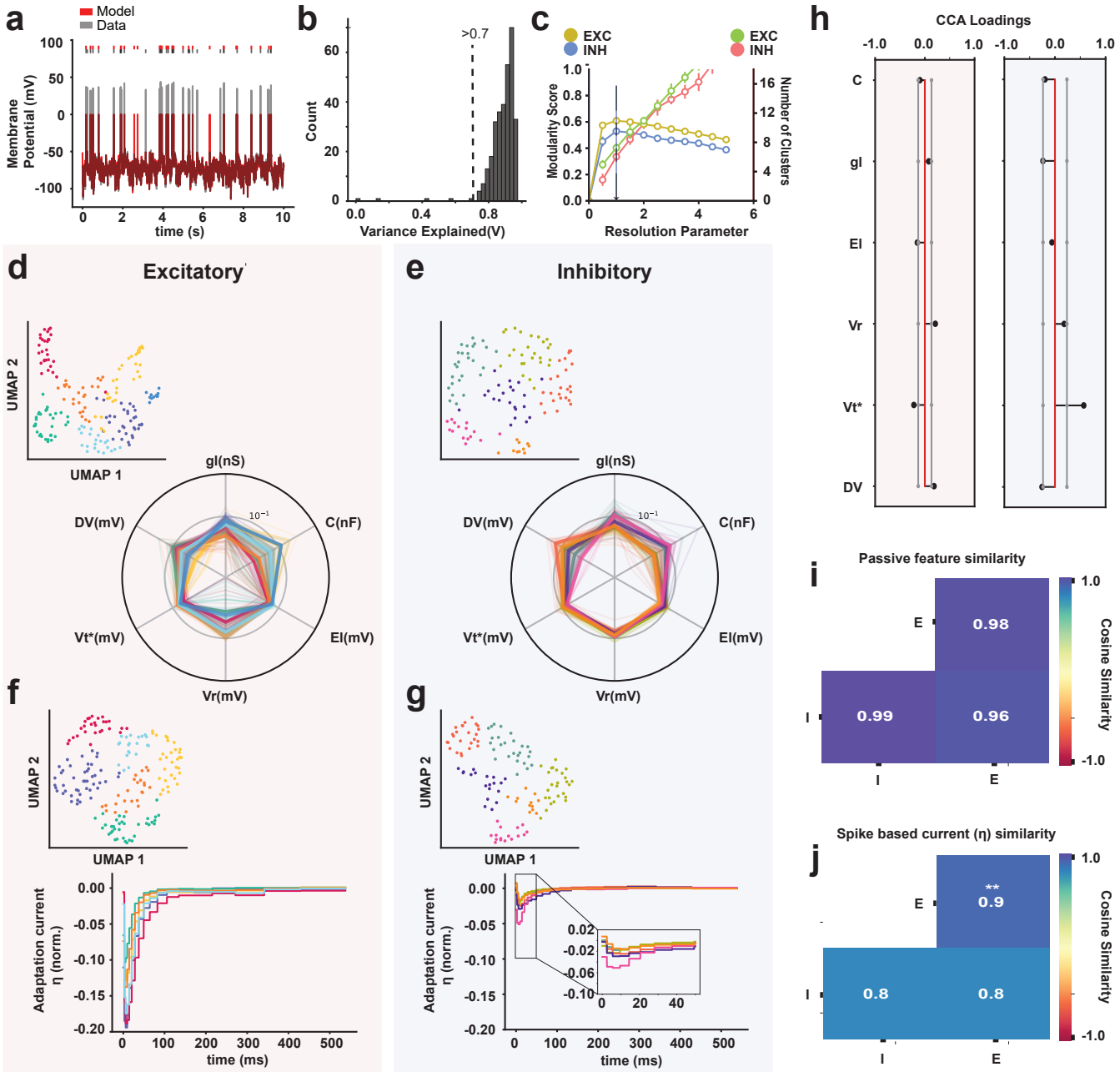

a
Frozen Noise
UMAP 2
UMAP 1
6
2
7
5
3
1
4
3 ms
b
Step and Hold
4
3
2
5
7
6
1
3 ms
c
Frozen Noise
Vm (mV)
40
0
-40
-80
0
50
100
150
200
time (ms)
d
Step and Hold
Vm (mV)
60
20
-20
-60
time (ms)
e
FN Cluster ID
SH Cluster ID
1 2 3 4 5 6 7
0.85 0.13 0.41 0.85 0.77 0.24 0.81
0.99 0.92 0.27 0.58 0.97 0.55
0.95 0.97 0.54 0.79 0.93 0.78
0.85 0.94 0.96 0.9 0.35 0.9
0.31 -0.08 0.99 0.98 0.64 0.97
0.53 0.15 0.94 0.98 0.97 0.63
0.83 0.56 0.7 0.85 0.96 0.99
0.6 0.83 -0.46 -0.24 0.18 0.94
0.5 0.15 0.9 0.91 0.77 -0.17 0.94
f
Waveform Similarity
SH Cluster ID
FN Cluster ID
Cosine Similarity
1.0
0.0
-1.0
0.88 0.27 0.54 0.93 0.88 0.37 0.91
0.83 -0.11 0.17 0.83 0.65 -0.016 0.68
0.18 0.98 0.93 0.3 0.6 0.97 0.59
0.4 0.92 0.97 0.5 0.77 0.94 0.77
0.71 0.67 0.85 0.82 0.94 0.72 0.95
0.66 -0.5 -0.24 0.55 0.28 -0.39 0.34
0.4 0.86 0.89 0.44 0.68 0.91 0.72
g
FN
SH
FN
SH
Modularity Score
Number of Clusters
Resolution Parameter
1.0
0.8
0.6
0.4
0.2
0.0
16
12
8
4
0
0
2
4
6
h
Co-clustering likelihood
SH Cluster ID
FN Cluster ID
Classification Probability
1.0
0.5
0.0
0.24 0.05 0.02 0.13 0.26 0.02 0.26
0.32 0.05 0.13 0.16 0.07 0.18 0.07
0.13 0.53 0 0.2 0.03 0.03 0.06
0.07 0.15 0.31 0.19 0.03 0.23 0
0.04 0.09 0.24 0.04 0.43 0.04 0.09
0.42 0 0.21 0.16 0 0.21 0
0.07 0.14 0.29 0.07 0.07 0.36 0

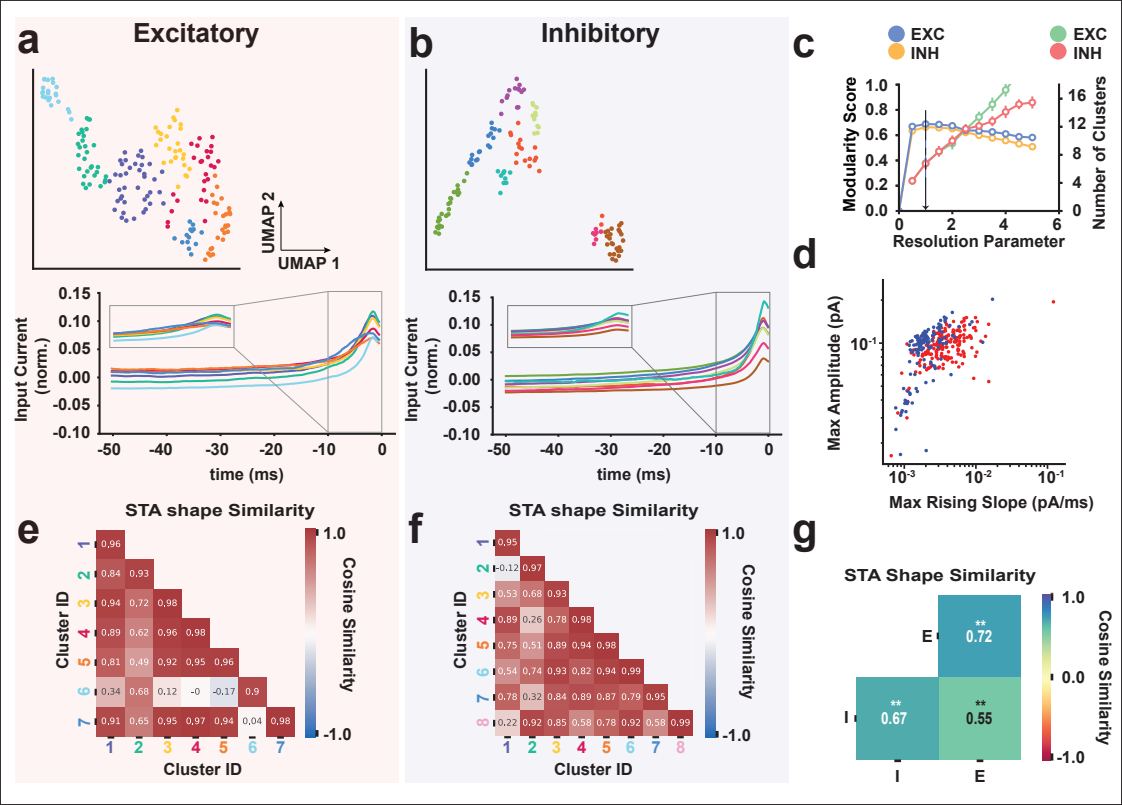

a
Frozen Noise
UMAP 2
UMAP 1
6
2
7
5
3
1
4
3 ms
b
Step and Hold
4
3
2
5
7
6
1
3 ms
c
Frozen Noise
Vm (mV)
40
0
-40
-80
0
50
100
150
200
time (ms)
d
Step and Hold
Vm (mV)
60
20
-20
-60
time (ms)
e
FN Cluster ID
SH Cluster ID
1 2 3 4 5 6 7
0.85 0.13 0.41 0.85 0.77 0.24 0.81
0.99 0.92 0.27 0.58 0.97 0.55
0.95 0.97 0.54 0.79 0.93 0.78
0.85 0.94 0.96 0.9 0.35 0.9
0.31 -0.08 0.99 0.98 0.64 0.97
0.53 0.15 0.94 0.98 0.97 0.63
0.83 0.56 0.7 0.85 0.96 0.99
0.6 0.83 -0.46 -0.24 0.18 0.94
0.5 0.15 0.9 0.91 0.77 -0.17 0.94
f
Waveform Similarity
SH Cluster ID
FN Cluster ID
1 2 3 4 5 6 7
0.88 0.27 0.54 0.93 0.88 0.37 0.91
0.83 -0.11 0.17 0.83 0.65 -0.016 0.68
0.18 0.98 0.93 0.3 0.6 0.97 0.59
0.4 0.92 0.97 0.5 0.77 0.94 0.77
0.71 0.67 0.85 0.82 0.94 0.72 0.95
0.66 -0.5 -0.24 0.55 0.28 -0.39 0.34
0.4 0.86 0.89 0.44 0.68 0.91 0.72
Cosine Similarity
1.0
0.0
-1.0
g
FN
SH
FN
SH
Modularity Score
Number of Clusters
Resolution Parameter
0.0
0.2
0.4
0.6
0.8
1.0
0
4
8
12
16
0
2
4
6
h
Co-clustering likelihood
SH Cluster ID
FN Cluster ID
1 2 3 4 5 6 7
0.24 0.05 0.02 0.13 0.26 0.02 0.26
0.32 0.05 0.13 0.16 0.07 0.18 0.07
0.13 0.53 0 0.2 0.03 0.03 0.06
0.07 0.15 0.31 0.19 0.03 0.23 0
0.04 0.09 0.24 0.04 0.43 0.04 0.09
0.42 0 0.21 0.16 0 0.21 0
0.07 0.14 0.29 0.07 0.07 0.36 0
Classification Probability
1.0
0.5
0.0

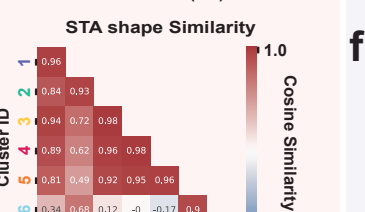

a
FN
UMAP 2
UMAP 1
b
SH
UMAP 2
UMAP 1
c
FN
SH
FN
SH
Modularity Score
Number of Clusters
Resolution Parameter
d
e
Spiking Dynamics
Spike Threshold
AP Height and Width
Current at first spike
AP Count
Firing Rate
Inst. Firing Rate
Time to first spike
Mean ISI
Median ISI
Max ISI
Min ISI
First Thr
Mean Thr
Median Thr
Min Thr
Max Thr
Mean Width
Median Width
Max Width
Min Width
Mean Amplitude
Median Amplitude
Min Amplitude
Max Amplitude
f
AP Attributes Similarity
SH Cluster ID
FN Cluster ID
Cosine Similarity
1 2 3 4 5 6 7
0.53 0.57 0.27 0.27 0.68 0.48 0.24
0.64 0.63 0.18 0.2 0.74 0.44 0.3
0.69 0.67 0.17 0.2 0.76 0.44 0.33
0.58 0.6 0.22 0.22 0.72 0.45 0.26
0.3 0.35 0.35 0.37 0.33 0.38 0.2
0.55 0.58 0.26 0.27 0.68 0.47 0.26
0.65 0.64 0.2 0.23 0.73 0.45 0.32
g
Co-clustering likelihood
SH Cluster ID
FN Cluster ID
Classification Probability
1 2 3 4 5 6 7
0.16 0.08 0.65 0 0.02 0.02 0.05
0.24 0.39 0 0.15 0.12 0.03 0.06
0.38 0.1 0 0.34 0.17 0 0
0.07 0.43 0 0.14 0.32 0 0.03
0.33 0.04 0.17 0.21 0.04 0.17 0.04
0.35 0.04 0 0.08 0 0.35 0.17
0.33 0.25 0 0.17 0.08 0 0.17

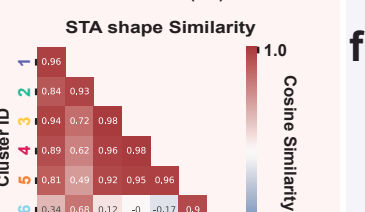

a
3
6
5
8
1
7
4
2
5 ms
UMAP 2
UMAP 1
Excitatory
Inhibitory
20 mV
2 ms
d
Waveform Similarity
E
I
**
0.96
**
0.92
**
0.73
I
E
1.0
-1.0
Cosine Similarity
Putative Class
Excitatory
Inhibitory
b
Firing rate (Hz)
25
20
15
10
5
0
-5
1
5
6
2
3
4
7
8
Cluster ID
c
AP half-width (ms)
1.0
0.8
0.6
0.4
0.2
1
5
6
2
3
4
7
8
Cluster ID

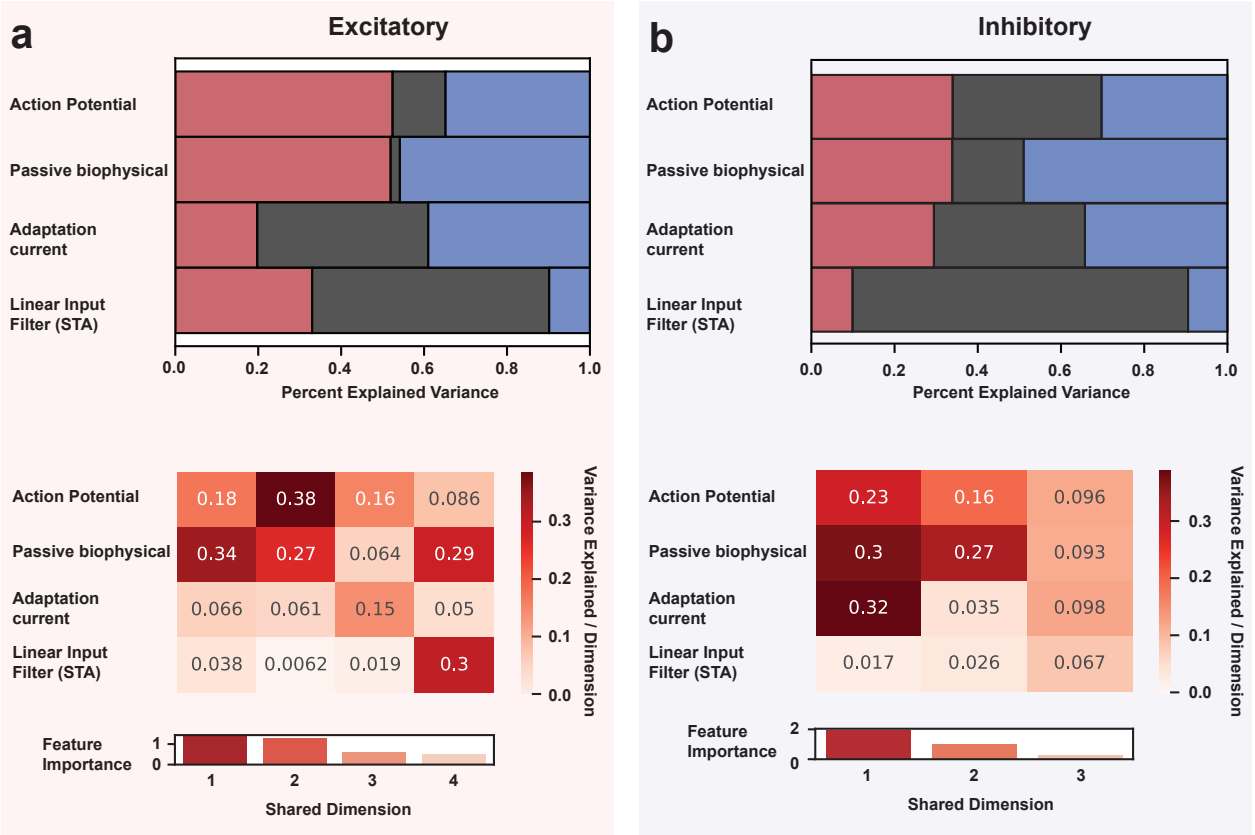

a
Excitatory
UMAP 2
UMAP 1
b
Inhibitory
UMAP 2
UMAP 1
c
EXC
INH
EXC
INH
Modularity Score
Number of Clusters
Resolution Parameter
d
CCA Loadings
Current at first spike
AP count
Firing Rate
Instant. Firing Rate
Time to first spike
Mean ISI
Median ISI
Max ISI
Min ISI
First Threshold
Mean Threshold
Median Threshold
Min Threshold
Max Threshold
Mean Width
Median Width
Max Width
Min Width
Mean Amplitude
Median Amplitude
Min Amplitude
Max Amplitude
e
f
Spiking Dynamics
Spike Threshold
AP Height and Width
Current at first spike
AP count
Firing Rate
Inst. Firing Rate
Time to first spike
Mean ISI
Median ISI
Max ISI
Min ISI
First Thr
Mean Thr
Median Thr
Min Thr
Max Thr
Mean Width
Median Width
Max Width
Min Width
Mean Amplitude
Median Amplitude
Min Amplitude
Max Amplitude
g
Action Potential Feature similarity
E
I
0.96
0.79
0.74
1.0
-1.0

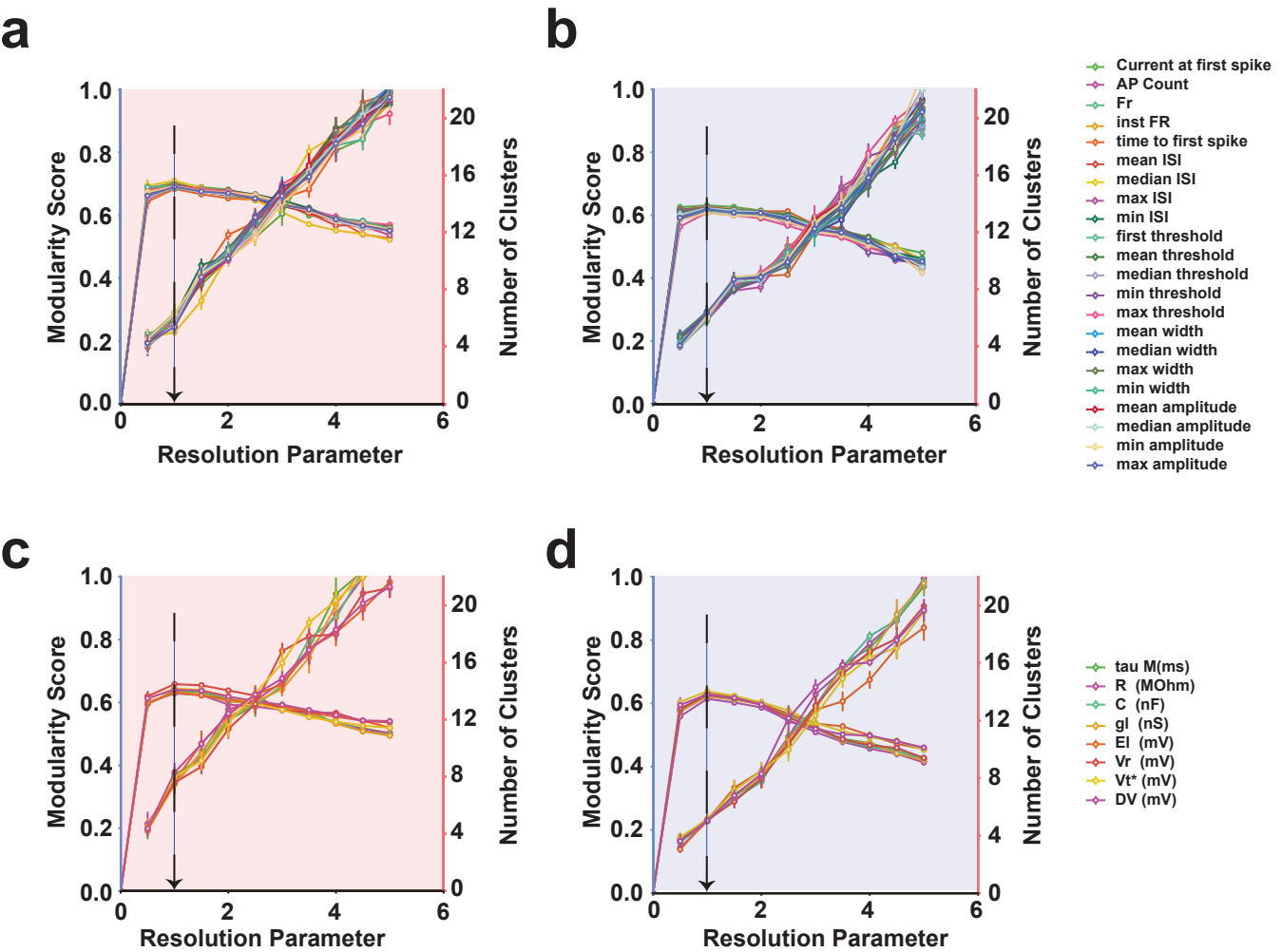

a
Model
Data
Membrane Potential (mV)
time (s)
b
>0.7
Count
Variance Explained(V)
c
EXC
INH
EXC
INH
Modularity Score
Number of Clusters
Resolution Parameter
h
CCA Loadings
C
gl
El
Vr
Vt*
DV
d
Excitatory
UMAP 2
UMAP 1
gl(nS)
C(nF)
El(mV)
Vr(mV)
Vt*(mV)
DV(mV)
e
Inhibitory
f
g
Adaptation current η (norm.)
time (ms)
i
Passive feature similarity
0.98
0.99
0.96
Cosine Similarity
j
Spike based current (η) similarity
**
0.9
0.8
0.8

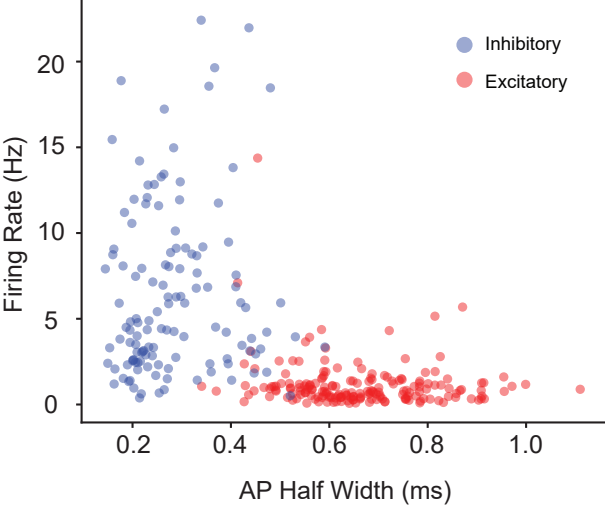

a
Excitatory
UMAP 2
UMAP 1
Input Current (norm.)
time (ms)
b
Inhibitory
Input Current (norm.)
time (ms)
c
EXC
INH
EXC
INH
Modularity Score
Number of Clusters
Resolution Parameter
d
Max Amplitude (pA)
Max Rising Slope (pA/ms)
e
STA shape Similarity
Cluster ID
Cosine Similarity
f
STA shape Similarity
Cluster ID
Cosine Similarity
g
STA Shape Similarity
E
I
0.72
0.67
0.55
Cosine Similarity


STA shape Similarity
Cluster ID
1
2
3
4
5
6
0,96
0,84 0,93
0,94 0,72 0,98
0,89 0,62 0,96 0,98
0,81 0,49 0,92 0,95 0,96
0,34 0,68 0,12 -0 -0,17 0,9
1.0
Cosine Similarity

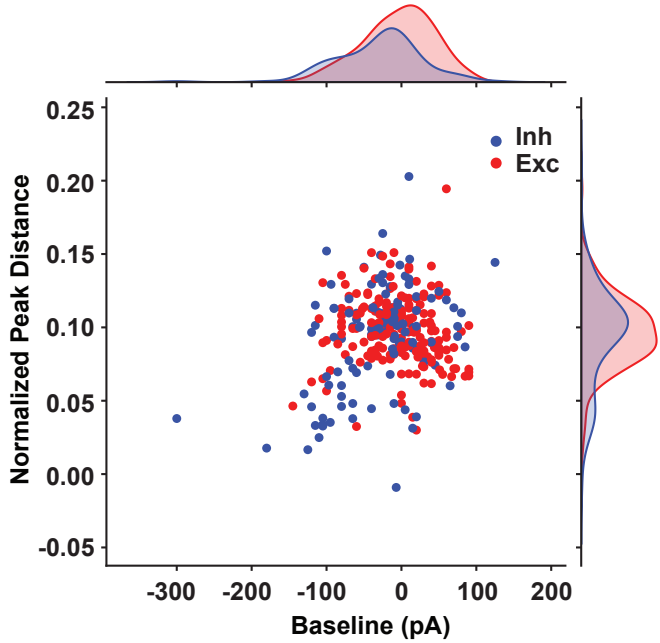

STA shape Similarity
Cluster ID
1
2
3
4
5
6
0,96
0,84 0,93
0,94 0,72 0,98
0,89 0,62 0,96 0,98
0,81 0,49 0,92 0,95 0,96
0,34 0,68 0,12 -0 -0,17 0,9
1.0
Cosine Similarity


a
Excitatory
Action Potential
Passive biophysical
Adaptation current
Linear Input Filter (STA)
0.0
0.2
0.4
0.6
0.8
1.0
Percent Explained Variance
Action Potential 0.18 0.38 0.16 0.086
Passive biophysical 0.34 0.27 0.064 0.29
Adaptation current 0.066 0.061 0.15 0.05
Linear Input Filter (STA) 0.038 0.0062 0.019 0.3
Variance Explained / Dimension
0.0
0.1
0.2
0.3
Feature Importance
1
0
1
2
3
4
Shared Dimension
b
Inhibitory
Action Potential
Passive biophysical
Adaptation current
Linear Input Filter (STA)
0.0
0.2
0.4
0.6
0.8
1.0
Percent Explained Variance
Residual
Private
Shared
Action Potential 0.23 0.16 0.096
Passive biophysical 0.3 0.27 0.093
Adaptation current 0.32 0.035 0.098
Linear Input Filter (STA) 0.017 0.026 0.067
Variance Explained / Dimension
0.0
0.1
0.2
0.3
Feature Importance
2
0
1
2
3
Shared Dimension

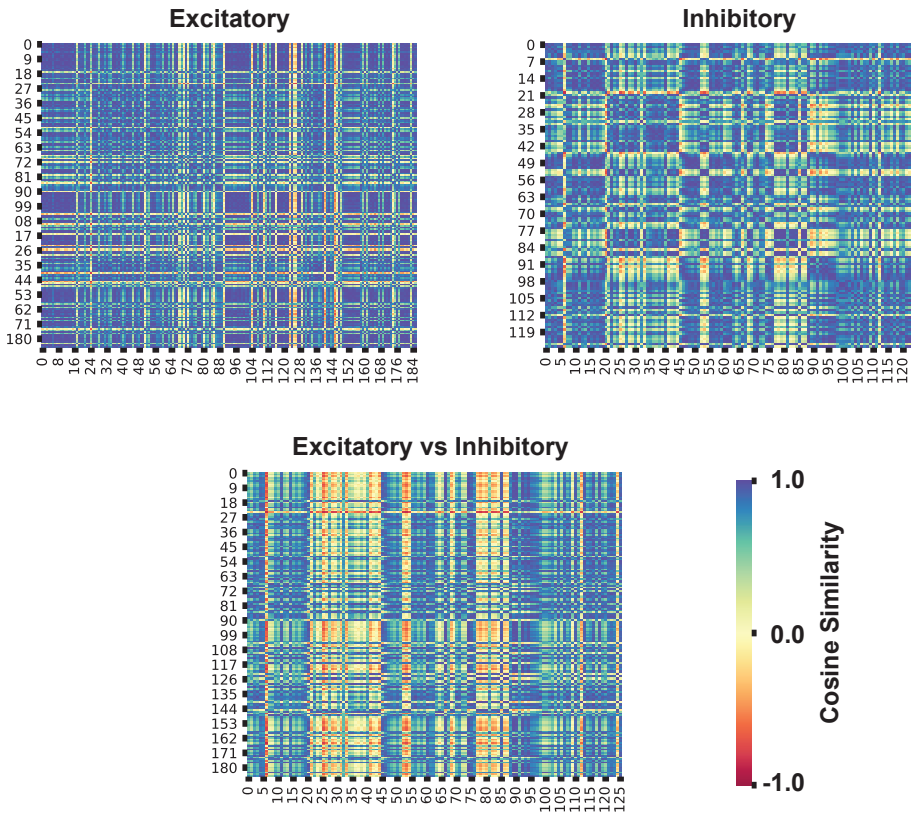

a
b
c
d
Modularity Score
Number of Clusters
Resolution Parameter
1.0
0.8
0.6
0.4
0.2
0.0
20
16
12
8
4
0
2
4
6
Current at first spike
AP Count
Fr
inst FR
time to first spike
mean ISI
median ISI
max ISI
min ISI
first threshold
mean threshold
median threshold
min threshold
max threshold
mean width
median width
max width
min width
mean amplitude
median amplitude
min amplitude
max amplitude
tau M(ms)
R (MOhm)
C (nF)
gl (nS)
El (mV)
Vr (mV)
Vt* (mV)
DV (mV)

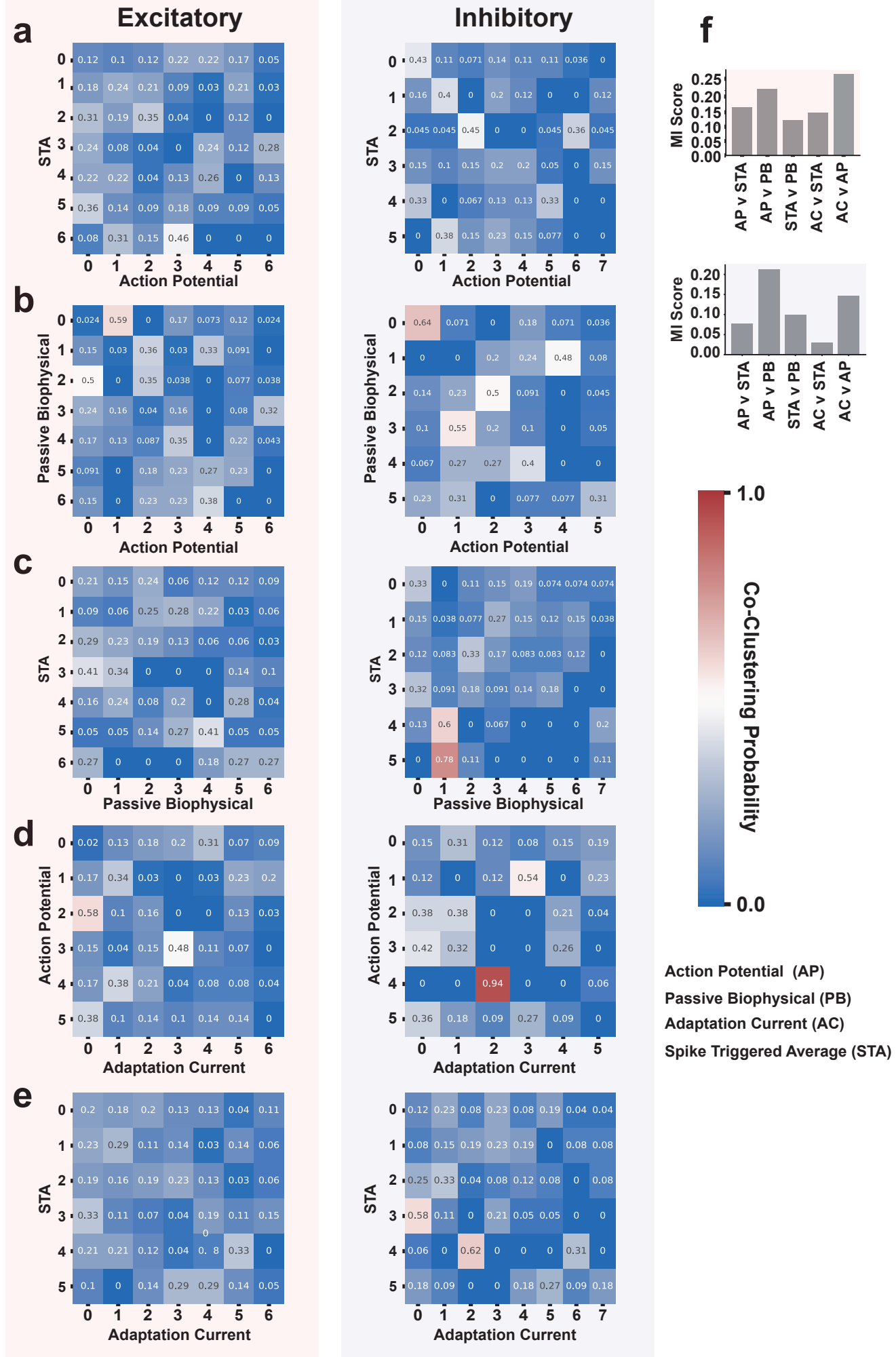

Inhibitory
Excitatory
Firing Rate (Hz)
0
5
10
15
20
0.2
0.4
0.6
0.8
1.0
AP Half Width (ms)

a


Inh
Exc
Normalized Peak Distance
0.25
0.20
0.15
0.10
0.05
0.00
-0.05
-300
-200
-100
0
100
200
Baseline (pA)

b


Excitatory
Inhibitory
Excitatory vs Inhibitory
Cosine Similarity
1.0
0.0
-1.0


Excitatory
Inhibitory
a
STA
Action Potential
b
Passive Biophysical
Action Potential
c
STA
Passive Biophysical
d
Action Potential
Adaptation Current
e
STA
Adaptation Current
f
MI Score
AP v STA
AP v PB
STA v PB
AC v STA
AC v AP
Co-Clustering Probability
1.0
0.0
Action Potential (AP)
Passive Biophysical (PB)
Adaptation Current (AC)
Spike Triggered Average (STA)