Politecnico di Torino

Master's Degree in Computer Engineering

Graduation Session March 2026

Activity Landscape and Plasticity Signatures (ALPS)

A Data-Driven Bioinformatic Analysis of Neuronal Activity Regimes and Neuronal Plasticity from single-cell Patch-seq data

Supervisor

Roberta Bardini

Co-supervisors

Gianluca Amprimo
Lorenzo Martini
Stefano Di Carlo
Alessandro Savino

Candidate

Serena Canavero

Abstract

Intrinsic neuronal excitability provides a functional proxy of a neuron's activity state and reflects its capacity for adaptation through intrinsic plasticity. This thesis aims to discover and characterize electrophysiological excitability regimes as dynamic functional states by constructing a data-driven excitability landscape directly from patch-clamp recordings and assessing their relationship with neuronal types and subtypes defined by Patch-seq.

Despite the central role of excitability in neuronal computation and plasticity, electrophysiology still lacks a unified characterization of excitability levels that quantifies where a neuron lies within a continuous spectrum of responses when subjected to stimulation. This thesis addresses this gap by constructing a data-driven excitability landscape directly from patch-clamp recordings and by testing whether electrophysiological excitability regimes align with, or cut across, transcriptomically defined neuronal types and subtypes measured with Patch-seq.

At the single-spike level, the Neuronal Spike Shape (NSS) framework is extended by adding additional first-spike action potential features that are relevant to intrinsic plasticity. In particular, the analysis includes descriptors of the after-hyperpolarization phase measured under both long square and ramp stimulation protocols.

To complement this spike-based view, the membrane-potential responses to ramp stimulation are also analyzed as continuous voltage traces. These signals are embedded into low-dimensional state spaces using derivative-based coordinates and Takens delay embeddings. From the resulting representations, a set of geometry-based descriptors is computed, including Poincaré and increment Poincaré features, which characterize aspects of the temporal organization of the voltage trajectory beyond spike shape alone.

Unsupervised clustering of these features reveals distinct excitability regimes that are largely independent of transcriptomic neuronal types and subtypes, supporting the idea that neurons from different molecular identities can occupy shared electrophysiological states and transition among them through intrinsic plasticity. Finally, the relationship between excitability regimes and plasticity is investigated by examining the expression of plasticity-related genes across the inferred excitability landscape.

Overall, this work characterizes neuronal excitability regimes as shared, dynamic functional states rather than fixed cell-type properties and provides a reproducible, electrophysiology-derived landscape for probing intrinsic excitability modulation, with transcriptomic data used for joint analysis on intrinsic and synaptic plasticity markers.

Acknowledgements

A mio fratello

Table of Contents

Chapter 1

Introduction

Neuronal excitability is a fundamental property of the nervous system that determines how neurons transform synaptic and injected inputs into electrical outputs. Through the generation and timing of action potentials, intrinsic excitability shapes information processing, circuit dynamics, and behavioural responses. At the biophysical level, excitability emerges from the interaction of passive membrane properties and voltage-dependent ionic conductances, including sodium, potassium, calcium, and hyperpolarization-activated currents, whose collective dynamics govern spike initiation, waveform morphology, and firing stability [1, 2].

Despite its central importance, intrinsic excitability remains difficult to characterize in a unified way and experimental electrophysiology commonly summarizes neuronal responses using a small set of scalar descriptors such as rheobase current, firing rate, spike width, or adaptation indices. These quantities are informative, but they provide only a partial description of neuronal responses as they compress complex voltage dynamics into isolated measurements and therefore fail to see how the membrane potential evolves during sustained stimulation.

As a consequence, neurons with markedly different intrinsic dynamics may appear similar when described using conventional feature sets. This raises a broader question: how can excitability be represented in a way that captures the temporal structure of neuronal responses rather than reducing them to a few static measurements?

From a theoretical perspective, neuronal firing is inherently dynamical. Conductance-based models describe spike generation as the evolution of a nonlinear system in which membrane voltage interacts with multiple gating variables. In this framework, firing regimes correspond to trajectories in a high-dimensional state space,

and transitions between quiescence, spiking, or bursting arise from bifurcations in the underlying dynamics [3, 4].

However, in electrophysiological recordings only the membrane potential is directly observed and instead the internal state variables remain hidden. But since the recorded voltage trace still reflects the evolution of the underlying dynamical system, the use of time-series and state-space reconstruction methods to extract dynamical information directly from measured voltage signals is thus motivated [5, 6, 7].

This dynamical interpretation is particularly relevant in the context of *intrinsic plasticity*. Unlike synaptic plasticity, which modifies the strength of connections between neurons, intrinsic plasticity alters the neuron's own excitability through regulation of ion-channel expression and kinetics. Such mechanisms enable neurons to transition between different firing regimes while preserving circuit-level stability, suggesting that excitability states may be shared across neurons with distinct molecular identities [8, 9, 10].

Recent large-scale initiatives combining patch-clamp electrophysiology with transcriptomic profiling, most notably Patch-seq, have provided unprecedented multimodal datasets linking neuronal function to molecular cell types [8, 9, 10]. However, transcriptomic classifications are inherently static and do not necessarily reflect the functional state of a neuron at the time of recording. Moreover, electrophysiological properties themselves often cut across transcriptomic categories [11, 12], raising the question of whether intrinsic excitability should be understood primarily as a cell-type property or as a dynamic, state-dependent dimension.

The goal of this thesis is to address this gap by constructing a *data-driven excitability landscape* directly from patch-clamp recordings. Rather than assigning neurons to fixed electrophysiological types, this work aims to identify recurrent excitability regimes as dynamical states and to investigate how these states relate to transcriptomic identity and plasticity-associated gene expression. To this end, we develop an integrated analysis framework that combines interpretable spike-based features with dynamical descriptors computed from whole membrane-potential traces.

This thesis thus focuses on understanding how intrinsic excitability can be characterized within neuronal populations. Its first question is whether distinct excitability regimes can be identified directly from patch-clamp recordings using data-driven approaches. It then asks how these electrophysiological regimes relate to transcriptomic cell types, and whether they align with molecular classifications or instead cut across them. And the last aspect it touches upon concerns the relationship between these excitability regimes and the expression of genes associated with neuronal plasticity.

The remainder of this thesis is organised as follows. Chapter 2 reviews the relevant background on intrinsic excitability, patch-clamp electrophysiology, Patch-seq, electrophysiological classification, and dynamical-systems approaches to neuronal activity. Chapter 3 describes the datasets and the methodological pipeline developed in this work, including quality control, feature engineering, dynamical embeddings, and clustering strategies. Subsequent chapters present the experimental results and their biological interpretation, followed by a discussion of implications and future directions.

Chapter 2

Background and State of the Art

2.1 Intrinsic Neuronal Excitability and Plasticity

Intrinsic neuronal excitability refers to the propensity of a neuron to translate synaptic or injected inputs into changes in membrane potential and, when threshold is reached, into action potentials. This property does not depend on a single variable, but emerges from the interaction between passive membrane properties and the particular combination of ion channels expressed by the cell, including sodium, potassium, calcium, and hyperpolarization-activated currents [1, 2]. In this sense, excitability is a structured physiological property rather than a single scalar quantity. At the same time, it is also degenerate: different combinations of conductances can produce similar firing behaviors, so comparable outputs do not necessarily imply identical underlying mechanisms [13].

A second important point is that excitability is not fixed. Neurons can adjust their responsiveness through activity-dependent and neuromodulatory processes, a phenomenon usually referred to as *intrinsic plasticity*. These changes can involve channel density, subcellular distribution, or kinetics, and can modify spike threshold, firing gain, adaptation, and afterhyperpolarization dynamics without directly changing synaptic weights [8, 14]. In this way, intrinsic plasticity complements synaptic plasticity and contributes to learning, development, and homeostatic regulation of circuit activity [8, 10].

Within this broader picture, the afterhyperpolarization (AHP) deserves special

attention. The AHP is not simply a passive tail of the spike: it is one of the variables that helps shape how readily a neuron can fire again, and it therefore contributes directly to intrinsic excitability. Several studies have described the post-burst AHP as an important determinant of neuronal excitability, and modulation of AHP-related currents has repeatedly been linked to changes in learning-related and state-dependent firing behavior [15]. Even the fast component of the AHP can influence gain and spike output in non-trivial ways, indicating that post-spike repolarization and recovery are functionally informative rather than secondary waveform details [16].

This point is relevant for electrophysiological phenotyping. Recent waveform-centered approaches such as the Neuronal Spike Shapes (NSS) framework have shown that compact action-potential descriptors can recover meaningful excitability structure from patch-clamp data [17]. At the same time, the importance of AHP-related mechanisms suggests that excitability cannot always be captured fully by features focused mainly on spike onset and peak geometry. Including descriptors of the post-spike phase may therefore help recover aspects of intrinsic state that are especially relevant for adaptation, recovery, and plasticity-related changes in responsiveness.

From a functional point of view, these observations indicate that excitability may be better interpreted as a dynamic functional dimension rather than a fixed cellular trait. Neurons can shift between lower and higher excitability states depending on recent activity, neuromodulatory context, and on the conductances that regulate recovery and firing stability. This perspective motivates approaches that represent excitability using structured feature spaces rather than relying only on single scalar measurements so as to better capture state-dependent variation in neuronal responses.

2.2 Patch-Clamp Electrophysiology as a Probe of Excitability

Patch-clamp electrophysiology remains the reference technique for studying the electrical behaviour of individual neurons with millisecond temporal precision [18]. In the whole-cell current-clamp configuration, controlled currents are injected through a glass pipette while the membrane potential is recorded. This configuration makes it possible to observe both subthreshold voltage dynamics and the generation of action potentials, providing direct information about how a neuron converts input current into electrical output [1].

Since the membrane potential reflects the combined effect of the underlying ionic conductances, patch-clamp recordings provide an experimental window onto intrinsic excitability. Features such as spike threshold, firing frequency, spike waveform, and adaptation patterns arise from the interaction of sodium, potassium, calcium, and other voltage-dependent currents shaping the membrane response [1, 19]. For this reason, systematic current-injection protocols are widely used to probe how neurons respond to controlled perturbations of their membrane potential.

The two stimulation paradigms common in this context are step-current protocols (often referred to as long-square stimuli), that inject a constant depolarizing current and are typically used to estimate properties such as rheobase, firing-rate gain, and spike-frequency adaptation, and ramp-current protocols that instead increase the injected current gradually over time, allowing threshold dynamics and spike initiation mechanisms to be probed under slowly varying drive [20]. These large datasets have also revealed an important limitation of traditional electrophysiological descriptors as many commonly used features reduce complex voltage traces to a small number of scalar measurements, which can obscure meaningful differences in the temporal organization of neuronal responses. Large-scale analyses in fact have shown that electrophysiological properties often vary continuously across neurons rather than forming sharply separated categories, suggesting that richer representations of voltage dynamics may be required to capture the full diversity of neuronal responses [11, 12].

In recent years several large-scale initiatives have extended patch-clamp analysis beyond individual laboratories to standardized datasets containing hundreds or thousands of neurons recorded under comparable protocols. For example, the Allen Institute Cell Types program has produced large collections of standardized patch-clamp recordings across multiple cortical regions and neuronal classes [21]. Datasets as this have made it possible to systematically compare electrophysiological properties across molecularly defined cell types and to explore population-level variability in neuronal dynamics.

2.3 Patch-seq and Multimodal Cell-Type Classification

The development of Patch-seq has provided a powerful framework for linking electrophysiological properties with molecular identity at the level of individual neurons. Patch-seq combines whole-cell patch-clamp recordings with single-cell RNA sequencing, and often with morphological reconstruction, allowing functional,

molecular, and anatomical information to be obtained from the same cell [22, 23].

Neuronal taxonomies have been refined significantly thanks to this multimodal approach. Analyses of large Patch-seq datasets have shown that transcriptomic subclasses are often associated with characteristic electrophysiological phenotypes, and have enabled systematic comparisons between gene expression profiles and functional properties across cortical cell types [22]. These studies have thus helped establish a framework in which molecular identity, morphology, and physiology can be integrated to define neuronal cell classes.

However, the relationship between transcriptomic identity and electrophysiological behaviour is not strictly one-to-one. Even within well-defined transcriptomic classes, neurons can exhibit substantial variability in firing patterns, spike waveform properties, and adaptation dynamics [12]. Conversely, electrophysiological features often vary continuously across neurons and may span multiple molecular subclasses [11]. Together these observations indicate that electrophysiological state and transcriptomic identity likely capture partially distinct aspects of neuronal organisation.

Recent studies have also explored whether transcriptomic data may contain signatures of neuronal activity or plasticity state. As an example machine-learning approaches like the NEUROeSTIMator [24] have been proposed to estimate neuronal activation from genome-wide expression profiles, revealing activity-dependent transcriptional programs that are conserved across cell types and experimental conditions. Such studies highlight the potential of multimodal datasets to investigate not only stable cell identity but also state-dependent aspects of neuronal function.

In this context, Patch-seq datasets offer a particularly useful framework to investigate how electrophysiological excitability relates to molecular identity and activity-dependent transcriptional programs. Instead of assuming a direct mapping between electrophysiological properties and transcriptomic classes, excitability can be interpreted as a dynamic functional dimension that may cut across molecular cell types. To explore this possibility analytical approaches capable of identifying excitability structure directly from voltage recordings and then comparing these electrophysiological regimes with transcriptomic annotations are needed.

2.4 Electrophysiological Typing and Spike-Based Feature Representations

A common way to analyze patch-clamp recordings is to convert voltage traces into a set of hand-crafted electrophysiological features. Typical examples include spike threshold, spike amplitude, spike width, firing rate, adaptation indices, and afterhyperpolarization-related measures [25, 21]. One reason these representations remain widely used is that they are interpretable: many of the extracted quantities can be related, at least approximately, to identifiable membrane mechanisms. For instance, repolarization and afterhyperpolarization features are shaped in part by potassium conductances and therefore provide information that is directly relevant to intrinsic excitability [1, 8].

But the feature-based electrophysiological approach has its limitations since reducing a voltage response to a small number of scalar measurements can hide differences in how the response unfolds over time; particularly when neurons occupy intermediate or context-dependent regimes rather than forming sharply separated electrophysiological classes. Moreover, this has been highlighted in research conducted on large-scale datasets where electrophysiological properties often vary along continuous axes and only partially align with transcriptomic classification [11, 12].

Within this broader literature, the Neuronal Spike Shapes (NSS) framework is particularly relevant as a starting point for the present work. NSS was introduced as a waveform-centered approach for exploring neuronal excitability states from action-potential shape, using a simple triangular representation of the first spike together with a compact set of derived electrophysiological descriptors [17]. In murine cortical interneurons, NSS showed that first-spike waveform features can recover biologically meaningful excitability partitions, including a fast-spiking state supported by both electrophysiological and transcriptomic evidence. In that study, voltage-gated potassium channel-related signatures were enriched in the identified partitions, further supporting the idea that compact waveform descriptors can capture functionally relevant differences in excitability [17].

NSS is therefore an important precedent, because it shifts the focus from cell-type classification to the identification of excitability states. However, it remains centered on a compact representation of the first action potential, with most of the emphasis placed on threshold, peak, and fast-trough geometry. This leaves less room for describing the broader post-spike recovery dynamics that also contribute to neuronal responsiveness. That omission matters because the afterhyperpolarization is not simply a passive tail of the spike: several studies have shown that AHP-related mechanisms influence spike-frequency adaptation, recovery from firing, and overall

intrinsic excitability [8, 15].

For this reason, extending spike-based phenotyping beyond the compact NSS representation is already useful at the single-spike level. In particular, adding descriptors of the post-spike phase, including the slow trough and AHP-related recovery, can capture aspects of excitability that are especially relevant for adaptation and plasticity-related modulation. More broadly, the limitations of first-spike summaries motivate the transition to representations that also preserve the temporal organization of the response beyond the isolated spike waveform.

2.5 A Dynamical-Systems View of Neuronal Firing

From a theoretical standpoint, neuronal firing is naturally described as a dynamical process. In conductance-based models, spike generation does not arise from a single threshold variable, but from the interaction of membrane voltage with multiple gating variables whose evolution defines a nonlinear dynamical system. In this framework, qualitatively distinct response patterns correspond to different regions of state space and to different dynamical regimes of the system. The Hodgkin-Huxley model provided the classical example of this view by showing how nonlinear interactions between voltage and gating variables give rise to threshold behavior and repetitive firing [3], while reduced models such as Morris-Lecar made it possible to analyze transitions between quiescence, tonic firing, and oscillatory behavior in terms of bifurcation structure [4].

Later work extended this dynamical perspective by showing that a wide range of neuronal firing behaviors can be organized in terms of a small number of canonical dynamical mechanisms [19]. Detailed conductance-based studies of cortical interneurons further demonstrated that modest changes in a few intrinsic parameters can move the same model neuron between qualitatively different regimes, such as quiescent, tonic, delayed, or stuttering firing [26]. These results highlight that excitability is not a single scalar quantity but rather reflects the position of a neuron within a structured dynamical landscape determined by its membrane conductances. More recent analyses have reinforced this view by emphasizing that neuronal excitability is a resilient property emerging from the interaction of multiple conductances operating across different timescales [27]. These results suggest that excitability states correspond to stable or metastable regimes of an underlying latent dynamical system, rather than to isolated scalar descriptors.

2.6 Nonlinear Time-Series Analysis and State-Space Reconstruction

In experiments, however, the full set of internal state variables is not directly accessible. Patch-clamp recordings provide the membrane potential, while most gating variables and other hidden degrees of freedom remain unobserved. This is precisely the setting in which tools from nonlinear time-series analysis become useful. Delay-coordinate reconstruction offers a principled way to study the geometry of an underlying dynamical system using a single observed signal. Under Takens-style assumptions, a scalar time series can be mapped into a higher-dimensional delay space in a way that preserves relevant properties of the attractor geometry [5, 6]. Throughout this work we refer to these embeddings as reconstructed state spaces, since the full set of dynamical variables governing the system is not directly observable.

For the present problem, delay embeddings are appealing because they treat the voltage trace itself as the observable output of an underlying dynamical process. Derivative-based embeddings provide a complementary view. Rather than reconstructing the system through delayed copies of the signal, they represent the trajectory through local temporal change, for example by combining the voltage with its first and second derivatives. This tends to emphasize rapid transitions, spike onset, and local curvature of the trajectory, and is well aligned with the general framework of observed-dynamics reconstruction discussed in the nonlinear-systems literature [7].

After reconstructing a trajectory in a suitable state space, its geometry can be summarized through several possible low-dimensional descriptors. One widely used example is the Poincaré plot that maps successive samples of a time series into a two-dimensional plane. It was originally developed in the context of heart-rate variability analysis and is able to effectively provide compact geometric summaries of short-timescale variability that characterize structured changes in temporal dynamics and autonomic regulation [28, 29].

There exist possible extensions of the basic construction like the increment-based variants that instead plot successive differences $(\Delta x_n, \Delta x_{n+1})$ rather than the raw signal values. These representations highlight rapid transitions while reducing the influence of slow baseline drifts since they emphasize local changes instead of absolute levels, and have been used to study physiological variability in heart-rate dynamics [30, 28]. Also, recently, second-order extensions have been introduced to capture additional temporal structure beyond the standard first-order version [31].

More broadly, geometric analyses of reconstructed trajectories have been applied in several biomedical contexts. For instance, in speech analysis for Parkinson's disease, state-space reconstruction has been used to study the attractor structure generated by speech dynamics. Changes in the geometry and temporal evolution of these reconstructed attractors have been shown to correlate with disease progression, providing dynamical markers that reflect alterations in motor control affecting speech production [32].

2.7 Limitations

Although these ideas are attractive, they cannot be imported into patch-clamp analysis without modification. The main reason is that the objects being analyzed are fundamentally different. HRV methods are typically built around sequences of inter-beat intervals, that is, series of discrete event timings. Patch-clamp recordings instead provide continuous voltage trajectories that include both subthreshold dynamics and detailed spike waveforms. As a result, the signal carries information not only in the timing between spikes, but also in the shape of the trajectory between and around those events.

Another important difference is that neuronal responses under current injection are often strongly nonstationary and this means that the voltage trace may change during the ramp or other sustained step because of adaptation, recruitment of slower conductances, or gradual shifts in threshold and recovery dynamics. Conventional scalar measures of intrinsic excitability can already become unreliable when the input context is changed, as shown by work comparing standard current-step protocols with more realistic synaptic drive [14], so this must be taken into account here since more complex measures are considered. This insight could suggest that dynamical descriptors for patch-clamp data should be sensitive to temporal evolution rather than assuming a single stationary regime throughout the recording.

And the last difference is that spike waveform and spike timing are coupled in neurons since they are both shaped by the same membrane conductances, refractory effects, and recovery dynamics. Because of this inter-spike intervals and isolated waveform features should not be analyzed alone as this might give back an incomplete picture of the response. A possible strategy used here is to work directly on the voltage trajectory and to derive descriptors that preserve both local waveform-related structure and the broader temporal organization of the response under the controlled stimulation.

Taken together, these considerations suggest that nonlinear time-series tools are best

used here as a source of representation and feature design, not as a plug-in transfer from HRV. The underlying idea is the same, to extract informative geometry from a trajectory, but the descriptors need to be adapted to the electrophysiological setting, where the observed object is a continuous, stimulus-driven membrane potential trace.

2.8 Positioning of This Work

The previous sections highlight several complementary perspectives on neuronal electrophysiology. Intrinsic excitability is a dynamic property shaped by interacting membrane conductances (section 2.1), and patch-clamp recordings provide direct access to these dynamics under controlled stimulation (section 2.2). At the same time, large-scale Patch-seq datasets make it possible to relate electrophysiological behaviour to molecular identity (section 2.3). However, most analyses of these recordings still rely on compact sets of hand-crafted electrophysiological features (section 2.4), which summarize voltage responses through a small number of scalar descriptors.

While such features are informative, they compress the temporal structure of the voltage trajectory and may obscure the dynamical organization of neuronal responses. This limitation is particularly relevant in light of the dynamical-systems view of neuronal firing (section 2.5), which emphasizes that excitability reflects the position of a neuron within an underlying dynamical landscape. Methods from nonlinear time-series analysis provide a natural framework for studying such systems from observed signals (section 2.6), but their direct application to patch-clamp data requires adaptations that account for the continuous, stimulus-driven nature of membrane potential traces (section 2.7).

The present work builds on these observations by approaching patch-clamp recordings as trajectories generated by an underlying dynamical system. Instead of relying solely on isolated spike descriptors, it explores geometric representations of reconstructed voltage dynamics that preserve both waveform structure and temporal organization. Within this framework, electrophysiological responses can be characterized through dynamical descriptors derived from the voltage trajectory, providing an alternative representation of intrinsic excitability that can be compared with transcriptomic annotations in multimodal datasets.

Chapter 3

Methods

3.1 Datasets

This study analyzes two publicly available datasets of intracellular electrophysiology from mouse visual cortex, chosen to provide both a high-quality electrophysiological reference dataset and a multimodal dataset enabling integration with transcriptomic information.

The first dataset is the PatchClampDataset, derived from the Allen Institute mouse visual cortex cell types resource and described by Gouwens and colleagues in their classification of electrophysiological and morphological neuron types in adult mouse visual cortex [21]. This dataset contains whole-cell patch-clamp recordings from **1,961 neurons**, together with morphological reconstructions for a subset of cells. In the present work, this dataset serves as the primary electrophysiology-only benchmark, because it provides large-scale standardized intrinsic physiology recordings acquired under controlled stimulation protocols and processed within a well-established experimental framework.

The second dataset is a Patch-seq mouse visual cortex dataset, corresponding to the multimodal resource associated with the study *Integrated Morphoelectric and Transcriptomic Classification of Cortical GABAergic Cells* [22]. This dataset is available through the DANDI archive and combines intracellular electrophysiology with single-cell transcriptomic measurements, enabling direct comparison between electrophysiological excitability structure and molecular identity within the same neurons. The Patch-seq dataset allows electrophysiological states to be interpreted in relation to transcriptomic subclass labels and to the expression of genes associated with intrinsic and synaptic plasticity.

For both datasets, electrophysiological recordings were obtained from the DANDI archive as NWB (Neurodata Without Borders) files [33]. Download was performed programmatically through the DANDI API. Raw NWB files were then parsed locally using `pynwb`, and sweep-level voltage/current traces were extracted directly from the NWB objects. To maximize compatibility across files, a multi-step extraction strategy was implemented, prioritizing: (i) `nwb.sweep_table` with unit/type-based identification of stimulus and response series, (ii) legacy name- and description-based heuristics within the sweep table, (iii) `nwb.intracellular_recordings` for NWB2 intracellular ephys tables, and (iv) a fallback scan of `acquisition` and `stimulus` objects grouped by `sweep_number`. This design ensured robust recovery of current and voltage traces even across heterogeneous NWB encodings.

After extraction, each sweep was stored in a standardized NumPy-based representation (`.npz`) containing the time axis, membrane potential, injected current, and associated metadata. The pipeline also detected the stimulation window directly from the current trace using thresholds based on the median absolute deviation [25]. This step allowed the sweeps to be cropped and aligned to stimulus onset in a consistent way across datasets.

Although the two datasets are used differently in the later stages of the analysis, their electrophysiological preprocessing starts from the same NWB files and follows the same general extraction procedure.

Only recordings acquired in whole-cell current-clamp configuration were considered. The analysis focused on stimulation paradigms designed to probe intrinsic excitability under controlled current injection. In the PatchClampDataset, both long square current injections and ramp current injections were analyzed, allowing comparison between discrete near-threshold step stimulation and continuously increasing depolarizing drive. In the Patch-seq dataset, the analysis focused on ramp current injections, which were the protocol consistently integrated into both the dynamical electrophysiology pipeline and the downstream transcriptomic comparison.

Together these two datasets allow intrinsic excitability to be studied both as a functional electrophysiological phenotype and as a property that can be related, but not reduced, to molecular cell identity.

3.2 Overview of the Analysis Pipeline

The methodological pipeline developed in this thesis aims to characterize intrinsic neuronal excitability as a dynamic functional state rather than a fixed cellular attribute. The pipeline consists of two complementary analysis stages:

1. **First spike-based analysis**, extending the Neuronal Spike Shape (NSS) framework through additional waveform features explicitly linked to intrinsic plasticity.

2. **Dynamical analysis of membrane-potential trajectories**, in which whole voltage responses are treated as continuous time series and embedded into low-dimensional state spaces.

These two stages both operate on the raw voltage traces extracted from the NWB files, but before feature extraction is carried out, a sweep-level quality control procedure is applied to ensure that the analyzed signals are both reproducible and biologically meaningful.

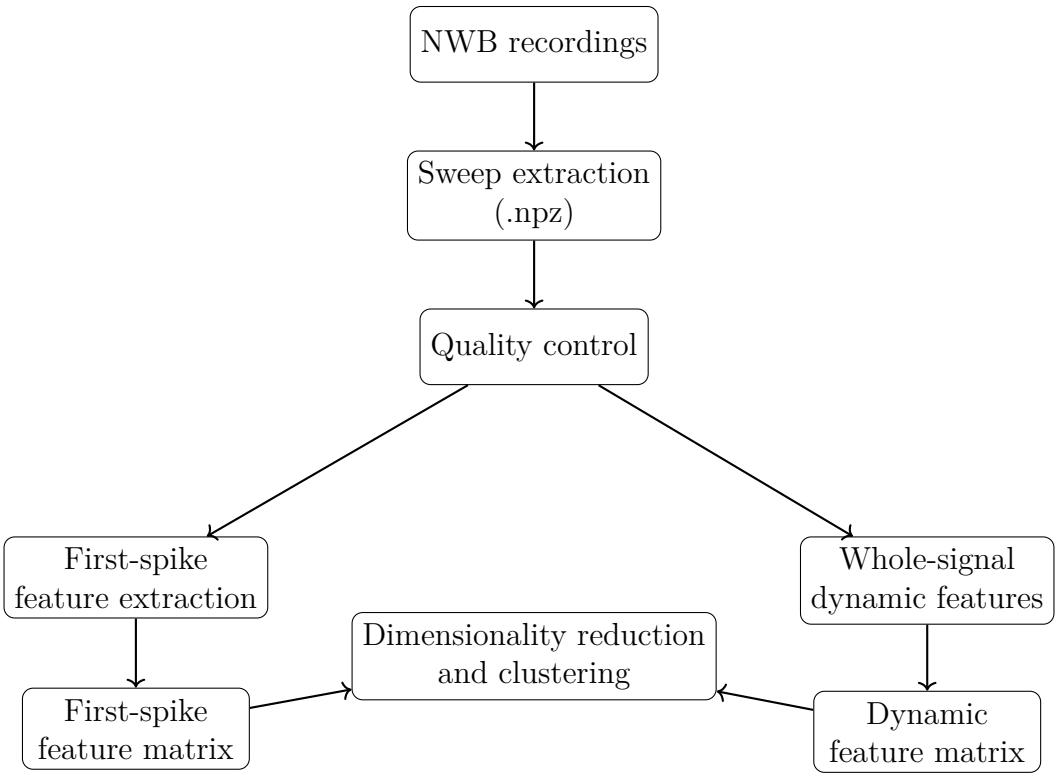


Figure 3.1: Overview of the electrophysiological analysis pipeline.

3.3 Electrophysiological Data Extraction and Pre processing

Raw intracellular recordings were processed through a custom Python pipeline designed to convert NWB files into a uniform intermediate representation suitable for large-scale electrophysiological analysis. Because the two datasets originate from different experimental pipelines and contain heterogeneous metadata structures, the preprocessing framework was designed to be robust to multiple NWB encodings while producing a consistent downstream representation.

Electrophysiological recordings were parsed directly from the NWB files using the `pynwb` library. Voltage and current time series were extracted sweep by sweep through a multi-stage parsing strategy designed to accommodate structural differences across NWB files. The pipeline first attempted to recover sweeps from the `nwb.sweep_table`, selecting stimulus and response series based on their physical units and object types. If this information was unavailable or incomplete, additional heuristics based on series names and stimulus descriptions were applied. When these approaches failed, sweeps were reconstructed from the NWB2 `intracellular_recordings` table, and as a final fallback the pipeline scanned the `acquisition` and `stimulus` groups, grouping time series according to their associated `sweep_number`. This hierarchical extraction strategy ensured that current stimulus and membrane-potential response traces could be recovered even from NWB files with heterogeneous internal organization.

Each extracted sweep was converted into a standardized compressed representation stored as a NumPy `.npz` file. This intermediate format detached downstream analysis from dataset-specific NWB structures while providing uniform access to electrophysiological traces. Each `.npz` file contained three primary numerical arrays:

- time axis t (seconds)
- membrane potential $v(t)$ (millivolts)
- injected current $I(t)$ (picoamperes)

Metadata associated with each sweep were stored in serialized form and included the specimen identifier, sweep identifier, NWB file path, stimulus label, session identifier, subject identifier, extraction method, and species. Current and voltage signals were converted to physical units using the conversion factors specified in

the NWB file. When the sampling frequency was not explicitly provided, it was inferred from the median temporal spacing between consecutive time samples.

Voltage traces were first baseline-corrected and aligned to the stimulation onset to obtain comparable signals across recordings. Subtracting the baseline reduces any differences in resting membrane potential across neurons, so that the following feature extraction focuses on stimulus-driven dynamics instead of absolute voltage offsets [25].

The baseline voltage was estimated as the median membrane potential within an initial pre-stimulus window of length n_0:

$$V_{\text{baseline}} = \text{median}(v[0 : n_0]).$$

The corrected signal was then defined as

$$v_0(t) = v(t) - V_{\text{baseline}}.$$

The traces were then aligned to the stimulus onset using the onset index stored in the sweep metadata, so that $t = 0$ corresponds with the beginning of the stimulation:

$$v_{\text{aligned}}(t) = v_0(t + t_{\text{onset}}).$$

The stimulation interval was detected directly from the injected current trace in order to standardize the analysis window across recordings. The onset and offset of the stimulation were identified using thresholds derived from the median absolute deviation of the baseline current signal which allowed reliable detection of stimulation epochs even when the stimulus annotations were incomplete. When the stimulation interval was identified, the voltage and current traces were cropped around it, with an additional padding to preserve both pre- and post-stimulus dynamics. These cropped sweeps were then used for all downstream electrophysiological analyses, ensuring consistent temporal alignment of stimulation onset across neurons.

An additional quality-control step performed after extraction and cropping removed sweeps that failed basic integrity checks, such as corrupted traces, missing stimulation windows, or invalid sampling information. The passing sweeps were recorded in specimen-level index files (`index_npz_qc.csv`) storing the paths and metadata

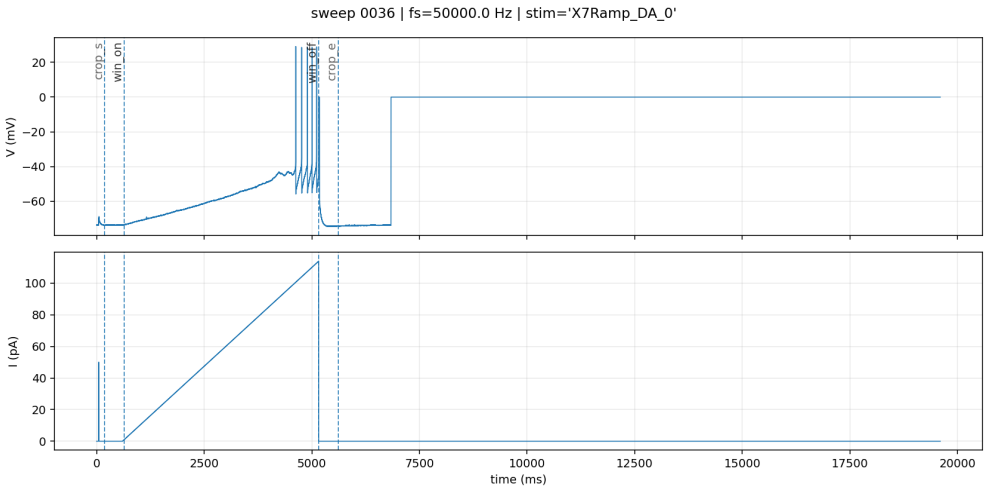


Figure 3.2: Example of cropped window for patchseq specimen_id 750189136, sweep 36

of all QC-passing sweeps and providing the entry point for the electrophysiological feature-extraction pipeline.

The preprocessing stage produced a structured dataset of standardized sweep representations organized by dataset, stimulus type, and specimen identifier. These intermediate files form the basis for the electrophysiological feature extraction procedures described in the following sections and ensure full traceability from the original NWB recordings to the final feature matrices used in the clustering analysis.

3.4 Sweep-Level Quality Control

Before feature extraction, a sweep-level quality control (QC) procedure was applied in order to retain only recordings that preserved biologically interpretable post-stimulus dynamics. This step operated directly on the cropped `.npz` sweeps and generated, for each specimen, three outputs:

- a table containing the computed QC metrics for each sweep;
- a table reporting QC pass/fail decisions;
- a filtered sweep index (`index_npz_qc.csv`) containing only the sweeps accepted for downstream analysis.

The sweeps that were rejected were moved to a dedicated `QC_rejected` subfolder thus ensuring that excluded recordings stayed traceable while also preventing their accidental inclusion in later stages of the analysis pipeline.

The carried-out quality control relied on a set of metrics computed from the membrane-potential trace after the stimulus onset and before it the trace was baseline-corrected. The baseline voltage was estimated as the median membrane potential within an initial pre-onset segment and then subtracted from the signal. This was done because working with a baseline-corrected trace reduces offsets between sweeps and retains the temporal structure of the neuronal response.

In order to assess signal integrity and dynamical content several complementary quantities were calculated:

- **Post-onset duration**, measuring the signal that remains available after the detected stimulus onset. Short durations typically indicate truncated or poorly cropped recordings.

- **Robust voltage range**, measuring the difference between the 95th and 5th percentiles of the post-onset voltage distribution. This statistic is less affected by isolated outliers and provides a more stable estimate of the signal's dynamical excursion with respect to a simple max–min range [6].
- **Windowed activity measures**: each post-onset trace was divided into four windows of equal duration and the robust voltage range was computed inside each window.
- **Late-to-early activity ratio**, computed as the ratio between the robust voltage range in the last window and that in the first window, helping with the spotting of traces in which the dynamical activity progressively decreases during the stimulation.

The QC script also calculated several additional diagnostic quantities that were kept for diagnostic inspection but were not used in the rejection rules directly; among these drift measures, tail statistics and indicators of spike occurrence.

Since the long-square and ramp stimuli protocols probe neuronal dynamics in different ways, different rejection criteria were applied to them.

Long square stimuli: for long-square current injections, the main failure mode corresponds to truncated recordings in which too little signal remains after stimulus onset. Sweeps were rejected when the post-onset duration was shorter than 100 ms. No amplitude-based rejection criterion was applied at this stage so that valid subthreshold responses were preserved.

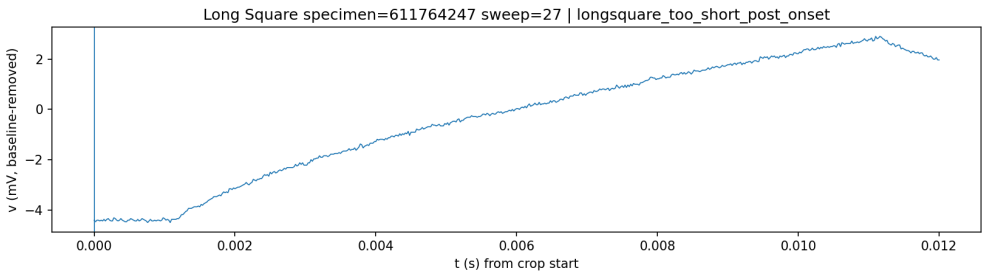


Figure 3.3: Long-square sweep rejected during quality control due to insufficient post-onset duration.

Ramp stimuli: Ramp sweeps required stricter filtering because the analysis depends on the full temporal evolution of the membrane response. A ramp sweep

was rejected if any of the following conditions were met:

- post-onset duration shorter than 200 ms;
- robust voltage range smaller than 6 mV;
- evidence of progressive loss of activity, defined by both
  - late-window activity smaller than 2 mV, and
  - a late-to-early activity ratio smaller than 0.3.

Sweeps that are too short, nearly flat, or that lose meaningful dynamics during the ramp stimulation were thus removed.

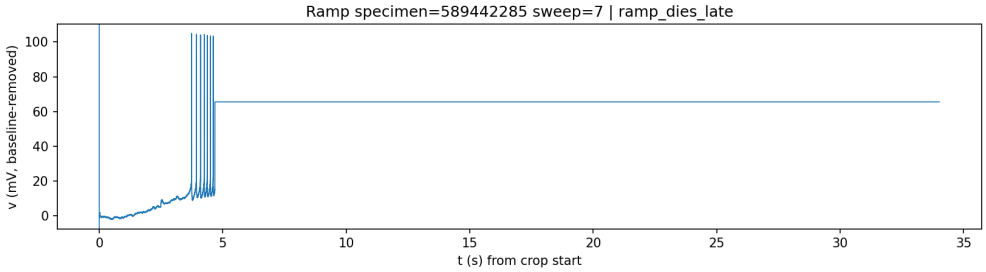


Figure 3.4: Example of a ramp sweep rejected during quality control. The voltage trace shows very limited dynamical activity during the stimulation period.

This quality-control stage produced the set of valid sweeps used for all subsequent first-spike analyses. The pipeline preserves reproducibility and allows every cell-level descriptor to be traced back to the exact subset of sweeps that contributed to it because it explicitly separates extraction, QC, and feature computation.

3.5 FirstSpike Electrophysiological Feature Analysis

To characterize intrinsic excitability at spike onset, a first-spike-centered electrophysiological analysis was performed on the quality-controlled sweeps. This stage transformed collections of valid voltage traces into per-cell electrophysiological feature vectors describing the morphology and short-timescale dynamics of the first evoked action potential. Conceptually, the analysis extends the Neuronal Spike Shape (NSS) framework by combining classical first-spike waveform descriptors with additional quantities designed to capture afterhyperpolarization dynamics and recovery processes.

The direct input to this stage consisted of the quality-controlled cropped `.npz` sweeps referenced in specimen-level `index_npz_qc.csv` files. Each accepted `.npz` file contained the time axis (`t_s`), membrane potential trace (`v_mV`), injected current trace (`i_pA`), and serialized metadata. These arrays formed the basis for all subsequent spike detection and waveform measurements.

For each valid sweep, spike detection and initial waveform extraction were performed using one of two supported backends:

- **AllenSDK backend:** spike features were extracted using the AllenSDK class `EphysSweepFeatureExtractor`[25]. This routine identifies standard spike landmarks and reports descriptors such as threshold time, threshold voltage, threshold index, spike width, and the upstroke/downstroke ratio.
- **IPFX backend:** spike features were alternatively computed with `Spike FeatureExtractor`. Detection followed the Allen-style derivativebased scheme, using parameters including a derivative cutoff (`dv_cutoff = 20`), threshold fraction (`thresh_frac = 0.05`), a minimum peak criterion (`min_peak = -30 mV`), and temporal filtering of the voltage trace before feature extraction.

Feature extraction was applied directly to the cropped voltage/current traces read from the `.npz` files. The stimulation interval was defined from the detected onset

index to the end of the cropped trace, so that ramp responses were analyzed over their full post-onset evolution. This choice is consistent with the use of ramps as progressively increasing depolarizing stimuli rather than fixed-duration current steps.

Spike detection provided the basis for computing several sweep-level electrophysiological features. The following quantities were extracted for each sweep:

- the number of spikes emitted during the stimulation interval;
- the firing rate over the stimulation interval;
- the latency of the first spike relative to stimulus onset;
- the injected current at the first spike threshold crossing.

Whenever a sweep had at least two spikes also mean inter-spike interval, first inter-spike interval, last inter-spike interval and adaptation index were evaluated.

The adaptation index specifically was derived from the sequence of spike threshold times and was defined as the normalized difference between the last and first inter-spike intervals, providing a simple measure of how much firing slowed during the sweep [34].

The pipeline enriched the spike table with **derived waveform features** computed directly from the membrane potential traces by estimating the needed features via post-processing of the detected spike landmarks.

For each spike, a **slow trough** was identified as the minimum membrane potential occurring after the spike peak within a defined search interval. The search started a short time after the peak (5 ms in the implementation) and extended either to the next spike threshold or to the end of the stimulation. This quantity was introduced to capture the slower phase of the afterhyperpolarization beyond the canonical fast trough [8].

The pipeline also computed, whenever both fast-trough and slow-trough landmarks were available, a metric that provides a compact descriptor of post-spike recovery dynamics, that is to say the **AHP slope**, defined as the slope between the fast trough and the slow trough.

Another custom component of the pipeline estimated the **resting potential during stimulation** storing it in the code as `v_rest_stim_mV`. Importantly, this is *not* the pre-stimulus baseline and is instead defined as the median membrane

potential over the stimulation interval after excluding time windows surrounding each spike and its associated afterhyperpolarization. For each spike, samples were excluded from 2 ms before threshold to 2 ms after the slow trough. If the slow trough could not be estimated, the exclusion interval fell back to a short post-peak window. The in-stimulus resting potential was then defined as the median of the remaining samples within the stimulation epoch. This definition is particularly relevant for ramp responses where the effective operating membrane potential during stimulation may significantly differ from the pre-stimulus baseline.

The pipeline computed several additional waveform features using this in-stimulus resting potential, among which:

- **AP area**: the integral of voltage above the in-stimulus resting level between spike threshold and fast trough;
- **AHP area**: the integral of voltage below the same resting level between fast trough and the next spike threshold, or the end of stimulation for the last spike;
- **threshold-to-peak amplitude**: the voltage difference between spike threshold and spike peak;
- **AHP depth**: the difference between the in-stimulus resting potential and the slow-trough voltage.

These quantities extend the standard NSS representation by incorporating information about repolarization, hyperpolarization, and recovery, rather than only spike shape in a narrow sense.

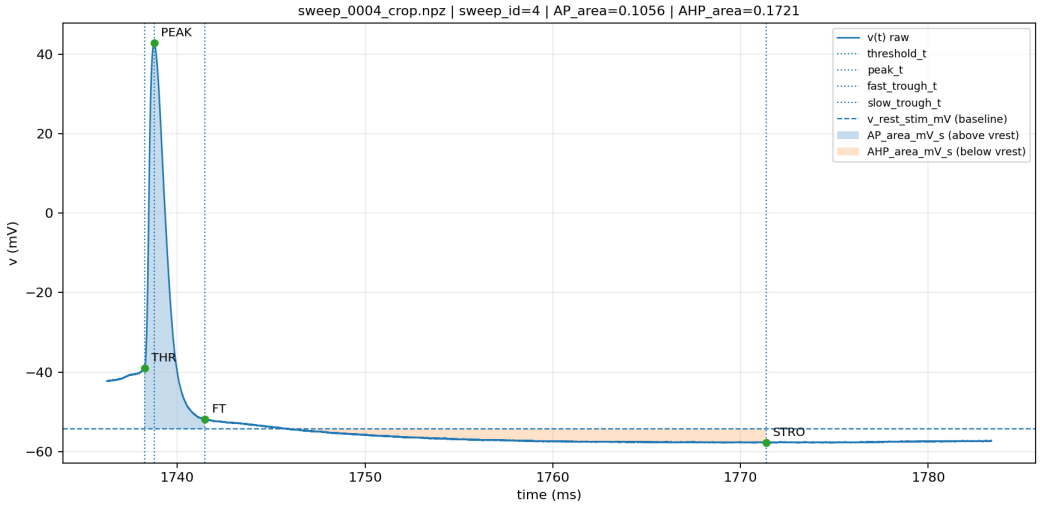
After the enriched spike table had been generated, an NSS-like spike representation was computed for each detected spike. This representation included waveform descriptors such as:

- spike width;
- spike height;
- depth of repolarization (`dV_deep`);
- voltage ratio (`dV_ratio`);
- slope-based repolarization measure (`Slope_deep`);

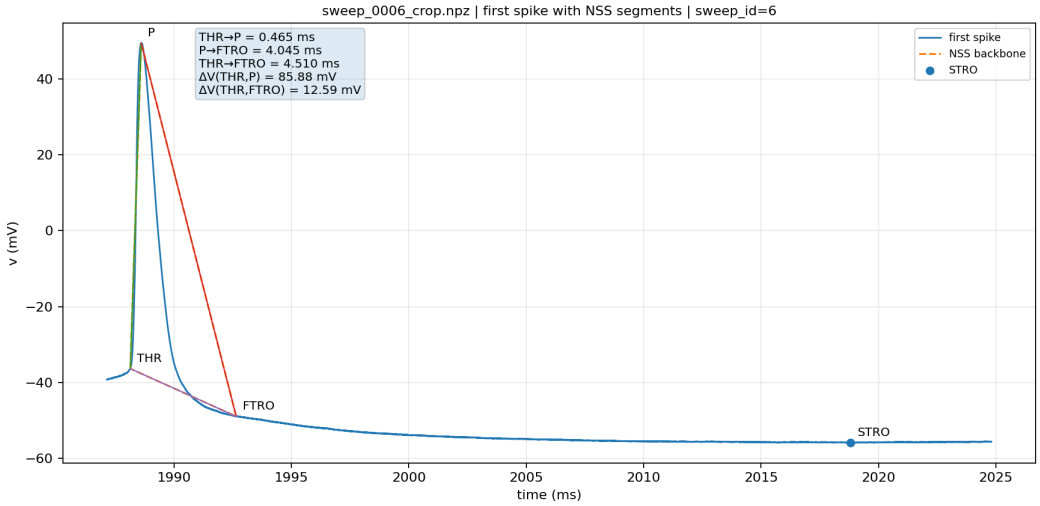
- upstroke/downstroke ratio.

In the implementation, these features were derived from threshold, peak, and fast-trough landmarks and can be appreciated in Figure3.5. For example, `Width` was computed from the temporal separation between the midpoints of the upstroke and downstroke phases, `Height` as the absolute peak-to-fast-trough voltage excursion, `dV_deep` as the absolute threshold-to-fast-trough drop, `dV_ratio` as the ratio between threshold-to-peak and peak-to-fast-trough excursions, and `Slope_deep` as the absolute voltage drop from threshold to fast trough divided by the corresponding time interval. To these traditional NSS variables, the present implementation added the extended AHP-related and recovery-related descriptors described above, including slow-trough-derived measures, AHP depth, AHP slope, in-stimulus resting potential, AP area, AHP area, and threshold-to-peak amplitude. The resulting file, `nss_spike_features.csv`, therefore preserves compatibility with the original first-spike perspective while broadening the physiological information encoded in the spike representation.

At the cell level, each neuron was represented by a single reference first spike extracted from a representative sweep. The selection rule depended on the stimulation protocol, but was implemented consistently across datasets at the export stage. For ramp stimulation, the reference sweep was selected among all spiking sweeps as the one with the lowest injected current at the moment of the first spike threshold crossing (`ramp_level_pA_at_first_spike`). When multiple sweeps shared the same value, the earliest sweep in acquisition order (`sweep_id`) was selected as a deterministic tie-breaker. This rule was used in the native *PatchSeq* branch and, for *PatchClamp*, in the stimulus-specific exporter `export_patchclamp_ep_row.py`. If this quantity was unavailable, the earliest spiking sweep by `sweep_id` was used as a fallback. This selection rule systematically favors sweeps in which spiking occurs at the lowest injected current level, thereby anchoring the representative sweep to a near-threshold firing regime. As a consequence, neurons are compared under comparable minimally suprathreshold conditions rather than under stronger depolarizing drive, reducing variability associated with stimulus intensity. For long-square stimulation in *PatchClamp*, the reference sweep was instead chosen as the rheobase sweep, preferentially the sweep for which `stim_level_pA` matched `rheobase_pA` and that contained at least one spike; if no exact match was available, the spiking sweep with the lowest stimulus amplitude was selected. Once the reference sweep had been identified, the representative spike was defined as the first spike in temporal order within that sweep, that is, the spike with the smallest threshold time (`threshold_t`). This procedure yielded a per-cell representation anchored to a minimally suprathreshold, rheobase-like excitability condition, while preserving protocol-specific differences between ramp and long-square stimulation.



(a) Detail of the first spike highlighting the main electrophysiological landmarks and associated areas: *threshold_t*, *peak_t*, *fast_trough_t*, *slow_trough_t*, *v_rest_stim_mV*, AP area, and AHP area.



(b) First spike in the NSS geometric framework showing the segments linking the threshold (THR), peak (P), and fast trough (FTRO) together with the slow trough (STRO).

Figure 3.5: Visualization of the first action potential extracted from a representative sweep. The top panel highlights the main spike landmarks and the areas computed with respect to the resting membrane potential during stimulation, while the bottom panel shows the corresponding NSS-based geometric representation of the spike waveform.

Once the reference sweep had been identified, the first spike within that sweep was used as the representative event for that neuron. All raw waveform measurements and NSS-derived descriptors associated with this spike were exported into the per-cell electrophysiological feature vector (`ep_row.csv`).

In the *PatchClamp* branch, the reference first spike was selected and exported through dataset-specific scripts called by the main analysis pipeline which generated stimulus-specific electrophysiological summaries for each neuron and supported both **long-square** and **ramp** stimulation protocols. The output is written in the same standardized per-cell feature tables used in the downstream analyses.

The final cell-level file `ep_row.csv` contains three main groups of features:

- **metadata and sweep-level summary features**: including the reference sweep identifier, the latency of the first spike, and the injected current at spike initiation;
- **raw first-spike markers** such as threshold, peak, fast-trough and slow-trough landmarks, the upstroke/downstroke ratio, in-stimulus resting potential, AP area, AHP area, threshold-to-peak amplitude, and AHP slope;
- **NSS-derived waveform features**.

Two complementary families of first-spike descriptors are considered:

- `firstspk_*`: features describing the first spike in the selected reference sweep;
- `firstspkmean_*`: features obtained by averaging the first spike across first spikes taken from all sweeps of the same neuron.

A third representation, denoted `combo`, combined both feature families into a unified per-cell vector. This distinction was preserved in downstream analyses, where `firstspk` emphasized a single first spike excitability snapshot, `firstspkmean` emphasized a more stable average first-spike tendency across sweeps, and `combo` integrated both perspectives.

The output of this stage was written at two levels. At the specimen level, each neuron had a dedicated feature directory containing:

- `sweep_features.csv`,
- `cell_features.csv`,

- `spike_table.csv`,
- `nss_spike_features.csv`,
- `ep_row.csv`.

At the dataset level, the per-cell rows were concatenated into global electrophysiological feature matrices. For *PatchClamp*, these included `PatchClamp_EP_features_long_square.csv`, `PatchClamp_EP_features_ramp.csv`, and optionally a merged `PatchClamp_EP_features.csv`. For *PatchSeq*, species-specific output was generated, `PatchSeq_EP_features_mouse.csv`.

These global electrophysiological matrices formed the input to the downstream clustering analysis.

3.5.1 Clustering analysis of first-spike electrophysiological features

Feature subsets were selected according to the analysis mode (`firstspk`, `firstspk mean`, or `combo`) and stimulation protocol. Before clustering, selected variables were converted to numeric form, features containing only missing values were discarded, low-variance features were removed, missing values were imputed using the median, and the feature matrix was standardized. Redundant variables were then pruned using a correlation threshold. The resulting feature space was analyzed either directly or after dimensionality reduction by principal component analysis (PCA), with the number of retained principal components determined from cumulative explained variance. Clustering was subsequently performed either in the original feature space or in PCA space, and the resulting neuronal states were visualized in PCA coordinates and related to subclass annotations derived from *cre*-type labels in PatchClamp and transcriptomic subclass labels in PatchSeq.

The script was applied independently to *PatchClamp* and *PatchSeq* electrophysiological feature tables, maintaining the distinction between **ramp** and **long-square** stimulation.

Three different feature representations were tested for each input matrix. In the `rheo_first` configuration the features describe the first spike measured in the selected reference sweep along with spike latency, the in-stimulus resting potential and the current level at spike initiation. The `firstspkmean` configuration, instead, uses the descriptors obtained by averaging the first spike across all available

sweeps of the same neuron. Ans lastly the `combo` configuration combines both representations into a single feature vector.

The selected variables were then converted to numeric format and filtered to remove unusable or redundant features. Columns containing only missing values were discarded and features with very small variance were removed. Any remaining missing entries were replaced using the median value of the corresponding feature. The resulting matrix was then standardized to zero mean and unit variance and to limit redundancy between variables, highly correlated features were also removed using an absolute correlation threshold of 0.85 (in the present implementation).

Clustering was at this point done either directly on the standardized feature matrix or on the lower-dimensional representation obtained through principal component analysis (PCA), which was fitted on the standardized data and which retained the minimum number of components required to explain 85% of the cumulative variance [35]. In this way clustering was examined both in the original standardized feature space and in the PCA space defined by the retained components.

The script supports agglomerative clustering, k-means, and Gaussian mixture models. For each dataset and feature representation, clustering solutions were evaluated across a range of candidate cluster numbers k. To assess clustering quality, three complementary internal validation metrics were computed for each value of k:

- **Silhouette score** [36]. For each observation, the silhouette coefficient compares the mean distance to points within the same cluster with the mean distance to points in the nearest neighboring cluster. The silhouette coefficient for observation i is defined as

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}},$$

  where $a(i)$ is the mean intra-cluster distance and $b(i)$ is the mean distance to the nearest neighboring cluster. The score ranges from -1 to 1, where higher values indicate better separation between clusters relative to their internal dispersion.

- **Calinski-Harabasz index** [37]. This index evaluates the ratio between between-cluster dispersion and within-cluster dispersion. It is defined as

$$CH = \frac{\mathrm{Tr}(B_K)}{\mathrm{Tr}(W_K)} \cdot \frac{N - K}{K - 1},$$

where B_K and W_K denote the between-cluster and within-cluster dispersion matrices, N is the number of observations, and K the number of clusters. Larger values indicate solutions in which clusters are compact and well separated in the feature space.

- **Davies-Bouldin index** [38]. The Davies-Bouldin index evaluates the similarity between clusters by comparing their internal scatter with the separation between cluster centroids. It is defined as

$$DB = \frac{1}{K} \sum_{i=1}^{K} \max_{j \neq i} \frac{S_i + S_j}{M_{ij}},$$

where S_i denotes the within-cluster scatter of cluster i, and M_{ij} is the distance between the centroids of clusters i and j. Lower values correspond to clusterings with better separation and smaller internal dispersion.

These three metrics were used to evaluate clustering quality, since each one captures a slightly different aspect of the cluster structure: the silhouette score mainly reflects how well clusters are separated, but in some datasets it can favor very small numbers of clusters; the Calinski-Harabasz and Davies-Bouldin indices provide complementary information by measuring cluster compactness and the relative dispersion between clusters.

When automatic cluster selection was enabled, the number of clusters was determined from the silhouette profile using a simple rule stating that among the tested values of k, the selected solution corresponded to the smallest k whose silhouette score reached at least a fixed fraction of the maximum silhouette value observed across the range. This favors compact and well-separated clusters while avoiding unnecessarily large values of k.

The script produced for each obtained clustering solution a table assigning every neuron to a cluster label together with several diagnostic outputs, including PCA score matrices, PCA loadings, scree plots, cluster-metric curves across k, silhouette plots, and two- or three-dimensional PCA visualizations colored by cluster identity. Also mutual-information rankings of features with respect to cluster labels and violin plots showing the distribution of individual features across clusters were generated.

Although the clustering workflow itself was shared across datasets, subclass annotation was handled differently for *PatchClamp* and *PatchSeq*. In *PatchClamp*, neurons were annotated using the `cre_type` field contained in the electrophysiological table, reduced to its first token. In *PatchSeq*, electrophysiological rows

were joined to external metadata by matching `specimen_id` in the EP table to `ephys_session_id` in the metadata table; subclass labels were then derived as the first token of `corresponding_AIT2.3.1_alias`. These subclass annotations were not used to define the clusters, but only to visualize cluster composition and aid biological interpretation.

3.6 Whole-Signal Dynamical Analysis

While the first-spike analysis focuses on the waveform geometry of a single action potential, intrinsic neuronal excitability is fundamentally a dynamical process unfolding over time during stimulation. To capture this aspect of neuronal behavior, the membrane-potential response to ramp stimulation was analyzed as a continuous time series and represented in reconstructed dynamical state spaces.

This analysis aims to characterize the geometry and temporal evolution of the membrane-potential trajectory during stimulation. Rather than reducing the electrophysiological response to a small set of spike-level descriptors, the voltage trace is analyzed as a trajectory generated by an underlying dynamical system, from which quantitative descriptors of its structure are extracted [39].

In the present work, whole-signal dynamical analysis was applied to ramp stimulation sweeps, which provide a continuously varying depolarizing drive and therefore reveal gradual transitions between subthreshold and spiking regimes.

3.6.1 Signal normalization

Dynamic feature extraction operated on the baseline-corrected and onset-aligned traces produced during preprocessing. Because geometric descriptors of the reconstructed state space depend on relative trajectory shape rather than absolute voltage scale, voltage traces were standardized before embedding.

The normalized signal was defined as

$$v_z(t) = \frac{v_{\text{aligned}}(t) - \mu}{\sigma},$$

where μ and σ denote the mean and standard deviation of the aligned voltage trace. This normalization keeps waveform geometry intact and removes absolute voltage offsets across neurons.

3.6.2 State-space embeddings

To analyze the voltage trace as a dynamical trajectory, the time series was embedded into low-dimensional state spaces and two complementary embeddings were used.

Derivative embedding Each time point was represented by the vector

$$(v(t), \dot{v}(t), \ddot{v}(t)),$$

where $v(t)$ denotes the membrane potential, $\dot{v}(t)$ its first time derivative, and $\ddot{v}(t)$ its second derivative.

The derivatives were obtained using Savitzky-Golay filtering, which provides smooth derivative estimates and limits the amplification of high-frequency noise [40]. Thanks to this coordinate system rapid voltage transitions, spike initiation, and local curvature of the trajectory become more easily visible.

Conceptually, derivative embeddings approximate the local phase space of the underlying dynamical system governing membrane potential dynamics [7].

Delay-coordinate embedding a second representation was obtained using a delay-coordinate embedding:

$$(v(t), v(t+\tau), v(t+2\tau)).$$

The delay τ was defined in milliseconds and converted to samples using the sampling frequency of the recording. In the present implementation a default value of $\tau = 5\,\text{ms}$ was used.

Delay embeddings are motivated by Takens' theorem, which states that the dynamics of an unknown deterministic system can be reconstructed from delayed observations of a single scalar variable [5]. In this context, the delay embedding captures the geometry of voltage trajectories without requiring explicit derivative estimation [6].

The derivative and delay embeddings together give two different views of the same voltage dynamics since the derivative representation focuses on local differential structure, highlighting rapid voltage changes and curvature of the trajectory, while

the delay embedding captures temporal recurrence and the global geometry of the trajectory in state space.

3.6.3 Dynamical feature families

Starting from the baseline-corrected voltage trace and the corresponding embedded trajectories, several groups of dynamical features were extracted. These quantities describe different aspects of the reconstructed signal, including its statistical variability, the geometry of the trajectory, and the way the dynamics evolve over time.

Poincaré descriptors A first set of features was derived from the Poincaré representation of the voltage signal. For a discrete time series x_n, the Poincaré map is obtained by considering pairs of consecutive samples (x_n, x_{n+1}).

From this representation the following quantities were computed:

- **SD1**: the standard deviation of successive differences, capturing short timescale variability;
- **SD2**: the standard deviation along the identity line, reflecting slower variations of the signal;
- the **correlation coefficient** r between consecutive samples;
- the **area of the Poincaré ellipse**, defined as

$$A = \pi\, SD1 \cdot SD2;$$

- the mean and standard deviation of the signal.

These descriptors provide a compact summary of the local temporal variability of the voltage trajectory.

Increment Poincaré descriptors: the same Poincaré statistics were also computed on the first difference of the voltage signal,

$$\Delta v_n = v_{n+1} - v_n.$$

Rapid changes in the signal are highlighted when working with voltage increments, rather than with its absolute level. This representation tends to emphasize spike-related transitions and other transient events in the membrane potential.

Embedding geometry features: the geometry of the reconstructed trajectories was characterized by several descriptors computed in both derivative and delay embedding spaces.

- **Covariance eigenvalues:** the eigenvalues of the covariance matrix of the embedded point cloud describe how the trajectory spreads along its principal directions.
- **Anisotropy:** anisotropy was defined as the ratio between the largest and smallest covariance eigenvalues

$$\text{anisotropy} = \frac{\lambda_1}{\lambda_3}.$$

  where large values indicate elongated trajectories, which typically arise when the dynamics evolve predominantly along a preferred direction.
- **Bounding-box dimensions:** axis-aligned bounding boxes were used to estimate the spatial extent of the trajectory along each embedding axis.
- **Bounding-box aspect ratios:** ratios between these bounding-box dimensions provide a measure of trajectory elongation and directional structure.
- **Convex hull geometry:** when the embedded trajectory contained a sufficient number of points, its convex hull was computed and convex hull volume and convex hull surface area were extracted, to capture the overall spatial spread of the trajectory within the reconstructed state space.

Alpha-shape mesh descriptors (optional): when optional dependencies were available, a more detailed geometric reconstruction was performed using alpha shapes. The alpha shape generalizes the convex hull by allowing concave boundaries controlled by a scale parameter α [41].

From the reconstructed mesh these geometric descriptors were calculated to give a more detailed description of the trajectory:

- mesh volume

- mesh surface area
- watertightness indicator
- Euler characteristic
- principal inertia components

They were treated as optional features in the analysis pipeline because their computation is more expensive.

3.6.4 Temporal windowing and adaptation-sensitive descriptors

Neuronal responses during ramp stimulation often evolve over time as the injected current increases. To capture this temporal evolution, each post-onset voltage trace was divided into a fixed number of temporal windows (four windows by default).

The complete set of dynamical features described above was recomputed separately for each window, thus producing window-specific descriptors that capture the early and late portions of the neuronal response.

The pipeline also computed difference features comparing the last and first windows in order to examine how the dynamics changed during stimulation:

$$\Delta f = f_{\text{last window}} - f_{\text{first window}}.$$

These `delta_last_first` features track how the geometry of the trajectory evolves over time and are useful for identifying effects such as adaptation, gradual depolarization, or transitions between different dynamical regimes.

3.6.5 Representative ramp sweep selection

Dynamic features were initially extracted at the level of individual sweeps, producing a sweep-level table (`features.csv`). Because ramp responses from the same neuron may vary substantially across repetitions, a representative sweep was selected for each neuron, selected among those that passed quality control.

The ramp chosen among these was the one with the lowest current level at the first detected spike, corresponding to the earliest ramp that successfully triggered spiking.

The identity of the selected sweep was stored in `best_ramp_sweeps_per_neuron.csv`, and its corresponding dynamical descriptors were exported to `neurons_features_best_ramp.csv`. These neuron-level feature vectors formed the input to the clustering analysis described below.

3.6.6 Conceptual role of the whole-signal analysis

Whereas the first-spike analysis characterizes neuronal excitability through waveform descriptors extracted at the onset of the first action potential, the whole-signal analysis treats excitability as a dynamical process unfolding over the entire duration of stimulation.

During ramp current stimulation, the injected current increases gradually, and the membrane potential evolves continuously in response. This protocol makes it possible to observe dynamical aspects of neuronal behavior that are not visible from a single spike alone since the stimulus here unfolds over time.

One common example is *spike-frequency adaptation*, where the firing rate progressively changes during sustained depolarization. This effect is typically linked to mechanisms such as slow potassium currents or sodium channel inactivation [34]. Ramp stimulation can also show slower shifts in the membrane potential baseline or gradual changes in spike-threshold dynamics during depolarization [42].

In this context, the voltage trace can be interpreted as a trajectory generated by the underlying dynamical system governing membrane potential dynamics. By reconstructing the trajectory of the signal in low-dimensional state spaces, it becomes possible to characterize the geometry and temporal organization of the neuronal response. State-space reconstruction techniques, such as delay-coordinate embeddings, provide a way to approximate the state-space structure of an unknown dynamical system using observations of a single variable. These representations have been widely used in nonlinear time-series analysis to study the geometry of physiological signals [5, 6, 7] and are shown in Figure 3.6.

The resulting dynamical features therefore capture multiple aspects of neuronal activity, including spike morphology, temporal variability, trajectory dispersion, and changes in dynamical structure over the course of stimulation. Rather than describing only the initiation of spiking, this approach provides a representation of

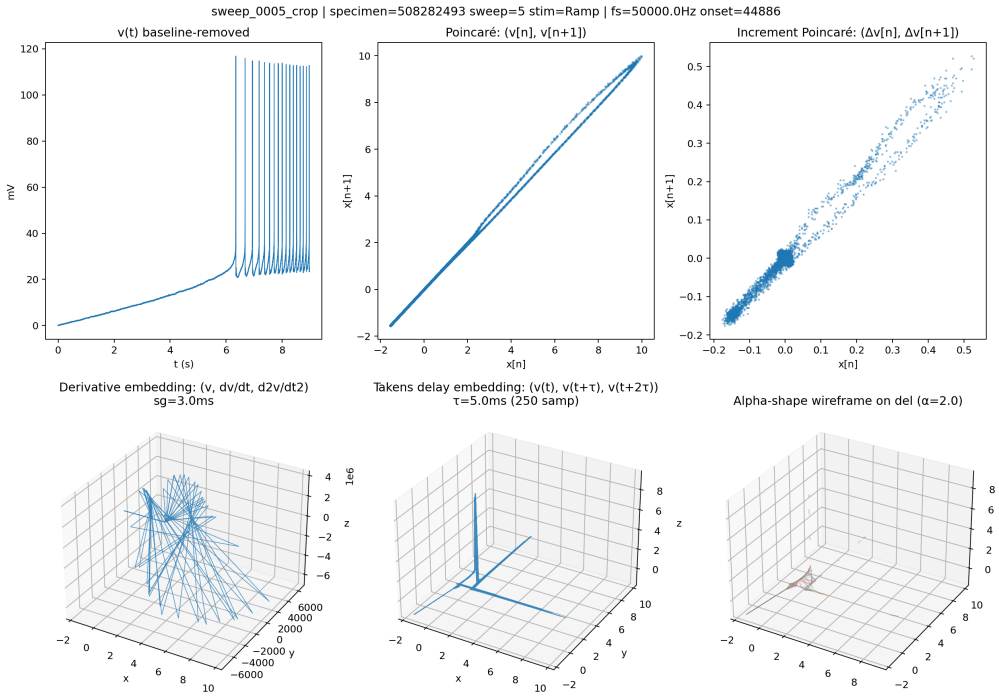


Figure 3.6: Exemplar embeddings of voltage trace.

how neuronal responses evolve as the system moves through different dynamical regimes during increasing depolarizing drive.

In the present work, these whole-signal descriptors complement the first-spike feature set by providing a richer characterization of neuronal excitability. The combined feature space formed the basis for the identification of distinct dynamical excitability regimes described in Chapter 4.

3.6.7 Rationale for the choice of whole-dynamics descriptors

In contrast to approaches that summarize neuronal excitability through isolated spike waveform measurements, the whole-signal analysis treats the membrane potential trace as the observable output of an underlying dynamical system. During current stimulation, the membrane voltage evolves according to the interaction of multiple hidden variables, including voltage-gated conductances, adaptation currents, and synaptic inputs [43]. Because intracellular recordings measure only the membrane potential, the observed signal represents a one-dimensional projection of a higher-dimensional dynamical process. Methods from nonlinear time-series analysis therefore provide a framework to characterize the geometry and variability of this trajectory [6].

A theoretical motivation for this approach comes from delay-coordinate reconstruction. Takens' embedding theorem shows that the state space of a deterministic dynamical system can, under suitable conditions, be reconstructed from time-delayed observations of a single scalar signal [5].

The voltage trace can thus be represented using delay coordinates such as $(v(t), v(t+\tau), v(t+2\tau))$ that provide an approximation of the latent dynamical manifold governing the membrane-potential response.

Within this reconstructed representation, several complementary families of descriptors were extracted in order to capture different aspects of neuronal dynamics. First, Poincaré descriptors quantify short-lag dependence between successive samples of a physiological time series. Originally introduced in heart-rate variability analysis, these descriptors measure local variability and serial correlation through the geometry of the (x_n, x_{n+1}) plot [44]. Their increment-based variants further emphasize rapid fluctuations by applying the same analysis to successive signal differences. Although developed in cardiac physiology, these descriptors are general measures of lag-dependent variability and are therefore applicable to membrane potential trajectories recorded during stimulation.

Second, the voltage trajectory was embedded in derivative coordinates

$$\left(v,\ \frac{dv}{dt},\ \frac{d^2v}{dt^2}\right)$$

This representation directly captures local slope and curvature of the signal and has been widely used in nonlinear analyses of physiological signals, such as voice dynamics, to reveal irregular or chaotic attractor structures [45]. In the context of neuronal recordings, these variables correspond to biologically meaningful quantities related to the approach to threshold, spike initiation dynamics, and repolarization kinetics.

Geometric descriptors of the reconstructed trajectory were also computed to summarize its global organization in state space. Among these, the eigenvalues of the covariance matrix and the resulting anisotropy describe the main directions of variability within the point cloud, while other measures based on bounding boxes and convex hulls estimate the spatial extent occupied by the trajectory.

Descriptors of this type are commonly used in attractor analysis to describe the morphology and dispersion of reconstructed dynamical systems [46]. To capture more complex structures, alpha-shape mesh descriptors were also included, allowing non-convex geometries of the reconstructed attractor to be represented beyond simple convex-hull approximations.

Finally, the temporal evolution of these descriptors was assessed by computing features within successive temporal windows and comparing early and late segments of the stimulation response. This strategy allows detection of gradual changes in excitability during stimulation, which are expected to arise from physiological mechanisms such as spike-frequency adaptation and dynamic spike-threshold modulation [34, 42].

These descriptor families form a compact representation of neuronal response dynamics that goes beyond information contained in single spike waveforms, capturing several complementary properties of the voltage trajectory including short-timescale variability, reconstructed state-space structure, the spatial extent of the dynamical attractor, and changes in excitability during stimulation.

An additional motivation for the whole-signal representation arises from the nature of ramp stimulation protocols. Classical spike-train descriptors such as firing rate, inter-spike intervals, or spike-count statistics are primarily defined for responses that contain sustained sequences of spikes. While these metrics are highly informative for long square-current stimuli that elicit stable firing regimes, they are less suitable for ramp protocols in which the neuron may produce only a few spikes or transition

gradually from subthreshold to suprathreshold dynamics. In such conditions, spike-train statistics provide only sparse information and may fail to capture the structure of the underlying membrane response.

Whole-signal descriptors instead operate directly on the continuous membrane-potential trajectory and are informative even when spiking activity is sparse or absent. They capture features of the response that are not accessible from individual spikes alone, such as subthreshold fluctuations, the gradual approach to spike threshold, spike initiation dynamics and post-spike recovery, because they analyze the entire voltage trace during stimulation, not just individual spikes.

This is particularly relevant for ramp stimuli, where neuronal excitability develops progressively as the injected current increases. Analyzing the full voltage trajectory therefore provides a broader description of neuronal dynamics across both subthreshold and suprathreshold regimes.

Descriptor family	**Typical literature application**	**Property measured**	**Rationale**
Poincaré descriptors (SD_1, SD_2, correlation, ellipse area)	Heart-rate variability analysis	Short-lag variability and serial dependence between consecutive observations	Captures local variability and short-timescale dependence in voltage trajectories during stimulation, reflecting rapid fluctuations and recovery dynamics [44].
Increment Poincaré descriptors	HRV variability analysis	Variability of successive increments rather than absolute values	Focuses on rapid changes in membrane potential and slope-to-slope variability with less sensitivity to absolute voltage offsets.
Delay-coordinate embedding $(v(t), v(t+\tau), v(t+2\tau))$	Nonlinear dynamical systems and attractor reconstruction	Reconstruction of latent state-space dynamics from a single observable	Provides a principled reconstruction of neuronal dynamics from the observed membrane potential signal according to Takens' theorem [5].

Descriptor family	**Typical literature application**	**Property measured**	**Rationale for neuronal voltage traces**
Derivative embedding (v, dv/dt, d^2v/dt^2)	Nonlinear physiological signal analysis and voice dynamics	Local slope and curvature of signal evolution	Directly captures biologically meaningful quantities associated with spike approach, threshold crossing, and repolarization dynamics [45].
Covariance eigenvalues and anisotropy	Attractor geometry analysis	Principal axes of variation and shape anisotropy of the point cloud	Distinguishes trajectories confined to narrow manifolds from more dispersed dynamical responses in neuronal state space [46].
Bounding-box descriptors	Geometric attractor characterization	Global extent of the trajectory along each dimension	Provides robust measures of the amplitude and dispersion of neuronal trajectories within reconstructed state space.
Convex-hull descriptors	Point-cloud geometry and attractor morphology	Volume and surface area of trajectory envelope	Measures the spatial region explored by the voltage dynamics during stimulation.
Alpha-shape mesh descriptors	Nonlinear attractor analysis in voice and physiological signals	Non-convex shape and structural complexity of attractors	Captures complex dynamical structures that cannot be represented by convex hulls alone.
Windowed features and early-late deltas	Temporal attractor evolution in physiological signals	Changes in dynamical properties across time	Allows detection of stimulus-dependent evolution of neuronal excitability associated with adaptation and threshold dynamics [34, 42].

Table 3.1: Summary of descriptor families used in the whole-dynamics analysis.

3.6.8 Dimensionality reduction, clustering, and report generation

The final stage of the whole-signal pipeline operated on the neuron-level table obtained after representative-ramp selection, with one row per neuron. This table contained the dynamic descriptors extracted from the selected ramp together with quality-control metadata and, when available, cell-type annotations. A dedicated analysis script was then used to perform feature selection, dimensionality reduction, clustering, and report generation in a fully reproducible manner.

The analysis proceeded as follows:

- **Feature selection:** several feature-selection modes were evaluated in order to assess how different families of dynamical descriptors organize the excitability landscape. A compact reference representation (`minimal_ramp`) retained a small set of interpretable descriptors summarizing key aspects of the voltage dynamics, including Poincaré and increment Poincaré statistics, anisotropy measures derived from derivative and delay embeddings, and selected early-late difference features.

  Additional modalities were constructed by selectively retaining specific descriptor families, such as Poincaré statistics, increment Poincaré statistics, or geometric descriptors derived from derivative embeddings. These alternative representations allowed us to test whether particular aspects of the reconstructed dynamical trajectory (e.g., local variability, geometric structure, or temporal evolution) were sufficient to recover meaningful excitability regimes.

- **Preprocessing of the feature matrix:** features containing only missing values were removed, and near-zero variance features were discarded. Remaining missing values were imputed by feature-wise median. The feature matrix was then standardized to zero mean and unit variance, and highly correlated features were pruned using a user-defined absolute-correlation threshold in order to reduce redundancy. We evaluated three correlation thresholds: 0.70, 0.80, and 0.90.

- **Principal component analysis:** principal component analysis (PCA) was performed on the standardized feature matrix. The maximum number of components was limited by the number of neurons and retained variables, while the subset of components used in downstream analyses was chosen using a cumulative explained-variance criterion. To document the dimensionality-reduction step, the pipeline generated scree plots, cumulative explained-variance curves, PCA loadings, neuron scores, and a reconstruction-error curve in scaled feature space.

- **Clustering:** clustering could be carried out either in the original preprocessed feature space, in PCA space, or in both. The implemented method was agglomerative clustering with a user-selectable linkage criterion; in practice Ward linkage was typically used, as it merges clusters while minimizing the increase in within-cluster variance. The number of clusters could be specified *a priori* or selected automatically over a user-defined range of k values.

- **Internal validation and model selection:** for each tested value of k, the script computed several internal validation metrics, including the silhouette score, the Calinski-Harabasz index, and the Davies-Bouldin index. In automatic mode, the chosen number of clusters corresponded to the smallest k whose silhouette score reached at least a fixed fraction of the maximum value observed across the tested range, thereby favoring parsimonious solutions. Additional k-means inertia elbow plots were also produced for both feature space and PCA space. When clustering was performed in both representations, agreement between the two labelings was quantified using the adjusted Rand index.

- **Visualization and biological annotation:** the final cluster labels were exported together with neuron identifiers and available metadata. PCA coordinates were used to generate two- and three-dimensional visualizations in which color indicated cluster membership and marker shape indicated neuronal subclass. External metadata could be joined using the electrophysiology session identifier in the PatchSeq dataset using a simplified subtype annotation corresponding to the first token of the transcriptomic alias.

- **Cluster interpretation:** the script produced cluster-size summaries, stacked bar plots of subclass composition, cluster centroids in standardized feature space and centroid-distance heatmaps. There is also an optional medoid module that identifies for each feature-space cluster the neuron closest to the cluster centroid and generates appendix plots showing its representative ramp trace, derivative embedding, delay embedding, Poincaré map, increment Poincaré map, and the optional alpha-shape visualization.

- **Reproducibility outputs:** all analysis settings including preprocessing choices, dimensionality, reduction parameters, clustering configuration, and output paths were stored in a machine-readable `report.json` file that was produced together with tabular summaries and figures generated at each stage of the pipeline.

3.6.9 Whole-signal feature configurations

In the whole-signal dynamics analysis, features were engineered to model the underlying dynamical structure of membrane-potential traces. These descriptors aim to capture properties of neuronal voltage dynamics such as short-timescale variability, spike recruitment dynamics, trajectory geometry in reconstructed state space, temporal evolution of firing responses during ramp stimulation, and changes in membrane-potential fluctuations associated with transitions between excitability regimes.

Each feature modality therefore corresponds to a specific subset of dynamical descriptors derived from the reconstructed voltage trajectories. These modalities were designed to isolate complementary aspects of neuronal dynamics, including signal variability, trajectory geometry in reconstructed state spaces, and the temporal evolution of voltage statistics along the ramp stimulus, and are summarized in Table3.2.

The different modalities isolate complementary aspects of the reconstructed voltage dynamics:

- **`minimal_ramp`:** a compact representation designed to summarize ramp responses using a small set of interpretable descriptors combining variability, trajectory geometry, and temporal evolution (`p_sd1`, `p_sd2`, `p_r`, `p_area`, `p_mean`, `p_std`, `ip_sd1`, `ip_sd2`, `ip_r`, `ip_area`, `ip_mean`, `ip_std`, `full_der_cov_anisotropy`, `full_del_cov_anisotropy`, `delta_last_first__p_mean`, `delta_last_first__p_std`, `delta_last_first__ip_mean`, `delta_last_first__ip_std`).

- **`poincare_only`:** descriptors derived from the Poincaré map (v_n, v_{n+1}), which quantify the short-timescale variability of successive voltage samples (`p_sd1`, `p_sd2`, `p_r`, `p_area`, `p_mean`, `p_std`).

- **`increment_poincare_only`:** Poincaré statistics computed on voltage increments $(\Delta v_n, \Delta v_{n+1})$, highlighting rapid voltage changes and spike-recruitment dynamics (`ip_sd1`, `ip_sd2`, `ip_r`, `ip_area`, `ip_mean`, `ip_std`).

- **`derivative_geometry_only`:** geometric descriptors of the trajectory in derivative state space $(v, dv/dt, d^2v/dt^2)$, describing curvature, anisotropy, and the spatial extent of the trajectory (e.g. `full_der_cov_anisotropy`, `full_der_cov_e1`, `full_der_cov_e2`, `full_der_bbox_dx`, `full_der_bbox_dy`, `full_der_bbox_dz`).

- **`delay_geometry_only`:** descriptors of the reconstructed attractor in delay coordinates $(v(t), v(t+\tau), v(t+2\tau))$, reflecting recurrence structure and global trajectory shape (e.g. `full_del_cov_anisotropy`, `full_del_cov_e1`, `full_del_bbox_dx`, `full_del_bbox_dy`, `full_del_bbox_dz`).

- **`dynamics_only`:** combination of Poincaré and increment-Poincaré descriptors and early-late temporal differences, describing both variability and its evolution along the ramp (`p_*`, `ip_*`, `delta_last_first__*`).

- **`geometry_plus_delta`:** geometric features of reconstructed trajectories combined with early–late differences, enabling the analysis to track how the trajectory structure changes during the ramp (`full_der_*`, `full_del_*`, `delta_last_first__*`).

- **`increment_plus_derivative`:** (`ip_*`, `full_der_*`).

- **`poincare_plus_derivative`:** (`p_*`, `full_der_*`).

- **`increment_derivative_delta`:** extension of the increment–derivative representation that also includes early-late temporal differences across the ramp (`ip_*`, `full_der_*`, `delta_last_first__*`).

- **`poincare_increment_derivative`:** (`p_*`, `ip_*`, `full_der_*`).

- **`global_full_only`:** global descriptors computed over the entire ramp response, treating the signal as a single dynamical trajectory (`full_*`, `p_*`, `ip_*`).

- **`mesh_hull_only`:** envelope-based geometric descriptors of the reconstructed trajectory quantifying the occupied state-space volume and surface (e.g. `mesh_*`, `hull_*`).

- **`extended_ramp`:** an expanded version of the compact ramp representation that combines global variability measures, trajectory geometry, early-late temporal differences, and selected late-window geometric descriptors, including mesh and convex-hull measures from the later phases of the ramp, capturing in this way both the overall dynamical organization of the response and the emergence of excitability-state differences in the most informative late portions of the stimulus (`p_*`, `ip_*`, `full_der_*`, `full_del_*`, `delta_last_first__*`, `w2__*`, `w3__*`, late-window `mesh_*` and `hull_*` descriptors).

Feature modality	Feature families included
`minimal_ramp`	Poincaré descriptors, increment Poincaré descriptors, embedding anisotropy, and temporal-difference features.
`poincare_only`	Poincaré descriptors (`p_*`, windowed `__p_*`).
`increment_poincare_only`	Increment-Poincaré descriptors (`ip_*`, windowed `__ip_*`).
`derivative_geometry_only`	Derivative embedding geometry (`full_der_*`, `__der_*`).
`delay_geometry_only`	Delay embedding geometry (`full_del_*`, `__del_*`).
`dynamics_only`	Poincaré + increment Poincaré descriptors with temporal deltas.
`geometry_plus_delta`	Derivative and delay embedding geometry with temporal deltas.
`increment_plus_derivative`	Increment-Poincaré descriptors combined with derivative embedding geometry.
`poincare_plus_derivative`	Poincaré descriptors combined with derivative embedding geometry.
`increment_derivative_delta`	Increment-Poincaré + derivative embedding geometry with temporal deltas.
`poincare_incr_derivative`	Poincaré + increment-Poincaré descriptors with derivative embedding geometry.
`global_full_only`	Global signal descriptors (`full_*`, `p_*`, `ip_*`).
`mesh_hull_only`	Mesh and convex-hull geometric descriptors of the reconstructed trajectory.
`extended_ramp`	Poincaré and increment-Poincaré descriptors, derivative and delay geometry, temporal deltas, late-window geometry, and mesh/hull descriptors.

Table 3.2: Feature-selection modalities for the whole-signal clustering analysis. Each modality corresponds to a different combination of descriptor families derived from reconstructed voltage trajectories.

3.6.10 Transcriptomic Analysis of Plasticity-Related Gene Expression

Conceptual framework: neuronal activity versus plasticity mechanisms

The electrophysiological and transcriptomic components of this study probe two complementary but distinct aspects of neuronal function. Electrophysiological recordings capture the instantaneous dynamical behavior of neurons, reflecting their current excitability regime and response properties at the time of measurement. Features extracted from membrane potential traces therefore describe the functional activity state of the neuron.

In contrast, transcriptomic measurements provide a snapshot of the molecular machinery expressed by the cell. In particular, genes associated with neuronal plasticity reflect regulatory and structural pathways through which neurons can modify their physiological properties over longer time scales. These include transcription factors, signaling molecules, ion channels, and extracellular matrix components involved in activity-dependent remodeling of neuronal circuits.

The presence of plasticity-related transcripts does not necessarily mean that a neuron is currently undergoing plastic change. Gene expression profiles instead reflect the molecular machinery that can support plastic modifications. For this reason electrophysiological activity states and transcriptomic plasticity markers are better interpreted as complementary sources of information rather than as a single unified neuronal state.

In this context the electrophysiological analysis is used to identify different dynamical regimes of neuronal excitability, while the transcriptomic analysis examines whether neurons belonging to these regimes differ in the expression of genes associated with plasticity-related mechanisms.

Patch-seq transcriptomic dataset

Transcriptomic measurements were obtained from the Patch-seq dataset of mouse visual cortical neurons reported by [22]. Patch-seq combines whole-cell patch-clamp electrophysiology with single-cell RNA sequencing (scRNA-seq), allowing electrophysiological properties and gene expression profiles to be measured from the same neuron.

The dataset provides raw gene expression count matrices together with metadata

linking transcriptomic profiles to electrophysiological recordings. In the count matrix, rows correspond to genes and columns correspond to individual cells. Metadata tables include specimen identifiers and transcriptomic cell type annotations.

Cells were matched between electrophysiological and transcriptomic modalities using the specimen identifiers provided in the dataset metadata as this allowed transcriptomic measurements to be matched with the electrophysiological excitability descriptors computed in the earlier stages of the analysis.

Construction of the transcriptomic analysis object

To facilitate downstream analysis, the transcriptomic dataset was converted into the `AnnData` data structure used by the Scanpy framework wolf2018scanpy. In this representation, cells correspond to observations and genes correspond to variables.

The expression matrix was stored in compressed sparse row format in order to reduce memory usage. Cell-level metadata were recorded in the observation table of the AnnData object. Simplified transcriptomic subclass labels derived from the original cell-type annotations were included and later used for visualization and interpretation.

RNA preprocessing

Transcriptomic preprocessing was performed using the Scanpy framework. Library size normalization was first applied to compensate for differences in sequencing depth across cells. For each cell, the total counts were rescaled to a target value of 10^4 using the `normalize_total` function.

The gene expression values were transformed using the natural logarithm of one plus the normalized counts (`log1p`) as this transformation helps stabilize variance and reduces the dominance of highly expressed genes, making the data more suitable for downstream dimensionality reduction and clustering.

So the expression matrix consisted of log-transformed normalized gene-expression values for each neuron.

Selection of plasticity-related genes

Rather than performing a genome-wide transcriptomic analysis, this study focused on a curated subset of genes associated with neuronal plasticity. Candidate genes were initially identified through Gene Ontology queries using the AmiGO interface, specifically targeting the term *regulation of synaptic plasticity* for the organism *Mus musculus*. Only annotations curated within the *go_central* ontology were retained.

After removing duplicate entries and supplementing the list with additional plasticity‡ related genes identified in the literature, a set of 135 candidate genes associated with plasticity-related processes was obtained. These genes encompass multiple biological mechanisms underlying neuronal plasticity, including transcriptional regulation, intracellular signaling pathways, ion channel modulation, and extracellular matrix remodeling.

For visualization and interpretation, a representative subset of sixteen genes frequently discussed in the literature on interneuron plasticity was selected. These include extracellular matrix components (*Acan*, *Bcan*, *Hapln1*, *Tnr*), signaling molecules (*Nrg1*, *Erbb4*), ion channel genes (*Kcnc1*, *Scn1a*, *Hcn1*, *Kcnq2*, *Cacna1c*), and neurotrophic signaling genes (*Bdnf*, *Ntrk2*, *Mtor*) [47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59]. Only genes present in the transcriptomic dataset were retained for downstream analysis.

Dimensionality reduction in plasticity gene expression space

Principal component analysis (PCA) was performed on the expression matrix restricted to the selected gene set in order to characterize the structure of plasticity-related gene expression across neurons. PCA identifies those orthogonal directions that capture maximal variance in the data and enables representation of transcriptomic variability in a lower-dimensional space.

Making use of the cumulative explained variance, the number of principal components used for downstream analysis was determined either manually or automatically based on it. The automatic case selected the smallest number of components required to explain a predefined fraction of the total variance (typically 80%) as this preserved the majority of transcriptional variability while avoiding unnecessary model complexity. The resulting PC coordinates define a low-dimensional embedding that summarizes variation in plasticity-related gene expression across neurons.

Graph construction and clustering

A k-nearest neighbor graph was constructed using the PCA representation of the plasticity gene expression matrix. In this graph, each neuron is connected to its nearest neighbors in transcriptomic space based on Euclidean distance in the PCA coordinates.

Community detection was then performed using the Leiden algorithm [60], which identifies densely connected groups of nodes in the graph and produces well-connected clusters. In environments where the Leiden implementation is unavailable, the Louvain algorithm can be used as a fallback [61].

The resulting cluster assignments represent groups of neurons that exhibit similar patterns of plasticity-related gene expression and are called *plasticity clusters*.

Visualization of transcriptomic structure

To visualize the organization of neurons in plasticity gene expression space, Uniform Manifold Approximation and Projection (UMAP) [62] was applied to the neighbor graph constructed from the PCA representation.

UMAP was used to obtain a two-dimensional embedding of the transcriptomic data that preserves local neighborhood relationships and where cells that have similar gene-expression profiles tend to appear close to one another, making it easier to visualize transcriptional gradients and cluster boundaries related to plasticity-associated genes.

Integration with electrophysiological excitability space

The transcriptomic analysis was integrated with the electrophysiological clustering results obtained in previous stages of the analysis pipeline. Each neuron therefore possesses both electrophysiological descriptors derived from intracellular recordings and transcriptomic measurements derived from single-cell RNA sequencing.

To investigate the relationship between these modalities, cluster labels derived from each analysis were projected onto the embedding generated from the other modality. In particular, electrophysiological cluster assignments were visualized on the UMAP embedding computed from plasticity-related gene expression. In this representation, neurons are positioned according to similarity in plasticity gene expression while colors indicate their electrophysiological excitability class.

Conversely, plasticity cluster assignments were projected onto the electrophysiological embedding derived from electrophysiological feature analysis. This reciprocal visualization allows inspection of whether neurons occupying similar excitability regimes exhibit coherent transcriptional signatures related to plasticity mechanisms.

Contingency tables were constructed to compare electrophysiological and transcriptomic cluster assignments in order to study the relationship between clustering structures across the modalities. The similarity between the two clusterings was evaluated using the Adjusted Rand Index (ARI) [63] and Normalized Mutual Information (NMI) [64] metrics.

Visualization of gene expression patterns

To facilitate interpretation of molecular patterns across the excitability landscape, the expression of individual plasticity-related genes was projected onto electrophysiological embeddings. Expression values were displayed using continuous color scales across neurons in the embedding space.

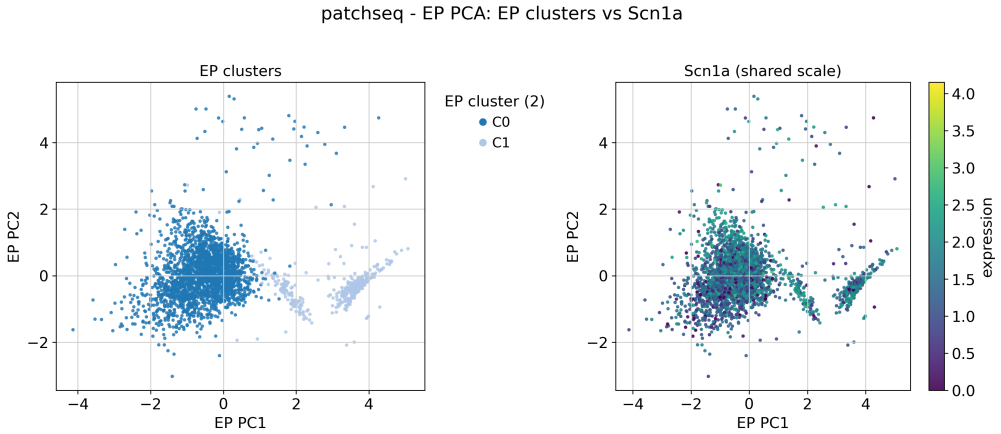


Figure 3.7: Example of Scn1a gene expression in EP PCA space.

In order to display the spatial distribution of multiple genes across the embeddings, multipanel visualization grids were generated and gene-expression patterns at the cluster level were further summarized using dot plots, matrix plots, and heatmaps, allowing comparison of transcriptional signatures across plasticity clusters.

Interpretation of plasticity-related gene expression patterns Plasticity-related genes were observed across multiple electrophysiological excitability regimes rather than being confined to a single transcriptomic cluster. This distribution

is consistent with the interpretation that gene expression reflects the molecular capacity for plastic change rather than a discrete plasticity state. In addition, transcriptomic measurements obtained through single-cell RNA sequencing represent a snapshot of gene expression at a single time point and are subject to technical limitations such as dropout events and limited transcript detection sensitivity. Consequently, the absence of detected expression for a given gene does not necessarily imply that the gene is not expressed in the cell [65, 66].

Furthermore, electrophysiological recordings capture neuronal dynamics on millisecond time scales, whereas transcriptional responses to neuronal activity typically unfold over longer time scales ranging from minutes to hours [67, 68]. As a result, electrophysiological activity states and transcriptional plasticity markers reflect complementary but temporally distinct aspects of neuronal function. The presence of plasticity-related transcripts across multiple excitability clusters therefore likely reflects a distributed regulatory landscape in which neurons across different dynamical regimes retain the molecular machinery required for activity-dependent plastic modification.

Chapter 4

Results

This chapter presents the results of the electrophysiological and transcriptomic analyses described in Chapter 3. Results are organized in three stages. First, I characterize excitability structure using first-spike features in the PatchClamp and Patch-seq datasets. Second, I show how whole-signal dynamical descriptors refine this organization by capturing ramp-response trajectory geometry. Third, I relate electrophysiological excitability structure to transcriptomic variation in plasticity-related gene expression in Patch-seq neurons.

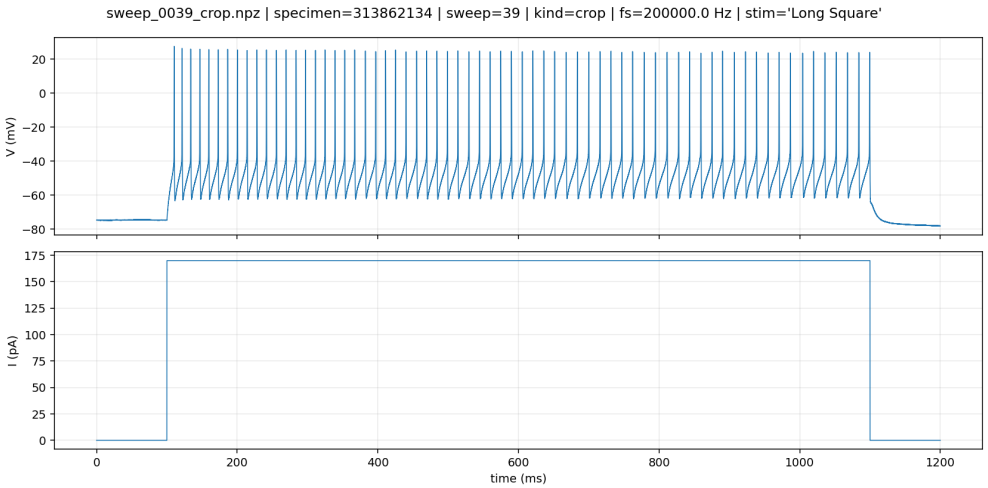
4.1 Dataset retention after extraction and quality control

The PatchClamp dataset contains 1,964 neurons reported in the original study [21], whereas the Patch-seq dataset comprises over 4,200 interneurons characterized using the Patch-seq pipeline [22]. The number of neurons and sweeps retained for the present analysis reflects the subset of recordings successfully processed and passing the sweep-level quality-control criteria described in Chapter 3.

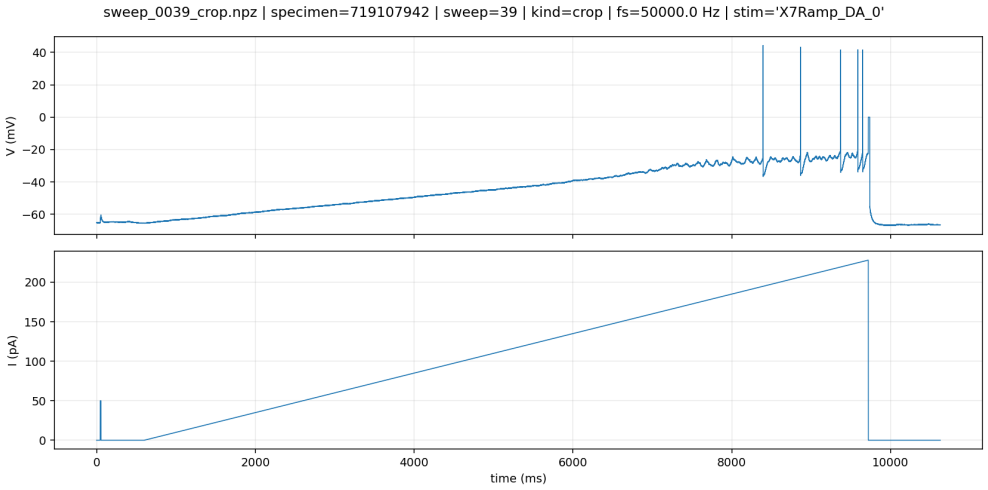
The sweeps that passed quality control are the ones that display stable post-stimulus dynamics and thus provide a sufficiently long recording window for biologically meaningful feature extraction.

Dataset	Stimulus	Neurons	Total sweeps	QC retained	QC rejected	Mean sweeps per neuron
PatchClamp	Long square	1964	27566	24962	2604	14 total / 12 QC
PatchClamp	Ramp	1964	5541	5094	447	2.8 total / 2.6 QC
Patch-seq	Ramp	2596	11599	9202	2397	4.5 total / 3.5 QC

Table 4.1: Summary of the datasets and sweep-level quality-control outcomes used in this study. For each dataset and stimulation protocol, the table reports the number of analyzed neurons, the total number of extracted sweeps, the number of sweeps retained after quality control, the number of rejected sweeps, and the average number of total and QC-retained sweeps per neuron. PatchClamp long-square and ramp stimuli were analyzed separately, whereas only ramp stimuli were considered for Patch-seq.



(a) Example long-square sweep retained after quality control



(b) Example ramp sweep retained after quality control

Figure 4.1: Representative electrophysiological sweeps retained after quality control. (A) Example long-square current injection response from the PatchClamp dataset. (B) Example ramp stimulation response illustrating the gradual depolarizing drive used in the dynamical analysis. Only sweeps passing the quality-control criteria described in Chapter 3 were included in downstream feature extraction.

4.2 First-spike electrophysiological organization of neuronal excitability

4.2.1 PatchClamp first-spike analysis

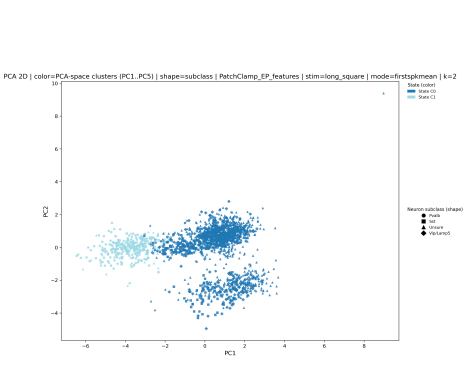
Stim.	Protocol	Space	k	PCs	Silh. ↑	DB ↓	CH ↑
Long-square stimulation							
Long-square	rheo-first	PCA	3	4	0.414	0.923	**1322.6**
Long-square	**firstspkmean**	PCA	2	5	**0.421**	0.864	1028.4
Long-square	firstspkmean	PCA	3	5	0.396	0.963	1090.8
Long-square	combo	PCA	2	5	0.412	**0.836**	991.4
Ramp stimulation							
Ramp	**firstspk**	PCA	2	4	**0.447**	**0.793**	**957.6**
Ramp	firstspkmean	PCA	2	5	0.404	1.022	664.1
Ramp	combo	PCA	2	5	0.443	0.886	888.9

Table 4.2: Summary of the selected clustering solutions across stimulation protocols and feature representations. All analyses used agglomerative clustering with Ward linkage and a feature-correlation pruning threshold of 0.85. The number of principal components (PCs) was selected as the minimum required to exceed 85% cumulative explained variance.

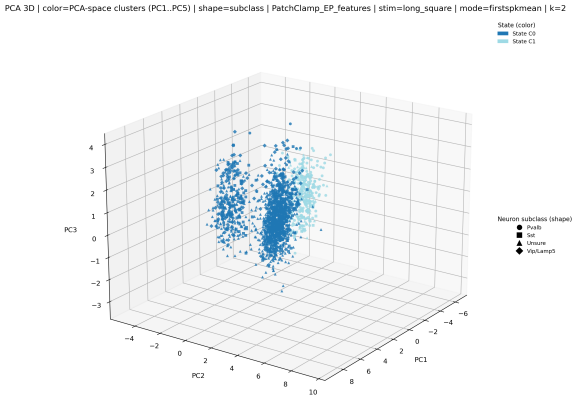
Clustering of first-spike electrophysiological features revealed reproducible low-dimensional excitability structure in both long-square and ramp stimulation conditions (Table 4.2). Across all tested representations, the selected solutions were obtained in PCA space after retaining the minimum number of principal components required to exceed 85% cumulative explained variance. For long-square stimulation, the `firstspkmean` representation yielded the most balanced solution, with a two-cluster partition obtained from the first five principal components and the highest silhouette score among the long-square models (0.421). Alternative long-square solutions showed comparable but slightly weaker separation, with the `combo` representation producing the lowest Davies-Bouldin index and the `rheo-first` representation reaching the highest Calinski-Harabasz score, indicating that the broad excitability structure was robust across feature parameterizations.

Stimulation	Representation	Features kept	Top discriminating features
Long-square	rheo-first	9	Latency, v_{rest}, (*Slope_deep*), spike width, upstroke/downstroke ratio.
Long-square	firstspkmean (k2)	9	Upstroke/downstroke ratio, (*Slope_deep*), spike width, dV_{deep}, AHP slope.
Long-square	firstspkmean (k3)	9	Latency, rheobase, AHP depth, upstroke/downstroke ratio, (*Slope_deep*).
Long-square	combo	10	Upstroke/downstroke ratio, (*Slope_deep*), spike width, AHP slope, latency.
Ramp	firstspk	9	(*Slope_deep*), upstroke/downstroke ratio, dV_{deep}, spike width, AHP slope.
Ramp	firstspkmean	8	Upstroke/downstroke ratio, dV_{deep}, spike width, AHP slope, latency.
Ramp	combo	10	Upstroke/downstroke ratio, (*Slope_deep*), dV_{deep}, spike width, AHP slope.

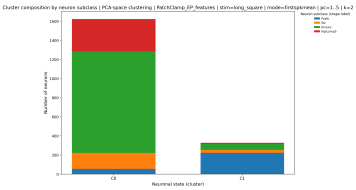
Table 4.3: Features with the highest mutual information with respect to the cluster labels were considered the top discriminating features for clustering.



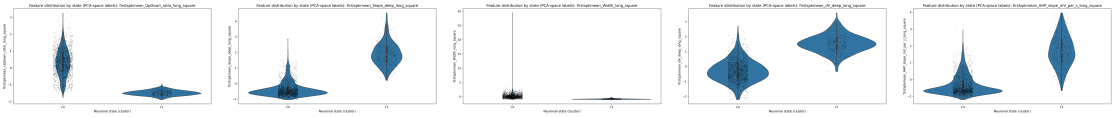
(a) PCA 2D



(b) PCA 3D

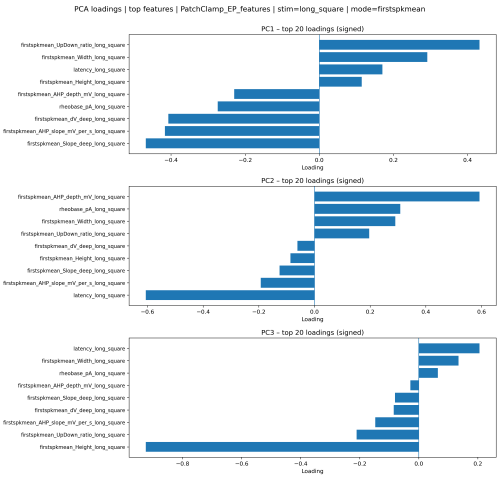


(c) Subclass composition

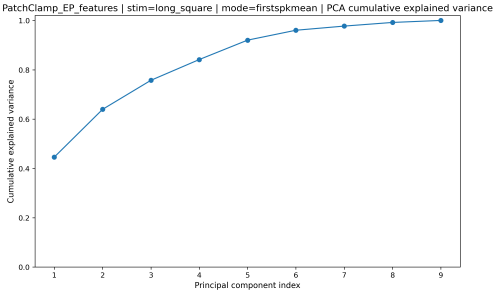


(d) UpDown ratio (e) Slope_deep (f) Width (g) dV_{deep} (h) AHP slope

Figure 4.2: Selected long-square first-spike clustering solution (`firstspkmean`, PCA space, $k = 2$, 5 PCs). The figure shows the PCA embedding in two and three dimensions, the cluster composition by neuronal subclass, and violin plots of the five most discriminating electrophysiological features.



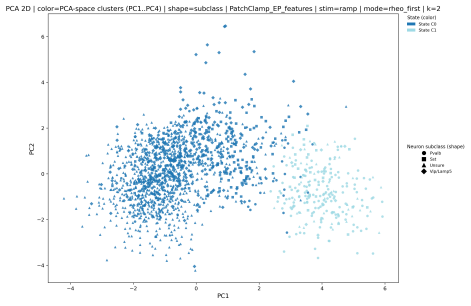
(a) PCA loadings



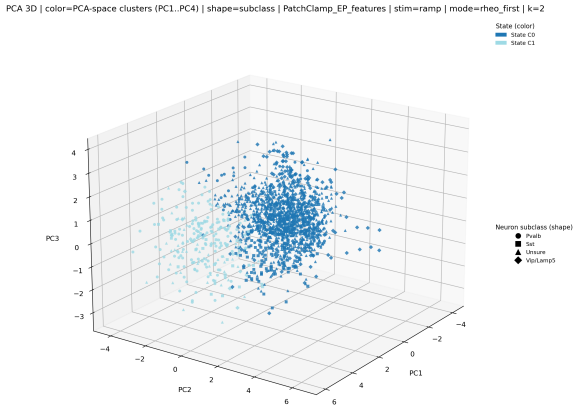
(b) Cumulative explained variance

Figure 4.3: PCA diagnostics for the selected long-square first-spike clustering solution (`firstspkmean`, PCA space, $k = 2$, 5 PCs), including the loadings of the main electrophysiological features on the first principal components and the cumulative explained variance used to select the retained dimensionality.

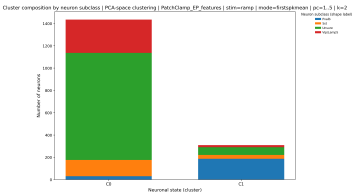
For ramp stimulation, the `firstspk` representation produced the strongest overall clustering solution, with a two-cluster partition in the first four principal components and the best values across all three internal validation metrics (silhouette = 0.447, Davies-Bouldin = 0.793, Calinski-Harabasz = 957.6). Inspection of mutual-information rankings indicated that cluster separation was driven primarily by spike-shape and repolarization descriptors, particularly *Slope_deep*, upstroke/downstroke ratio, spike width, dV_{deep}, and AHP slope, while latency, rheobase, and v_{rest} contributed more strongly in selected long-square solutions (Table 4.3). Together, these results indicate that first-spike features recover a compact but interpretable excitability structure, with ramp responses showing slightly clearer state separation than long-square responses.



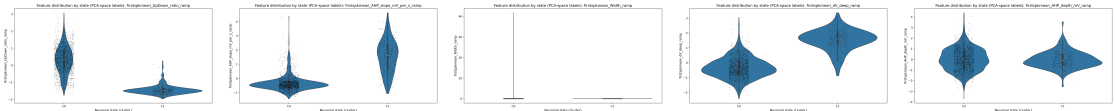
(a) PCA 2D



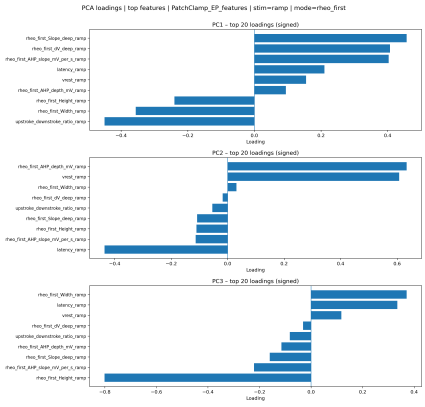
(b) PCA 3D



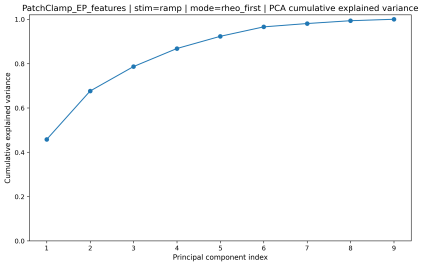
(c) Subclass composition



(d) UpDown ratio **(e)** AHP slope **(f)** Width **(g)** dV_{deep} **(h)** AHP depth



(i) PCA loadings



(j) Cumulative explained variance

Figure 4.4: Selected ramp first-spike clustering solution (`firstspk`, PCA space, $k = 2$, 4 PCs). The figure shows the PCA embedding in two and three dimensions, cluster composition by neuronal subclass, violin plots of the five most discriminating electrophysiological features, and PCA diagnostics including feature loadings and cumulative explained variance.

4.2.2 Patch-seq ramp-based first-spike analysis

The first-spike clustering analysis was next applied to the Patch-seq electrophysiological dataset, restricting the analysis to ramp stimulation responses, which, compared to long-square stimulation, probe a broader dynamic range and more closely resemble natural patterns of neuronal activation in the brain [14]. Across the three feature representations considered (`firstspk`, `firstspkmean`, and `combo`), clustering solutions consistently favored a two-cluster partition in PCA space (Table 4.4). The `firstspk` representation yielded the highest silhouette score (0.356), indicating the clearest separation between clusters, while the `combo` representation achieved the lowest Davies-Bouldin index and the highest Calinski-Harabasz score among the tested configurations.

Representation	**Space**	**k**	**PCs**	**Silh.↑**	**DB↓**	**CH↑**
firstspk	PCA	2	5	**0.356**	1.091	**1105.7**
firstspkmean	PCA	2	6	0.341	1.147	971.5
combo	PCA	2	6	0.329	**1.082**	838.8

Table 4.4: Summary of the selected clustering solutions for the Patch-seq first-spike analysis. Only ramp stimulation responses were considered in this dataset. All analyses used agglomerative clustering with Ward linkage and a feature-correlation pruning threshold of 0.85. The number of principal components (PCs) was selected based on a cumulative explained variance exceeding 85%.

Representation	**Features kept**	**Top discriminating features**
firstspk	9	Upstroke/downstroke ratio, *Slope_deep*, dV_{deep}, AHP slope, latency.
firstspkmean	10	Upstroke/downstroke ratio, dV_{ratio}, dV_{deep}, *Slope_deep*, AHP slope.
combo	10	*Slope_deep*, dV_{deep}, upstroke/downstroke ratio, AHP slope, spike width.

Table 4.5: Top discriminating electrophysiological features for the Patch-seq first-spike clustering solutions. Only ramp stimulation responses were analyzed in this dataset. Features listed correspond to those with the highest mutual information with respect to cluster labels.

Inspection of mutual-information rankings revealed that cluster separation was primarily driven by first-spike repolarization and recovery descriptors (Table 4.5). In

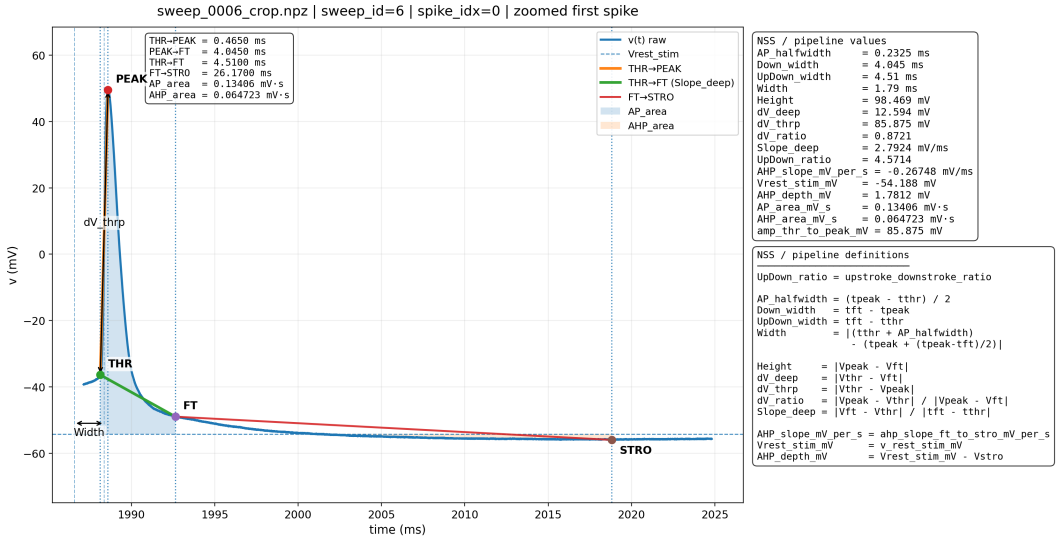
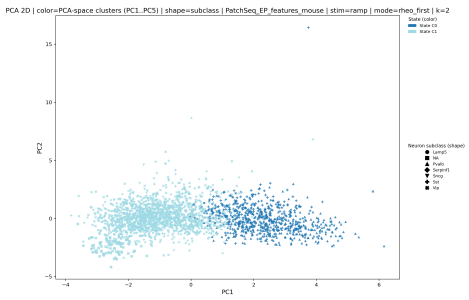


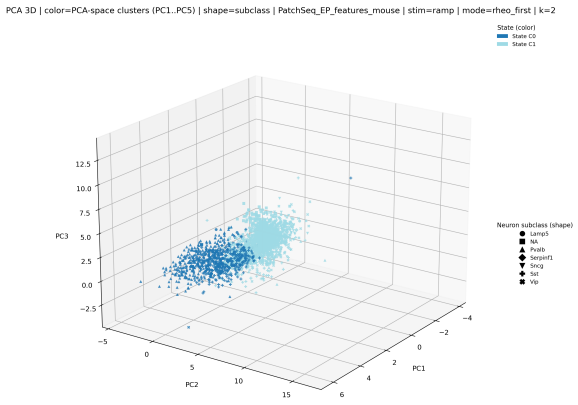
Figure 4.5: First action potential extracted from a current-clamp ramp stimulus. Key spike landmarks are indicated: threshold (THR), peak (PEAK), fast trough (FT), and slow trough (STRO). Highlighted segments correspond to the rise and recovery phases used to compute NSS spike features. The shaded regions represent the depolarization (AP area) and afterhyperpolarization (AHP area) relative to the resting membrane potential. Feature values and their mathematical definitions are reported in the annotation boxes.

particular, the upstroke/downstroke ratio, repolarization slope (*Slope_deep*), depth of repolarization (dV_{deep}), and afterhyperpolarization slope consistently appeared among the most discriminating features across all representations. Latency and spike width contributed in some representations but played a secondary role. These results closely mirror the feature importance patterns observed in the PatchClamp dataset, suggesting that early spike waveform geometry and repolarization dynamics constitute a robust axis of electrophysiological variability that is preserved across both electrophysiology-only and multimodal Patch-seq recordings.

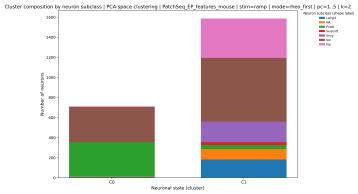
The selected clustering solutions are illustrated in Fig. 4.6. The figure shows the PCA embedding together with the composition of each cluster by transcriptomic subclass and the distributions of the electrophysiological features that most strongly distinguish the clusters.



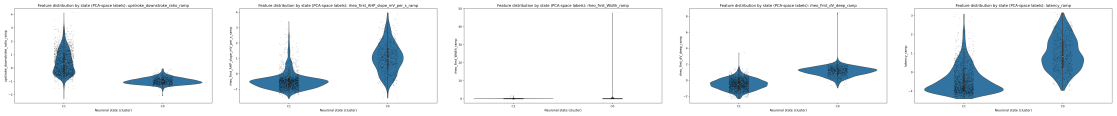
(a) PCA 2D



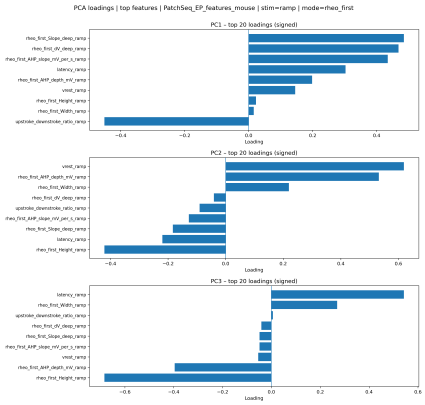
(b) PCA 3D



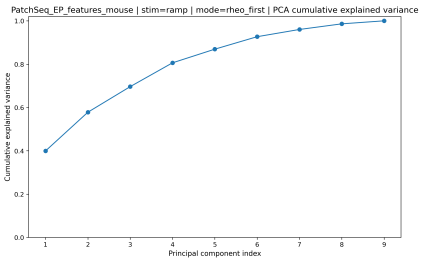
(c) Subclass composition



(d) UpDown ratio **(e)** AHP slope **(f)** Width **(g)** dV_{deep} **(h)** Latency



(i) PCA loadings



(j) Cumulative explained variance

Figure 4.6: Selected ramp first-spike clustering solution (`firstspk`, PCA space, $k = 2$, 5 PCs). The figure shows the PCA embedding in two and three dimensions, cluster composition by neuronal subclass, violin plots of the five most discriminating electrophysiological features, and PCA diagnostics including feature loadings and cumulative explained variance.

4.3 Whole-signal excitability regimes

To evaluate how different dynamical descriptors capture voltage dynamics, we compared multiple feature modalities defined in Section 3.6.9. The most informative configurations were then retained for in-depth analysis to probe the excitability states emerging from ramp responses.

4.3.1 PatchClamp whole-signal analysis

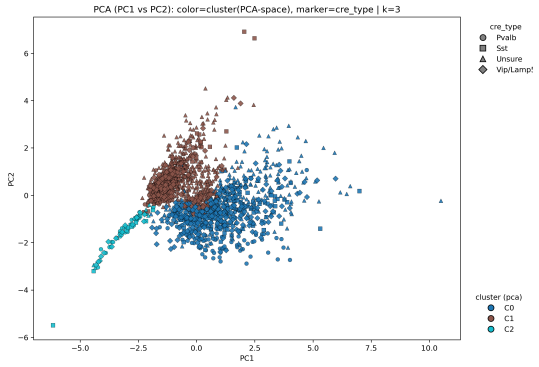
Whole-signal dynamical features extracted from ramp stimulation responses revealed distinct excitability regimes when analyzed in PCA space (Table 4.6).

Representation	**k**	**Feat.**	**PCs**	**Silh. ↑**	**DB ↓**	**CH ↑**
minimal_ramp	3	7	5	0.294	1.158	506.54
poincare_only	2	4	3	**0.370**	1.228	893.9
increment_poincare_only	9	10	6	0.212	1.080	339.3
derivative_geometry_only	2	3	3	**0.712**	**0.373**	**938.1**
dynamics_only	3	11	6	0.289	1.402	351.38
increment_plus_derivative	2	11	6	0.298	1.802	324.7
poincare_plus_derivative	6	6	4	0.288	0.966	**788.36**
increment_derivative_delta	2	11	6	0.298	1.802	324.7
poincare_increment_derivative	2	12	7	**0.537**	**0.747**	303.7
global_full_only	8	19	7	0.282	1.000	351.8
mesh_hull_only	2	25	17	0.701	1.706	240.42
extended_ramp	3	30	17	0.219	1.915	206.1

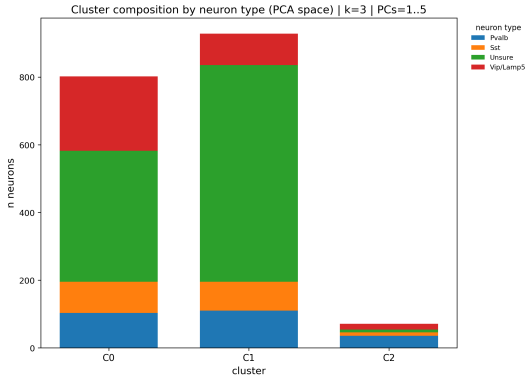
Table 4.6: Summary of clustering solutions obtained for the PatchClamp whole-signal analysis under different feature representations. Ramp stimulation responses were analyzed using agglomerative clustering with Ward linkage in PCA space. Reported metrics include silhouette score, Davies-Bouldin index (DB), and Calinski-Harabasz index (CH). Bold values highlight the best-performing solutions across validation metrics.

Among the evaluated feature representations, the `minimal_ramp` modality provided a compact and interpretable description of ramp dynamics, producing a stable

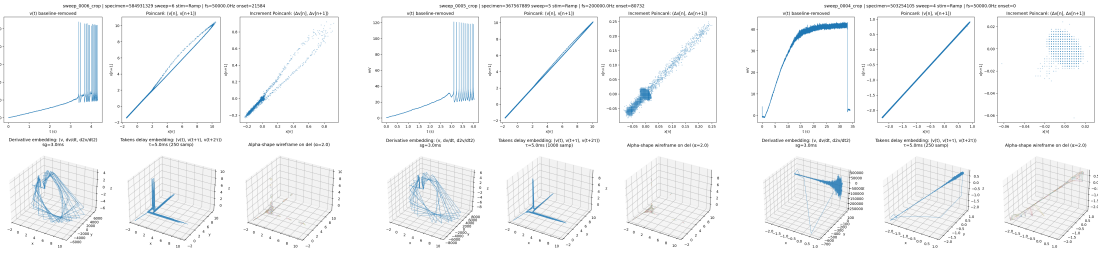
clustering solution with three clusters while relying on a small set of trajectory-based descriptors. Although some alternative modalities achieved higher individual validation metrics, the `minimal_ramp` representation offered a balanced compromise between clustering quality, dimensional compactness, and interpretability of the underlying dynamical features.



(a) PCA embedding (PC1 PC2) of neurons using the minimal ramp representation, colored by electrophysiological cluster



(b) Transcriptomic subclass composition of the electrophysiological clusters



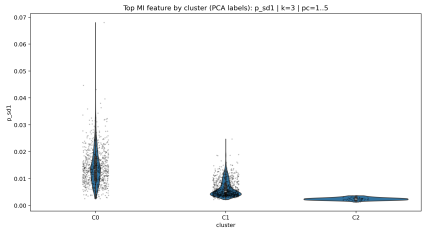
(c) State 0 **(d)** State 1 **(e)** State 2

Figure 4.7: Minimal ramp representation reveals three electrophysiological response regimes.

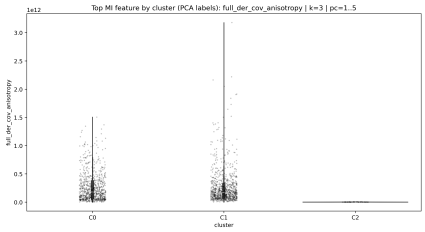
Two additional representations were retained for comparison because they capture complementary aspects of the voltage trajectory. The `poincare_only` representation yielded a clear two-cluster solution with a relatively high silhouette score, indicating that short-timescale variability of the voltage signal alone already captures a substantial portion of the structure underlying ramp responses. The `poincare_plus_derivative` representation, which integrates Poincaré variability statistics with geometric descriptors of the derivative embedding, produced a moderately separated but richer clustering structure, suggesting that

combining signal-level variability with phase-space geometry can reveal additional organization of the response trajectories. Together, these three representations provide complementary views of ramp-evoked dynamics: a compact summary of excitability regimes (`minimal_ramp`), a variability-driven representation (`poincare_only`), and a mixed variability-geometry description of trajectory structure (`poincare_plus_derivative`).

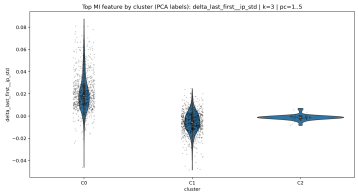
Inspection of mutual-information rankings indicated that cluster separation in the `minimal_ramp` representation was primarily driven by temporal changes in trajectory statistics and anisotropy measures, particularly features describing differences between early and late ramp dynamics (e.g. Δ late-early *ip_std* and Δ late-early *ip_mean*), together with variability descriptors such as *p_sd*1 and increment-based statistics (*ip_mean*). These features reflect how the statistical structure of the voltage trajectory evolves as the ramp stimulus progresses as highlighted in 4.7.



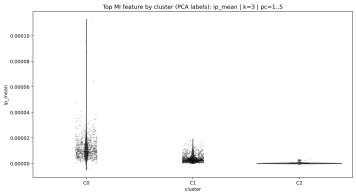
(a) *p_sd*1



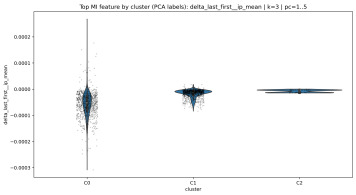
(b) *full_der_cov_anisotropy*



(c) Δ late-early *ip_std*



(d) *ip_mean*



(e) Δ late-early *ip_mean*

Figure 4.8: Features driving cluster separation in the `minimal_ramp` representation

Feature	Technical definition
Δ late-early *ip_std*	Difference between the standard deviation of the increment Poincaré signal in the late and early time windows, computed from successive voltage increments (Δv_n). Indicates how the variability of rapid voltage transitions changes during the ramp response.
*p_sd*1	SD1 of the Poincaré representation computed from the voltage trace itself, using pairs (v_n, v_{n+1}). Represents short-timescale variability of the membrane-potential trajectory.
ip_mean	Mean value of the increment Poincaré signal computed from successive voltage differences (Δv_n). Reflects the average magnitude of rapid voltage changes during the ramp response.
Δ late-early *ip_mean*	Difference between the late- and early-window mean of the increment Poincaré signal. Describes how the magnitude of rapid voltage changes evolves over the ramp.
full_del_cov_anisotropy	Anisotropy of the covariance matrix computed from the full delay-coordinate embedding $(v(t), v(t+\tau), v(t+2\tau))$. Indicates how elongated the reconstructed trajectory is in delay-embedding space.
full_der_cov_anisotropy	Anisotropy of the covariance matrix computed from the derivative embedding $(v, dv/dt, d^2v/dt^2)$. Describes directional structure in the geometry of the voltage trajectory.
*ip_sd*1	SD1 of the increment Poincaré representation computed from pairs of successive increments $(\Delta v_n, \Delta v_{n+1})$. Represents short-timescale variability of rapid voltage transitions.

Table 4.7: Minimal ramp-response dynamical descriptors kept after preprocessing. The notation p_* refers to descriptors computed from the Poincaré representation of the voltage trace, while ip_* refers to descriptors computed from the increment Poincaré representation based on successive voltage differences. The selected features capture variability of the voltage signal, its temporal evolution during the ramp, and the geometry of the reconstructed trajectory.

For the `poincare_only` representation, the most informative features were primarily measures of short-timescale variability of the signal, especially p_sd1 and its windowed variants across different ramp segments, indicating that cluster separation is largely driven by differences in local dispersion of the Poincaré map.

In the `poincare_plus_derivative` representation, the highest-ranking features combined variability statistics and geometric descriptors of the derivative embedding, including p_sd1, covariance eigenvalues of the derivative phase space, and bounding-box measures of trajectory extent. This suggests that clustering in this representation reflects both differences in signal variability and changes in the geometric structure of the reconstructed trajectory in derivative phase space.

4.3.2 Patch-seq whole-signal analysis

The Patch-seq whole-signal analysis was performed using the same approach adopted for the PatchClamp dataset, considering only ramp stimulation responses. The explored run configurations are summarized below.

Representation	**Corr. thr.**	**k**	**Feat.**	**PCs**	**Silh.↑**	**DB↓**	**CH↑**
minimal_ramp							
minimal_ramp	**0.70**	**2**	**4**	**3**	**0.619**	**0.660**	**2597.6**
minimal_ramp	0.70	5	4	3	0.380	0.839	1978.2
minimal_ramp	0.70	6	4	3	0.377	0.755	2124.2
minimal_ramp	0.85	2	5	3	0.467	0.612	2309.7
minimal_ramp	0.85	3	5	3	0.466	0.770	2717.8
minimal_ramp	0.90	2	6	3	0.557	0.747	2076.0
minimal_ramp	0.90	3	6	3	0.476	0.765	2628.8
minimal_ramp	**0.90**	**4**	**6**	**3**	**0.491**	**0.872**	**2347.6**
other modalities							
poincare_only		6	6	4	0.474	0.716	1177.9
incr_poincare_only	0.70	2	11	7	0.315	1.579	452.03
der_geometry_only	0.70	2	3	2	0.433	0.819	1278.8
dynamics_only	0.70	7	12	7	0.340	0.810	718.47
incr_plus_derivative	0.70	5	12	7	0.269	1.189	523.9
poincare_plus_der	0.70	7	7	5	0.404	0.681	858.6
incr_derivative_delta	0.70	4	12	7	0.262	1.342	510.5
poincare_incr_der	0.70	8	13	8	0.299	0.893	554.3
global_full_only	0.70	2	12	9	0.448	1.403	545.9
mesh_hull_only	0.70	2	27	17	0.484	1.714	264.47
extended_ramp	0.70	3	34	19	0.373	1.288	307.7

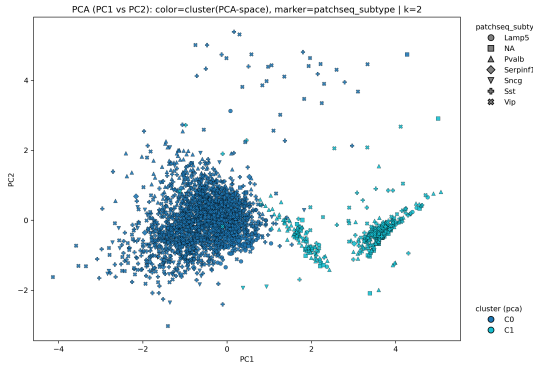
Table 4.8: Comparison of clustering solutions for the Patch-seq whole-signal analysis across feature representations and correlation-pruning thresholds. Selected solutions used in downstream analyses are highlighted in bold. Clustering was performed in PCA space using agglomerative clustering with Ward linkage. Internal validation was evaluated using the silhouette score, Davies-Bouldin index (DB), and Calinski-Harabasz index (CH).

Overall, the compact `minimal_ramp` representation produced the strongest and most consistent clustering solutions across the explored correlation thresholds. In particular, the solution with feature pruning correlation threshold 0.70 and $k = 2$

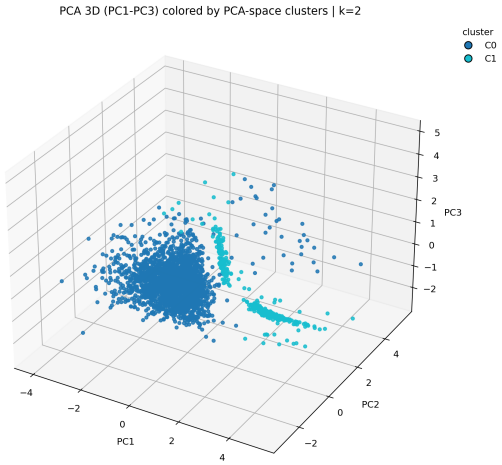
achieved the best overall internal validation profile, with the highest silhouette score and Calinski-Harabasz index together with a low Davies-Bouldin value, indicating a robust separation into two major excitability regimes.

A second `minimal_ramp` solution was retained with a feature-correlation threshold of 0.90 and $k = 4$. Compared with the $k = 2$ configuration, this solution produces a more detailed partition of the same compact feature space, while still maintaining strong internal validation scores and provides therefore a way to test whether the coarse two-state organization can be resolved into finer excitability subregimes without sacrificing interpretability.

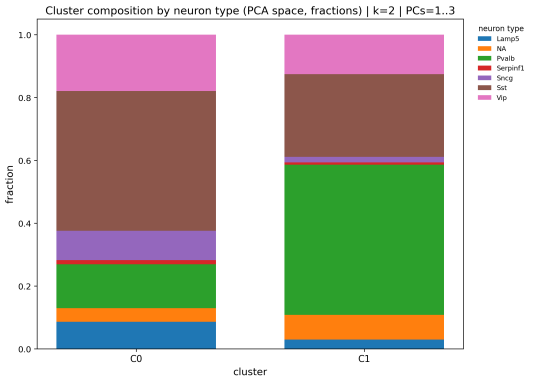
The compact partition (`minimal_ramp`, $k = 2$) and the finer subdivision within the same descriptor space (`minimal_ramp`, $k = 4$) are the two winning configurations on which subsequent analyses focus as they are the most informative representations of whole-signal excitability states in the PatchSeq dataset.



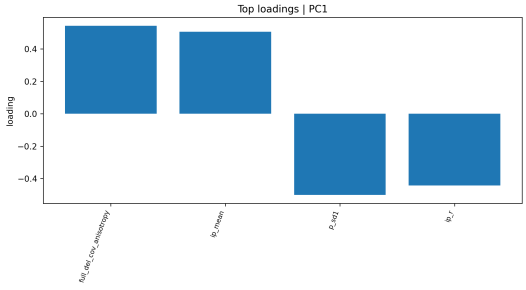
(a) PCA 2D



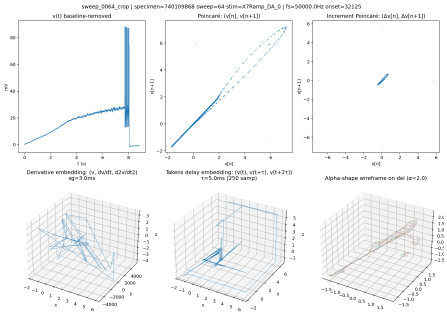
(b) PCA 3D



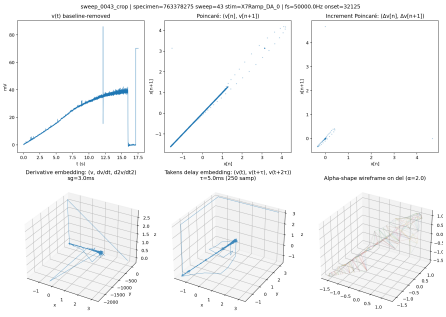
(c) Subclass composition



(d) PCA loadings

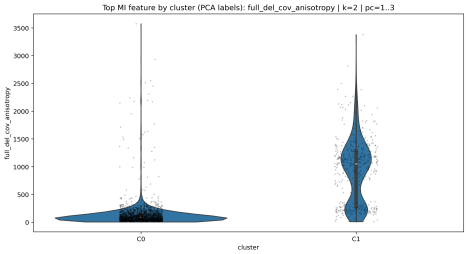


(e) Cluster C0

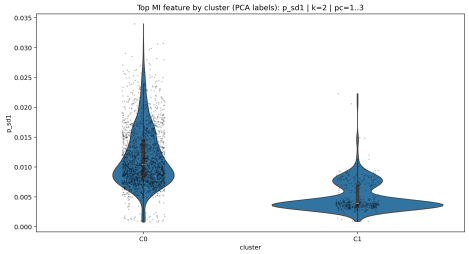


(f) Cluster C1

Figure 4.9: Minimal ramp, k=2

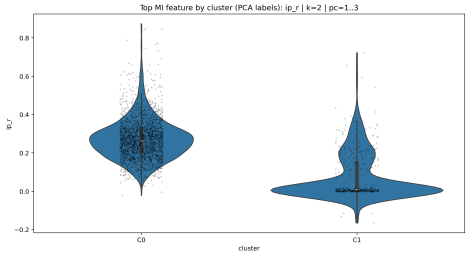


(a) *full_der_cov_anisotropy*



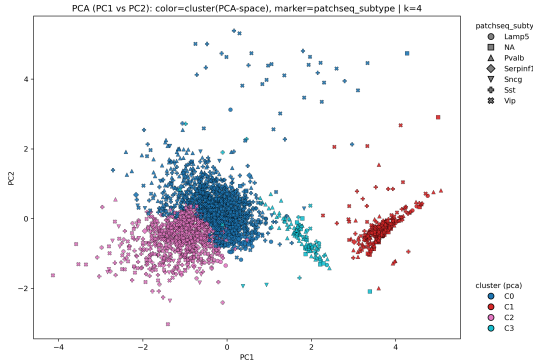
(b) *p_sd1*

(c) *ip_mean*

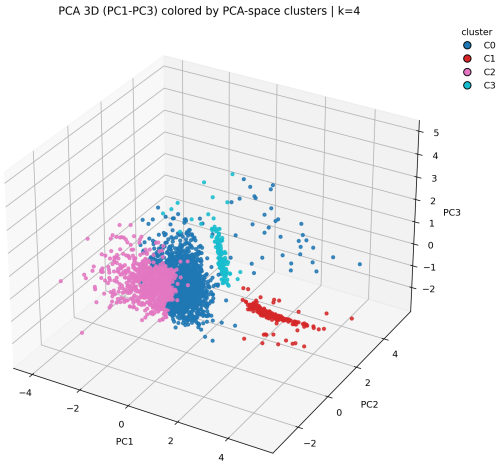


(d) *ip_r*

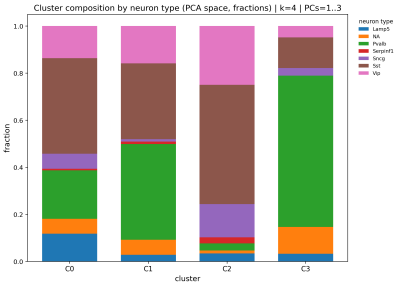
Figure 4.10: Features driving cluster separation in the `minimal_ramp` (k=2) representation



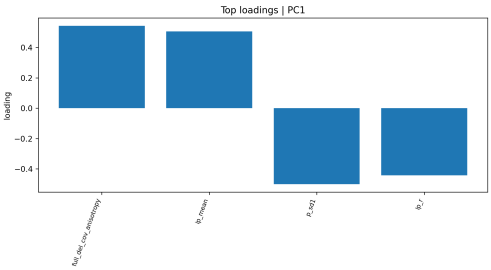
(a) PCA 2D



(b) PCA 3D

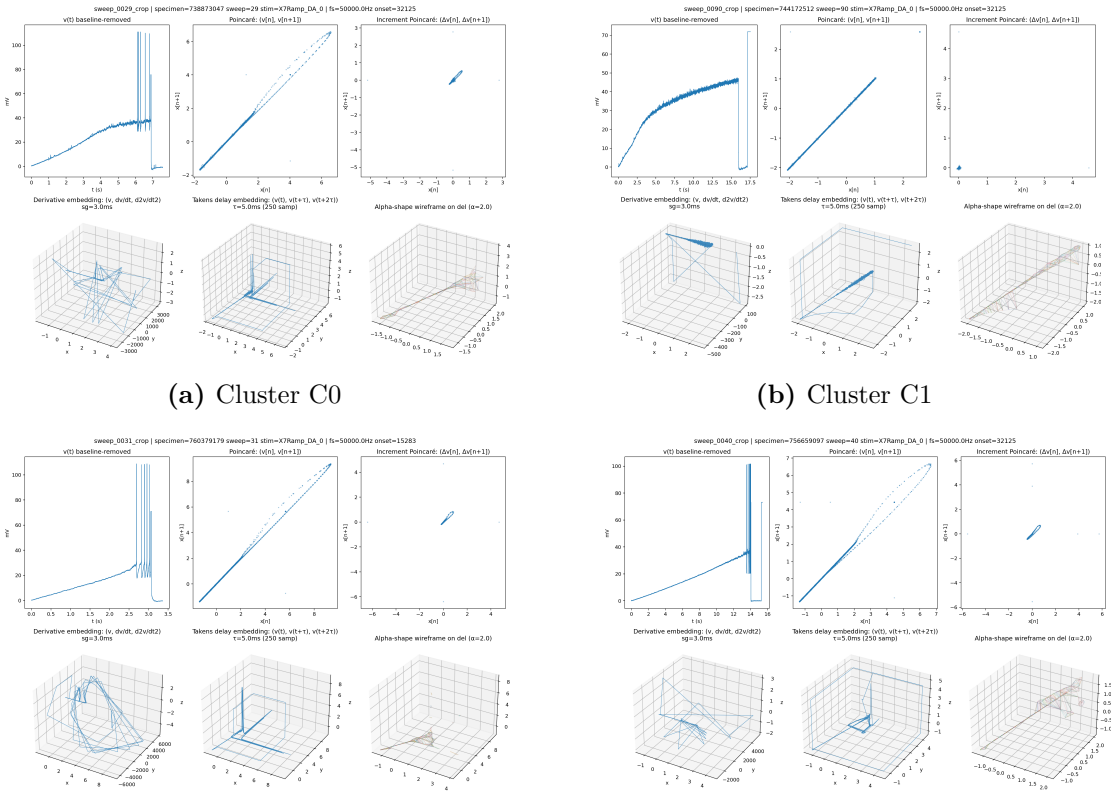


(c) Subclass composition



(d) PCA loadings

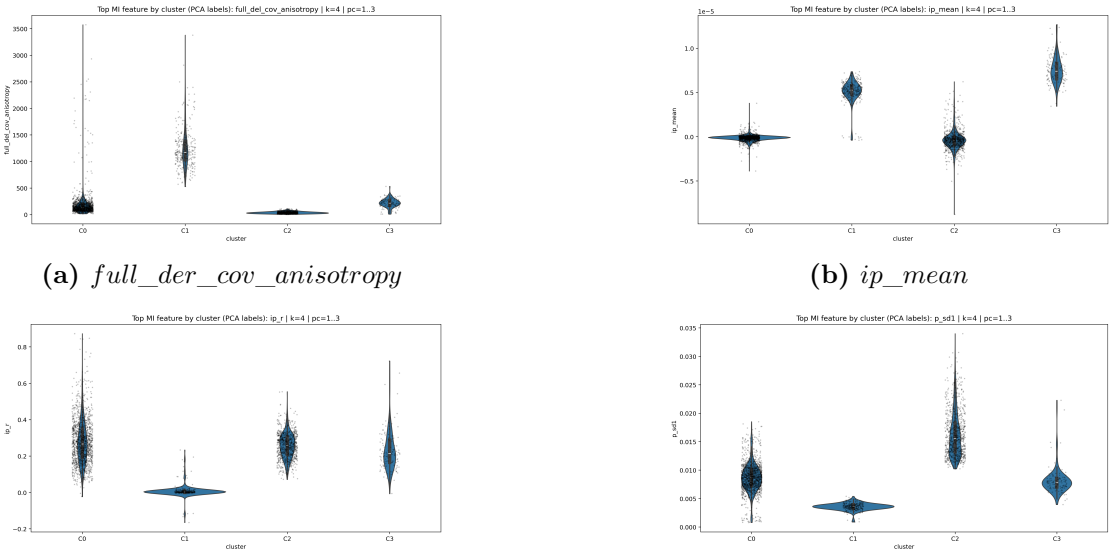
Figure 4.11: Minimal ramp, k=4



(a) Cluster C0 **(b)** Cluster C1

(c) Cluster C2 **(d)** Cluster C3

Figure 4.12: Minimal ramp EP state embeddings, k=4



(a) *full_der_cov_anisotropy*

(b) *ip_mean*

(c) *ip_r*

(d) *p_sd*1

Figure 4.13: Features driving cluster separation in the `minimal_ramp` (k=4) representation

4.4 Comparison between first-spike and whole-signal excitability analyses

The first-spike and whole-signal analyses together provide complementary views of neuronal excitability.

The first-spike analysis captures a compact near-threshold snapshot of excitability, primarily driven by waveform geometry and repolarization-related features such as upstroke/downstroke ratio, *Slope_deep*, dV_{deep}, spike width, and afterhyperpolarization descriptors. These features consistently emerged among the most discriminating variables across both PatchClamp and Patch-seq ramp recordings, indicating that early spike-shape geometry defines a robust axis of electrophysiological variability. Instead, the whole-signal analysis approaches excitability from a different perspective because rather than focusing on a single action potential, it considers the membrane-potential trace during the entire ramp stimulus and analyzes how its statistical and geometric structure evolves over time. The most informative whole-signal configurations had cluster separation mainly associated with temporal-difference descriptors and variability measures that reflect changes in voltage fluctuations, spike recruitment dynamics, and the overall organization of the trajectory during the ramp response, including Δ late-early *ip_std*, Δ late-early *ip_mean*, p_{sd1}, *ip_mean*, and anisotropy metrics derived from reconstructed derivative state-space embeddings.

In both datasets, the excitability regimes identified from electrophysiology showed only partial correspondence with transcriptomic subclasses, suggesting that excitability reflects a dynamic functional state rather than a fixed cell-type identity.

The first-spike framework focuses on local waveform properties at spike onset, while the whole-signal analysis captures the broader dynamical regime of the neuron during sustained depolarizing stimulation and for this reason, the two approaches are better interpreted as complementary rather than alternative analyses. A quantitative recap of the main clustering results across datasets is reported in Table 4.9, and the conceptual differences between the two analytical frameworks are summarized in Table 4.10.

Dataset	Analysis	Best solution	Top MI features
PatchClamp	First-spike ($k = 2$)	`firstspk` mode; Fig. 4.4; 4 PCs retained; silhouette = 0.447; DB = 0.793; CH = 957.6	UpDown ratio, AHP slope, width, dV_{deep}, AHP depth
PatchClamp	Whole-signal ($k = 3$)	`minimal_ramp` mode; Fig. 4.7; 5 PCs retained; silhouette = 0.294; DB = 1.158; CH = 506.54	Δ late–early *ip_std*, Δ late–early *ip_mean*, p_{sd1}, *ip_mean*, full derivative covariance anisotropy
Patch-seq	First-spike ($k = 2$)	`firstspk` mode; Fig. 4.6; 5 PCs retained; silhouette = 0.356; DB = 1.091; CH = 1105.7	UpDown ratio, *Slope_deep*, dV_{deep}, AHP slope, latency
Patch-seq	Whole-signal ($k = 2$)	`minimal_ramp` mode; Fig. 4.9; 3 PCs retained; silhouette = 0.619; DB = 0.660; CH = 2597.6	full derivative covariance anisotropy, *ip_mean*, ip_r, p_{sd1}
Patch-seq	Whole-signal ($k = 4$)	`minimal_ramp` mode; Fig. 4.11; 3 PCs retained; silhouette = 0.491; DB = 0.872; CH = 1177.9	full derivative covariance anisotropy, *ip_mean*, ip_r, p_{sd1}

Table 4.9: Recap of the main clustering results obtained across datasets and electrophysiological analysis strategies. For each dataset and analysis type, the table reports the selected feature representation, the corresponding figure, the number of retained principal components, internal clustering metrics, and the most informative features according to mutual-information ranking.

Aspect	**First-spike analysis**	**Whole-signal analysis**
Unit of analysis	First evoked action potential during the ramp stimulus	Entire ramp-evoked voltage trajectory
Main discriminating features	Upstroke/downstroke ratio, *Slope_deep*, dV_{deep}, spike width, AHP slope/depth, latency	Δ late-early *ip_std*, Δ late-early *ip_mean*, p_{sd1}, *ip_mean*, derivative/delay anisotropy
What the features describe	Waveform geometry of the first spike, repolarization steepness, afterhyperpolarization, and near-threshold spike timing	Evolution of signal variability across the ramp, spike recruitment dynamics, short-timescale fluctuations, and trajectory geometry in reconstructed state space
Temporal scale captured	Local, near-threshold excitability snapshot	Extended dynamical regime unfolding during sustained depolarizing drive
Best-performing structure in this study	Mostly compact two-cluster solutions in both PatchClamp and Patch-seq ramp analyses	Coarse two-state and finer multi-state organizations depending on feature modality and clustering resolution
Biological interpretation	Describes how neurons initiate firing, with strong emphasis on spike-shape differences and repolarization/AHP mechanisms	Describes how neurons sustain and modulate their response over time, including adaptation, variability changes, and dynamical trajectory structure

Table 4.10: Conceptual comparison between first-spike and whole-signal excitability analyses used in this study. The first-spike approach summarizes neuronal excitability through the waveform geometry of the first action potential, whereas the whole-signal approach characterizes the full dynamical evolution of ramp-evoked responses.

4.5 Plasticity-related transcriptomic structure in Patch-seq neurons

This analysis aimed to determine whether electrophysiological excitability states correspond to differences in plasticity-related molecular programs. To address this question, we analyzed the expression of genes associated with synaptic and intrinsic plasticity in Patch-seq neurons and examined their relationship to electrophysiological ramp-response dynamics [47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59].

4.5.1 Plasticity-gene expression defines a high-dimensional regulatory landscape

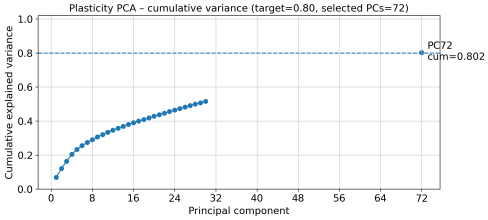
Principal component analysis (PCA) of a curated set of 135 plasticity-related genes from literature and the amiGO platform revealed a high-dimensional structure in transcriptomic space (Fig. 4.14A). Approximately 72 principal components were required to explain 80% of the total variance, indicating that plasticity-related transcriptional programs are distributed across many partially independent molecular pathways rather than being dominated by a small number of global axes.

Examining the genes with the highest loadings across the principal components revealed as shown in (Fig. 4.14B,C) that some components were strongly associated with activity-dependent transcriptional regulators such as *Egr1* and *Arc*, while other components were mainly driven by genes involved in postsynaptic signaling and synaptic scaffolding, including *Dlg4*, *Syngap1*, and *Camk2a/b*.

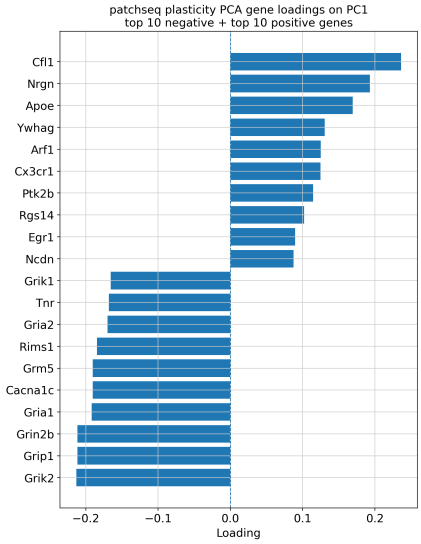
Additional components reflected variation in glutamatergic receptor signaling, with strong contributions from NMDA and AMPA receptor subunits (*Grin*, *Gria*, and *Grik* gene families). A distinct group of components was enriched for neuromodulatory signaling pathways, including dopamine, serotonin, and adrenergic receptor genes such as *Drd1*, *Htr7*, and *Adrb1*. Finally, several components captured variation in extracellular matrix and structural plasticity genes including *Acan*, *Bcan*, *Hapln1*, and *Tnr*, which are key constituents of perineuronal nets [48].

Overall, the PCA structure suggests that plasticity-related gene expression is organized into several partially independent molecular programs rather than forming a single continuum of “plastic” versus “non-plastic” neurons.

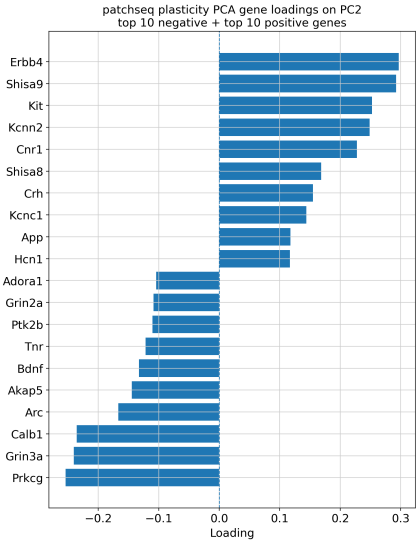
The main gene families contributing to the principal components are summarized in Table 4.11.



(a) Plasticity PCA variance ratio



(b) Top genes contributing to PC1



(c) Top genes contributing to PC2

Figure 4.14: Structure of the plasticity-related transcriptomic space. (A) Variance explained by the leading principal components of the plasticity-gene expression matrix. (B) Top positive and negative gene loadings for PC1. (C) Top positive and negative gene loadings for PC2. The gradual decay of explained variance indicates that plasticity-related transcriptional structure is distributed across many partially independent molecular axes.

Gene family / group	**Representative genes**	**Functional role**
Activity-dependent transcription	*Egr1*, *Arc*	Immediate early genes regulating activity-dependent plasticity and transcriptional responses
Postsynaptic signaling and scaffolding	*Dlg4*, *Syngap1*, *Camk2a/b*	Postsynaptic density organization and synaptic signal transduction
Glutamatergic receptors	*Grin*, *Gria*, *Grik*	NMDA, AMPA, and kainate receptor subunits mediating excitatory synaptic transmission
Neuromodulatory receptors	*Drd1*, *Htr7*, *Adrb1*	Dopamine, serotonin, and adrenergic modulation of cortical circuits
Extracellular matrix / PNN	*Acan*, *Bcan*, *Hapln1*, *Tnr*	Structural stabilization of synapses and regulation of critical-period plasticity [48]
Intrinsic excitability and plasticity signaling	*Kcnc1*, *Scn1a*, *Hcn1*, *Kcnq2*, *Cacna1c*, *Nrg1*, *Erbb4*, *Bdnf*, *Ntrk2*, *Mtor*	Ion channels and signaling pathways controlling interneuron firing properties and synaptic plasticity

Table 4.11: Representative gene families highlighted by PCA.

4.5.2 Plasticity gene expression reveals multiple transcriptional plasticity states

To identify recurring combinations of plasticity-related molecular programs, we applied Leiden community detection, a graph-based clustering algorithm that identifies densely connected groups of similar observations, to the plasticity gene-expression embedding [60]. This analysis identified nine transcriptomic clusters representing distinct plasticity-related transcriptional states, defined as characteristic patterns of coordinated gene expression across cells (Fig. 4.17).

Cluster sizes ranged from 14 to 727 neurons, indicating substantial heterogeneity in plasticity-related gene expression across the dataset. Examination of cluster-averaged gene expression revealed that individual clusters were enriched for different subsets of plasticity-related genes (Fig. 4.15). For example, some clusters were enriched for glutamatergic receptor genes (*Grik1*, *Grid2*, *Gria2*), suggesting specialization in excitatory synaptic signaling, whereas other clusters showed elevated expression of structural remodeling genes such as *Cfl1* and *Mmp9*. A separate cluster displayed strong expression of canonical activity-dependent plasticity regulators including *Camk2* isoforms, *Creb1*, *Arc*, and *Bdnf*.

The results suggest that neurons can occupy multiple transcriptional plasticity states, each characterized by different combinations of synaptic, signaling, and structural plasticity pathways.

4.5.3 Relationship between plasticity transcriptional states and electrophysiological excitability regimes

Coarse electrophysiological regimes revealed by minimal ramp clustering (k=2)

The minimal ramp representation selected as most promising configuration in the whole-signal excitability state analysis identified two principal electrophysiological regimes in the Patch-seq dataset (clusters C0 and C1; Fig. 4.9a). The dominant cluster C0 contained the majority of neurons, whereas cluster C1 formed a smaller but clearly separated branch in electrophysiological PCA space, indicating a distinct dynamical regime of ramp-evoked voltage trajectories.

Projection of plasticity clusters onto the electrophysiological embedding revealed that plasticity-defined transcriptomic states were distributed across both excitability regimes (Fig. 4.16). Individual electrophysiological clusters contained neurons from

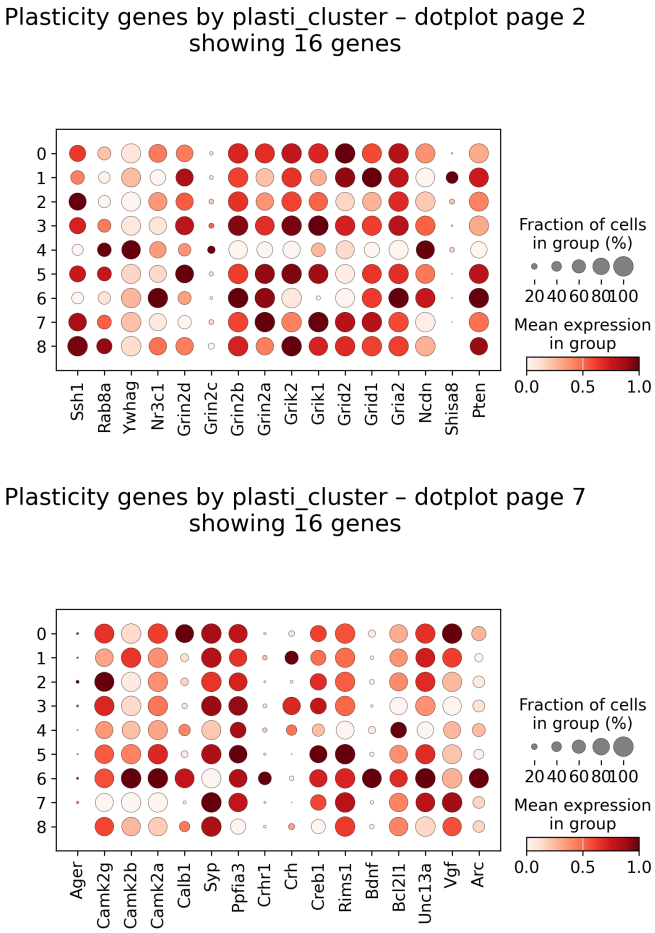


Figure 4.15: Plasticity-related transcriptional programs across Leiden clusters. Dotplots summarize the distribution of selected plasticity-related genes across transcriptomic clusters identified from the plasticity gene-expression embedding. Dot size indicates the fraction of neurons expressing each gene within a cluster, while color intensity represents average normalized expression. Distinct clusters show enrichment for different molecular programs.

multiple plasticity clusters, indicating that transcriptional plasticity programs do not map uniquely onto a specific electrophysiological state.

One plasticity cluster (cluster 2) was strongly enriched in the region of the plasticity embedding corresponding to Pvalb-positive interneurons. Cells in this cluster showed higher expression of several genes involved in activity-dependent synaptic signaling and structural plasticity, including *Camk2g*, *Shisa9*, *Cnr1*, *Ssh1*, and *Mtor*, which are associated with calcium-dependent signaling, AMPA receptor modulation, cytoskeletal remodeling, and translational control of synaptic plasticity, indicating that this cluster likely corresponds to a transcriptional state linked to increased plasticity in fast-spiking interneurons.

Interestingly, neurons belonging to the Pvalb subclass were also distributed across both electrophysiological clusters in the ramp-derived embedding, appearing on both sides of the boundary between the two electrophysiological regimes. This indicates that electrophysiological ramp-response states do not map directly onto classical interneuron subclasses. Instead, neurons of the same molecular subclass can occupy multiple electrophysiological regimes, supporting the view that electrophysiological dynamics and transcriptional plasticity programs represent partially independent axes of neuronal organization.

Conversely, projection of electrophysiological cluster labels onto the plasticity-gene embedding showed that neurons from both electrophysiological regimes were intermingled across multiple transcriptomic clusters (Fig. 4.17). These observations suggest that electrophysiological excitability states and transcriptional plasticity programs capture complementary but partially independent dimensions of neuronal organization. In this representation, the C1 electrophysiological regime was preferentially concentrated in the region enriched for Pvalb interneurons, whereas the broader C0 regime spanned multiple transcriptomic subclasses, most prominently Sst- and Vip-associated regions. Given that Pvalb interneurons are classically associated with fast-spiking, high-frequency firing dynamics, while Sst and Vip interneurons typically display more heterogeneous and often more adaptive firing behaviors, this pattern suggests that the C1 state captures a more Pvalb-enriched electrophysiological regime, whereas C0 reflects a broader dynamical mode shared across multiple interneuron subclasses. These observations further support the view that electrophysiological states cut across canonical cell-type boundaries rather than mapping one-to-one onto transcriptomic subclasses.

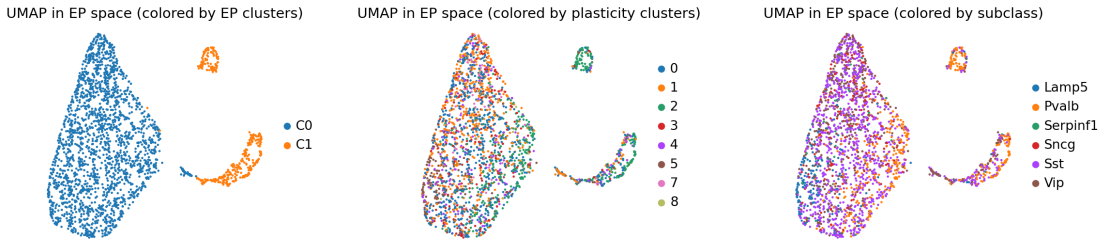


Figure 4.16: UMAP embedding in electrophysiological space. Left: electrophysiological clusters derived from minimal-ramp dynamic features, minimal ramp (k=2). Middle: projection of plasticity-gene Leiden clusters onto electrophysiological space. Right: transcriptomic subclass annotation projected onto the same space. Plasticity-defined molecular states are distributed across multiple electrophysiological regimes.

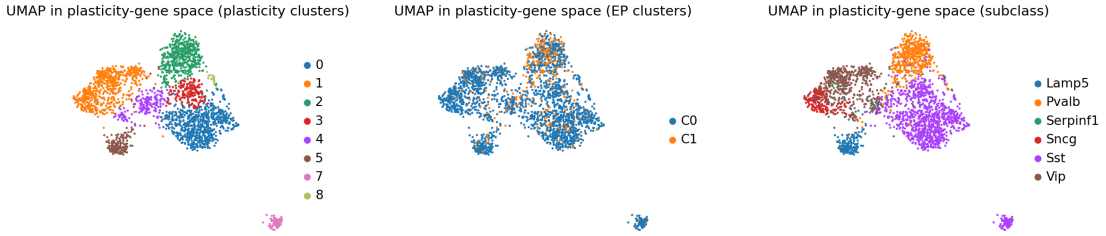


Figure 4.17: UMAP embedding in plasticity-gene expression space, minimal ramp (k=2). Left: Leiden clusters derived from plasticity-gene expression. Middle: projection of electrophysiological cluster labels onto the plasticity embedding. Right: transcriptomic subclass annotation projected onto the same space. Plasticity space shows clear transcriptomic structure while only partially aligning with electrophysiological excitability states.

Finer dynamical organization of ramp responses at higher clustering resolution (k=4)

When electrophysiological clustering was performed at higher resolution (k=4) in the minimal ramp configuration, four distinct dynamical regimes were identified (Fig. 4.12A). The largest cluster (C0) contained approximately 57% of neurons, whereas three smaller clusters (C1, C2, and C3) formed spatially distinct regions in the electrophysiological embedding space.

These clusters correspond to qualitatively different ramp-response dynamics rather than representing a simple continuum of excitability. Importantly, neurons belonging to the same transcriptomic subclass were distributed across multiple electrophysiological regimes. For example, Pvalb neurons were present in electrophysiological states C3, C1, and C0, while Sst and Vip neurons were observed across C0, C2, and C3.

Conversely, individual electrophysiological states could contain mixtures of transcriptomic subclasses. In particular, the dominant state C0 was enriched with neurons from several subclasses, including Pvalb, Sst, and Lamp5.

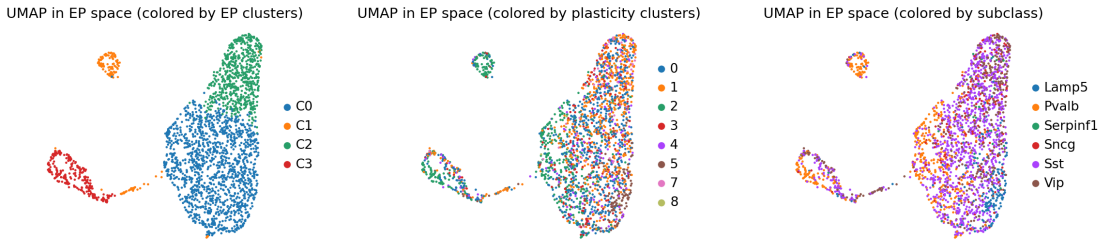


Figure 4.18: UMAP embedding in electrophysiological space, minimal ramp (k=4).

Projection of electrophysiological states onto the plasticity-gene embedding revealed a partially structured relationship between electrophysiological dynamics and transcriptional plasticity programs (Fig. 4.19). One electrophysiological regime (C1) was strongly confined to the upper region of plasticity space corresponding to plasticity cluster 2, which is enriched in Pvalb interneurons. In contrast, other electrophysiological regimes spanned broader regions of plasticity space. Plasticity cluster 5, corresponding primarily to Lamp5 neurons, was largely associated with electrophysiological state C0, whereas plasticity cluster 7, enriched in Sst neurons, mapped preferentially to state C2.

However, the central region of plasticity space comprising clusters 0, 1, 3, and 4 showed substantial overlap between electrophysiological states C0 and C2. These

observations show that while certain electrophysiological regimes can be enriched within specific transcriptomic subclasses, most ramp-response dynamics cut across multiple transcriptional plasticity states rather than mapping one-to-one onto molecular identity.

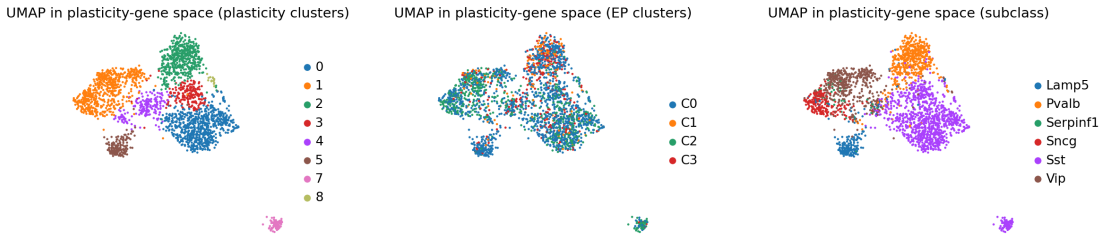


Figure 4.19: UMAP embedding in plasticity-gene expression space, minimal ramp (k=4).

4.5.4 Spatial distribution of plasticity-related genes across electrophysiological space

Visualization of selected plasticity-related genes across the electrophysiological embedding revealed heterogeneous but spatially structured expression patterns. Extracellular matrix genes associated with perineuronal nets (*Acan*, *Bcan*, *Hapln1*, *Tnr*) showed enriched expression in the upper region of electrophysiological space corresponding to *Pvalb* interneurons. This pattern was consistent across both clustering resolutions (k=2 (Fig. 4.20) and k=4 (Fig. 4.21) and is consistent with the well-established association between perineuronal nets and fast-spiking inhibitory interneurons.

In contrast, signaling and ion channel genes involved in intrinsic excitability regulation (*Scn1a*, *Kcnc1*, *Hcn1*, *Kcnq2*, *Cacna1c*) were broadly distributed across the electrophysiological embedding. Although slight enrichment was observed in regions corresponding to fast-spiking neurons, their expression extended across multiple electrophysiological regimes, indicating that molecular mechanisms controlling intrinsic membrane excitability are shared across diverse interneuron populations.

Neuromodulatory and synaptic signaling genes showed more heterogeneous spatial distributions across the electrophysiological embedding. For instance the *Erbb4* gene which is a key component of neuregulin signaling was strongly enriched in the region associated with *Pvalb* interneurons, while *Nrg1* appeared more broadly distributed across the embedding. A similar pattern was observed for *Reln*, whose expression was enriched in areas corresponding to Vip- and Lamp5-associated territories.

Activity-dependent plasticity regulators such as *Bdnf*, together with signaling pathway genes including *Ntrk2* and *Mtor*, showed either sparse or widely distributed expression across electrophysiological space. These observations indicate that while some molecular programs are spatially enriched in specific interneuron subclasses, many plasticity-related genes vary gradually across the electrophysiological embedding rather than being confined to discrete excitability states.

Importantly, increasing clustering resolution from k=2 to k=4 did not substantially alter these gene-expression landscapes. Instead, higher-resolution electrophysiological clustering revealed finer dynamical subdivisions within regions that already exhibited similar molecular expression profiles, indicating that electrophysiological response regimes can partition neurons that share broadly similar transcriptional plasticity programs.

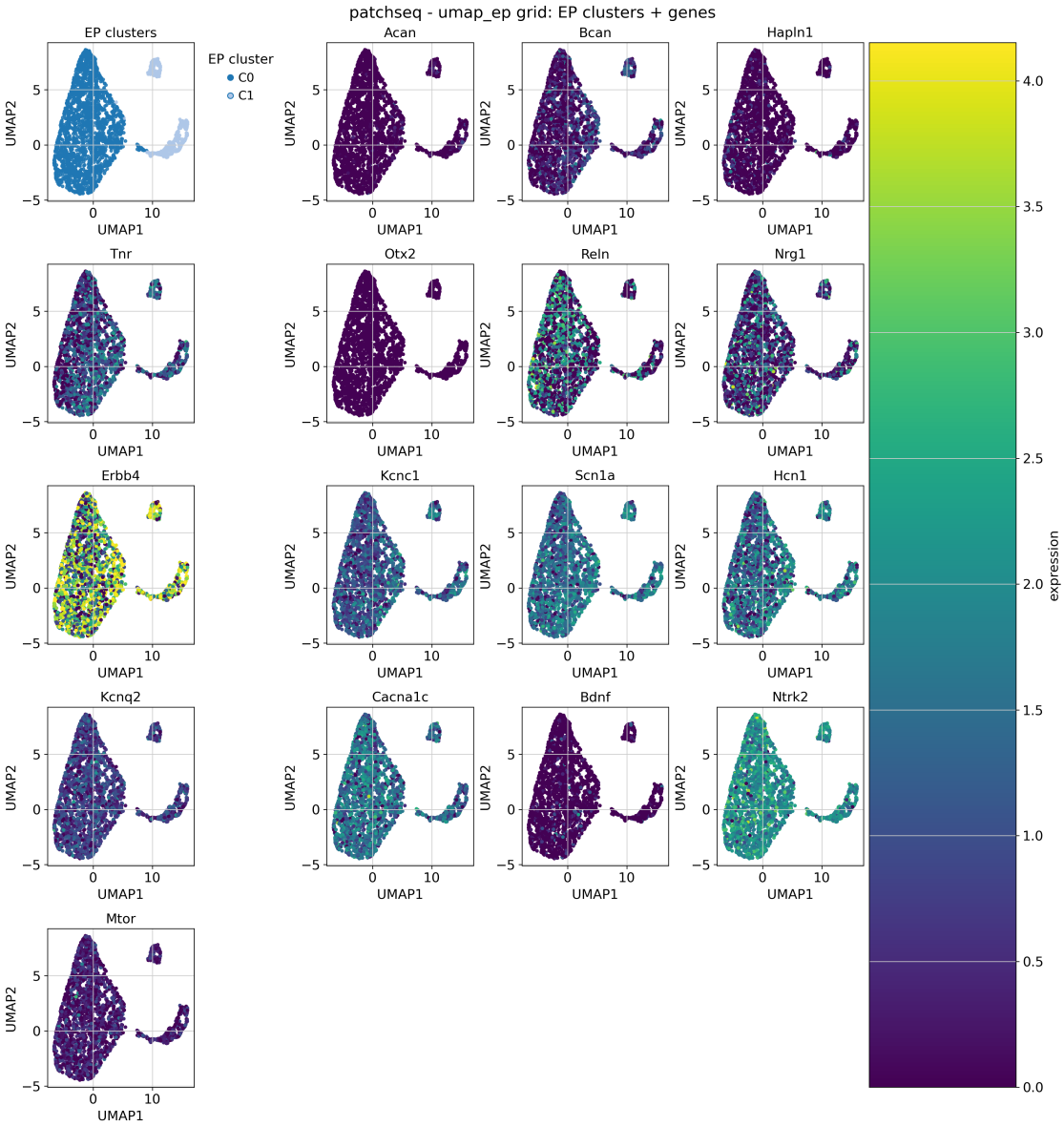


Figure 4.20: Expression of selected plasticity-related genes across the electrophysiological UMAP embedding for the minimal ramp k=2 case. Extracellular matrix genes associated with perineuronal nets (*Acan*, *Bcan*, *Hapln1*, *Tnr*) show enriched expression in restricted regions of the embedding, whereas signaling and ion-channel genes (*Nrg1*, *Erbb4*, *Scn1a*, *Kcnc1*, *Hcn1*, *Kcnq2*, *Cacna1c*) are distributed more broadly across excitability states.

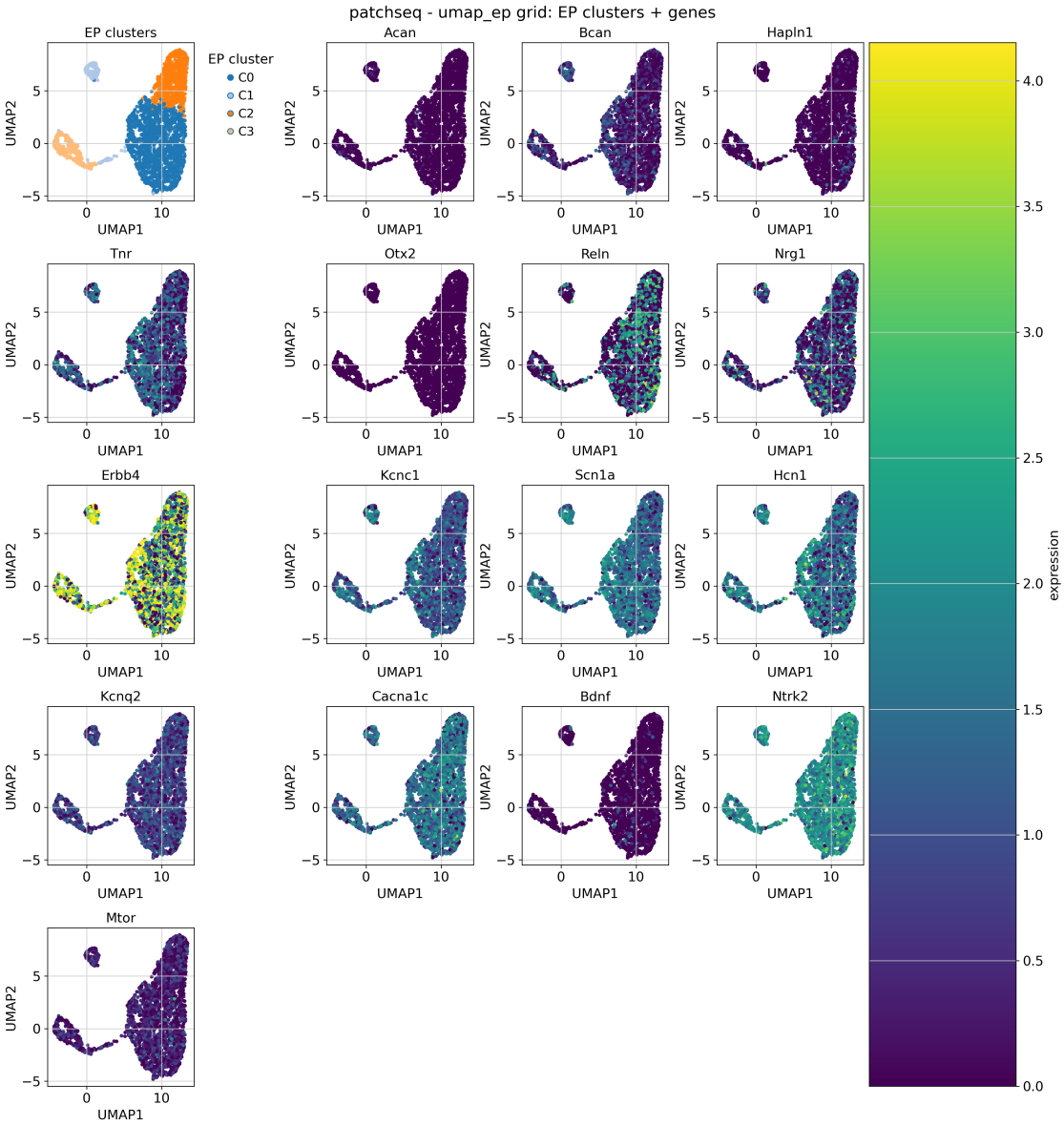


Figure 4.21: Expression of selected plasticity-related genes across the electrophysiological UMAP embedding for the minimal ramp k=4 case

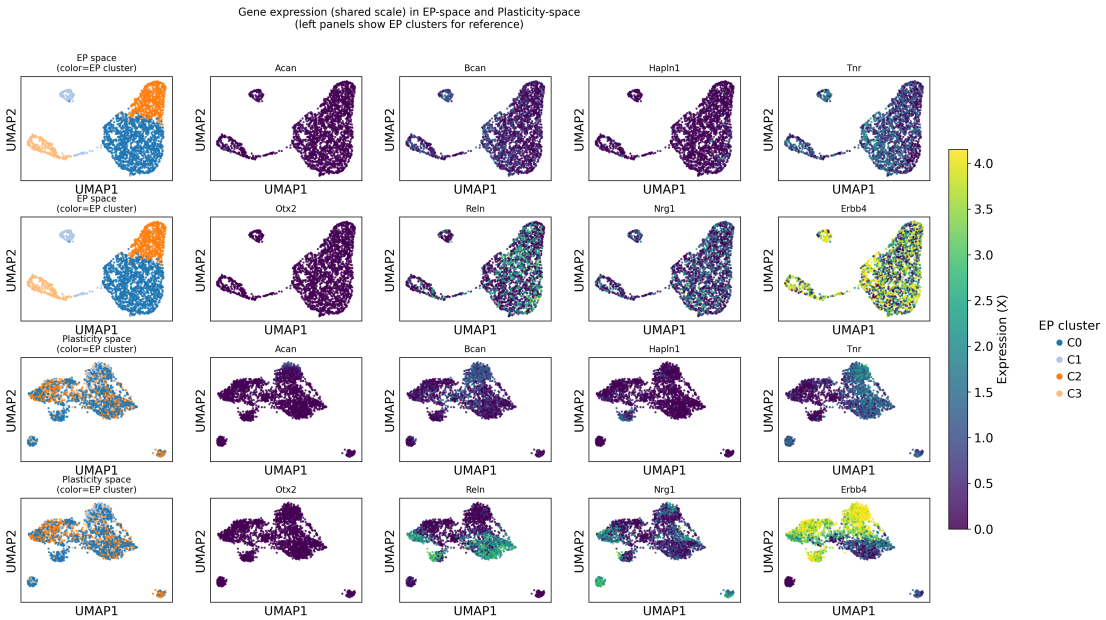


Figure 4.22: Comparison of gene-expression patterns across electrophysiological and plasticity-gene embeddings (minimal ramp, $k = 4$). Each gene is visualized using the same color scale in both spaces. Left panels show the electrophysiological UMAP colored by EP clusters for reference. The upper rows display gene expression projected onto electrophysiological space, while the lower rows show the same genes projected onto plasticity-gene space.

4.5.5 Excitability states and plasticity-related molecular programs

These analyses suggest that electrophysiological excitability regimes and plasticity-related transcriptional programs capture complementary aspects of neuronal organization. While electrophysiological clusters reflect the dynamical response of neurons to depolarizing stimulation, plasticity-related gene expression describes the molecular machinery enabling synaptic and intrinsic plastic modifications.

The partial correspondence between these modalities indicates that neurons with similar molecular plasticity profiles can exhibit different electrophysiological response dynamics, and conversely that similar electrophysiological regimes can arise from heterogeneous molecular backgrounds. These findings support a model in which neuronal identity, intrinsic excitability, and molecular plasticity capacity represent partially independent but interacting dimensions of cortical interneuron organization.

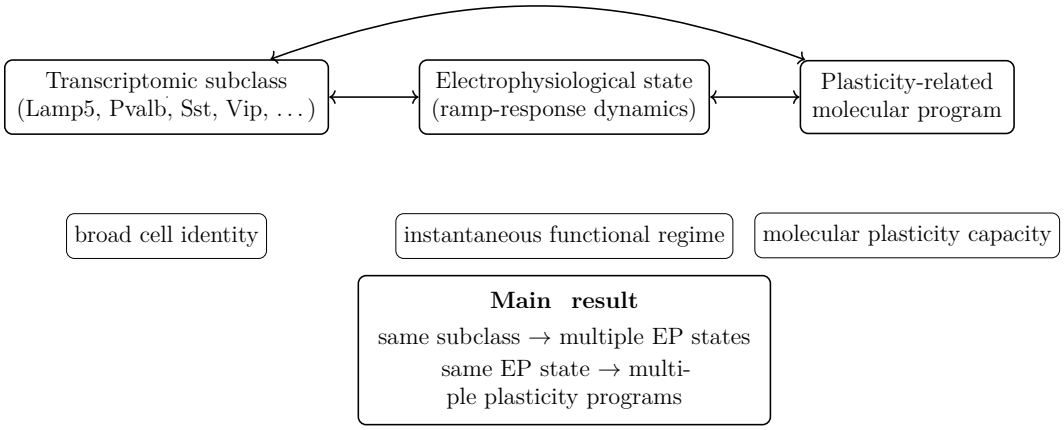


Figure 4.23: Conceptual summary of the three partially independent organizational axes identified in Patch-seq neurons. Transcriptomic subclass captures broad neuronal identity, electrophysiological clustering captures intrinsic excitability regimes, and plasticity-related gene expression captures molecular programs associated with plasticity.

Chapter 5

Conclusions

The main goal of this thesis was to investigate whether intrinsic neuronal excitability can be described through recurrent electrophysiological regimes identified directly from patch-clamp recordings, and whether these regimes relate in a systematic way to transcriptomic cell identity and plasticity-related molecular programs. More specifically, the work set out to move beyond the standard view of excitability as a collection of isolated scalar measurements and to test whether a richer, trajectory-based representation of membrane-potential dynamics could reveal additional structure in neuronal responses.

The results support this general perspective. Across both the PatchClamp and Patch-seq datasets, first-spike waveform analysis recovered compact and interpretable electrophysiological organization, while whole-signal analysis revealed additional structure that became visible only when the membrane-potential trace was treated as a dynamical trajectory rather than reduced to a single spike event. Taken together, these findings suggest that intrinsic excitability is not adequately described as a fixed cell-type property alone. Rather, it is better understood as a physiological dimension of neuronal responsiveness whose expression under stimulation gives rise to recurrent electrophysiological regimes. In this thesis, these regimes are interpreted as dynamic functional states of neuronal response.

5.1 Main findings

At the first-spike level, clustering analyses identified robust low-dimensional excitability structure in both datasets (Tables 4.2 and 4.4). In the PatchClamp

dataset, both long-square and ramp stimulation produced compact solutions, with the clearest separation observed for ramp stimulation in the `firstspk` representation (Fig. 4.4). In the Patch-seq dataset, a similar two-cluster organization emerged from ramp-based first-spike features (Fig. 4.6). Across datasets, cluster separation was consistently driven by waveform descriptors related to repolarization and recovery, especially upstroke/downstroke ratio, *Slope_deep*, dV_{deep}, spike width, and AHP-related measures. This result is consistent with the general motivation of the NSS framework, which showed that compact action-potential geometry can recover biologically meaningful excitability structure [17]. At the same time, the present results extend that view by showing that descriptors linked more explicitly to the afterhyperpolarization and post-spike recovery remain among the most informative variables for separating electrophysiological regimes.

This first result is important for two reasons. First, it confirms that a substantial portion of excitability-related variability is already visible in the early geometry of the first evoked spike. Second, it shows that this variability is not captured only by classical onset features, but also by the repolarization and AHP-related quantities that shape recovery from spiking. In this sense, the first-spike analysis supports the idea developed throughout the thesis that post-spike recovery is not a secondary detail, but part of the electrophysiological signature through which intrinsic excitability is expressed.

The whole-signal analysis added a second and conceptually distinct layer of information. Rather than focusing on the first action potential, this analysis treated the full ramp-evoked membrane response as a continuous time series and extracted descriptors from its reconstructed state-space geometry. In the PatchClamp dataset, the `minimal_ramp` representation identified a compact three-cluster organization (Table 4.6, Fig. 4.7), while `poincare_only` and `poincare_plus_derivative` showed that a considerable portion of this structure could already be captured by local variability descriptors and their combination with derivative-space geometry. In the Patch-seq dataset, the same compact `minimal_ramp` representation produced the strongest and most stable solutions across correlation-pruning thresholds, including both a coarse two-regime partition and a finer four-regime subdivision (Table 4.8, Figs. 4.9 and 4.12).

The analysis found that the most discriminating variables were measures such as p_{sd1}, *ip_mean*, ip_r, derivative or delay anisotropy, and early-late temporal differences such as Δ late-early *ip_std* and Δ late-early *ip_mean* (Fig. 4.13, Table 4.7). All such features summarize nicely the local short-timescale variability, the geometry of the reconstructed trajectory and the temporal evolution along the ramp. Thus, the whole-signal analysis captured how a neuron's membrane-potential trajectory evolves as depolarization progressively increases under ramp stimulation.

As shown in the comparison between the two analytical frameworks (Section 4, Table 4.10), the first-spike analysis provides a near-threshold snapshot of how intrinsic excitability is expressed at spike onset, strongly tied to waveform geometry and repolarization mechanisms. The whole-signal analysis instead describes broader electrophysiological regimes, reflecting variability, recruitment, anisotropy, and adaptation-sensitive evolution over the course of the stimulus. For this reason, the two approaches should not be interpreted as competing alternatives. Rather, they describe complementary levels of the same physiological phenomenon.

5.2 Excitability, transcriptomics and plasticity

A second major aim of the thesis was to determine whether electrophysiological excitability regimes correspond directly to known transcriptomic classes, or whether they cut across molecular boundaries. The results consistently support the second interpretation.

In both first-spike and whole-signal analyses, electrophysiological clusters showed only partial alignment with transcriptomic subclasses. This was already visible in the first-spike results, where the identified regimes were not reducible to subclass composition (Figs. 4.2, 4.4, and 4.6). It became even clearer in the whole-signal Patch-seq analysis, where the same transcriptomic subclass could appear across multiple electrophysiological regimes, and individual electrophysiological regimes contained mixtures of subclasses (Figs. 4.9 and 4.11). This is consistent with previous large-scale studies showing that electrophysiological and transcriptomic organization are related but not strictly one-to-one [11, 12, 22].

The transcriptomic plasticity analysis further strengthened this conclusion. PCA of the curated plasticity-related gene set revealed a high-dimensional expression landscape in which many partially independent molecular programs coexist (Fig. 4.14). The leading components were not dominated by a single "plasticity axis," but instead reflected multiple themes, including activity-dependent transcription, postsynaptic scaffolding, glutamatergic receptor signaling, neuromodulation, and extracellular-matrix-related programs (Table 4.11). Leiden clustering then identified multiple transcriptional plasticity states rather than a simple binary partition into plastic and non-plastic neurons (Fig. 4.15).

When these transcriptomic states were compared with the electrophysiological embeddings, the relationship was again only partial. Plasticity clusters were distributed across multiple electrophysiological regimes, and electrophysiological regimes could be projected onto multiple plasticity-related transcriptional regions

(Figs. 4.16, 4.17, 4.18, and 4.19). Some structured enrichments did emerge. For example, Pvalb-enriched regions were associated with more specific electrophysiological regimes, and extracellular matrix / perineuronal net genes such as *Acan*, *Bcan*, *Hapln1*, and *Tnr* showed restricted spatial enrichment in fast-spiking territories, consistent with prior literature on perineuronal nets and interneuron plasticity [48]. However, many other plasticity-related genes were distributed more broadly across electrophysiological space (Figs. 4.20 and 4.21), suggesting that the molecular capacity for plastic modulation is not confined to a single electrophysiological regime.

These results support a three-axis view of neuronal organization, summarized in Fig. 4.23, where transcriptomic subclass captures broad cell identity, electrophysiological clustering captures recurrent response regimes through which intrinsic excitability is expressed under stimulation and plasticity-related transcription captures molecular programs associated with the capacity for intrinsic and synaptic modification. These axes interact, but they are not reducible to one another.

5.3 Relationship to previous literature

The present findings fit naturally within, but also extend, existing work on excitability-state analysis. At the first-spike level, the results are broadly consistent with the NSS study, which showed that a compact action-potential representation can recover biologically meaningful excitability partitions, including a fast-spiking state supported by electrophysiological and transcriptomic evidence [17]. In the present thesis, this picture is strengthened by the explicit inclusion of slow-trough and AHP-related recovery descriptors, which repeatedly emerged among the most informative features across datasets and protocols. In this sense, the results suggest that the way intrinsic excitability is expressed at the single-spike level is not fully captured by onset and peak geometry alone, but also depends on how the neuron repolarizes and recovers after firing.

The whole-signal results go beyond what has so far been reported in waveform-centered approaches such as NSS. Here, intrinsic excitability is not only studied through the morphology of a single spike, but through the structure of the full ramp-response trajectory. The dominant whole-signal features point toward a view in which local variability, trajectory anisotropy, and early-late evolution during stimulation are central for identifying electrophysiological regimes. To our knowledge, this type of trajectory-based characterization has not yet become standard in patch-clamp analysis, even though it is strongly motivated by the nonlinear dynamical-systems literature [5, 6, 7] and by the idea that neuronal firing

reflects movement through a structured dynamical landscape [19, 26, 27].

The comparison with transcriptomics also extends recent multimodal state-oriented work. In particular, transcriptome-first approaches such as NEUROeSTIMator have shown that neuronal activation can be inferred from gene-expression structure and that these inferred activation signals relate to Patch-seq electrophysiological variables [24]. The present thesis takes the opposite direction: it starts from electrophysiological dynamics and asks how the resulting excitability regimes align with transcriptomic identity and plasticity-related transcription. The answer that emerges is not a one-to-one mapping, but a partial correspondence, with substantial crossing between axes. This is precisely why electrophysiological regimes, transcriptomic subtype, and plasticity-related molecular space need to be treated as complementary rather than interchangeable descriptions.

5.4 Limitations and Future Directions

Some limitations should be taken into account when interpreting the results of this thesis and should be the starting focus of further research stemming from this work.

Regarding the choice of the stimulation waveform, this work's ramp-based analyses represented each neuron by a single selected ramp, defined as the earliest ramp showing a spike response at the lowest current level. This choice gave a practical and uniform criterion, but does not capture the full range of responses available for each neuron. Since each neuron typically has available multiple ramp sweeps, this selected trace may reflect a particularly sensitive moment rather than the complete repertoire by which the cell's intrinsic excitability can be expressed. Plus, this criterion may systematically favor more excitable traces because earlier spiking is itself related to higher excitability. However the same rule was applied consistently across all neurons. A first priority thus would be to move beyond the one-ramp-per-cell representation and incorporate all available ramps for each neuron. This would allow within-cell variability to be quantified directly and would reduce the bias introduced by selecting only the earliest responsive sweep. Doing so, however, is not trivial: the whole-signal embeddings are defined on individual trajectories, and averaging embeddings from markedly different sweeps may not preserve meaningful biology. Future work will therefore need to develop principled ways of aggregating multiple ramp responses at the cell level. Related to this, it will be important to quantify the impact of the current ramp-selection strategy more explicitly. This could be done by comparing alternative criteria, such as the first ramp overall, the first spiking ramp, the most sensitive ramp, the median-response ramp, or

sweep-level aggregate summaries. Another useful analysis would be to measure the similarity between the selected ramp and the remaining ramps of the same neuron, in order to assess whether the chosen sweep is truly representative of the cell's broader excitability landscape.

Another point to keep in mind is that the current ramp protocol may not characterize all neuronal classes equally well. For example neurons with delayed recruitment, including some Pvalb cells, may require stronger or longer depolarization before entering their characteristic firing regime. In this analysis such neurons appeared as having a more difficult or delayed excitability profile, even when this could have possibly been due to the stimulation protocol rather than the actual absence of a fast-spiking regime. Removing these traces however excluded a significant biologically meaningful population, so in the end this analysis deliberately kept them within the inferred state space. Future work should test whether longer or stronger ramps alter their placement in electrophysiological state space. This would help disentangle actual biological low-excitability profiles from stimulation-protocol counter-effects. More realistic stimulation paradigms should also be tested, since standard current-step and ramp protocols remain imperfect estimators of activity under natural synaptic drive [14]. This is especially relevant if the goal is to characterize excitability regimes that generalize beyond highly controlled current injection.

Regarding the electrophysiological analysis, the descriptor set used in the whole-signal analysis could be enriched further. In particular, second-order Poincaré constructions or other higher-order trajectory-based variability measures may help characterize dynamical dependencies that are not visible in the current first-order representations.

Another point is that the transcriptomic integration remains constrained by the known limitations of Patch-seq and single-cell RNA sequencing more broadly. Transcriptomic measurements are noisy, affected by limited sensitivity, and represent a snapshot rather than a continuous molecular history. This necessarily reduces the confidence with which gene-expression structure can be related to electrophysiological regimes. In the present study, normalization and scale harmonization were handled carefully, but the biological interpretation still depends in part on those preprocessing choices.

Moving to the plasticity analysis, it must be noted that it relied on a manually, though literature informed, curated gene set. Thus, it inevitably introduces an element of subjectivity. This affects both the definition of the transcriptomic plasticity space and the interpretation of what counts as a plasticity-related molecular program. One possibility for future research is to use marker-based clustering

pipelines such as `cmCluster` [69], in order to build plasticity signatures using more standardized procedures. Differential-expression analysis focused on neighboring or electrophysiologically ambiguous clusters may also help determine whether some electrophysiological subdivisions are biologically meaningful or instead reflect over-segmentation.

Note that the present analyses infer excitability regimes across neurons, not state transitions within the same neuron over time. In other words, the results support the existence of recurrent dynamical response regimes in the population, but they do not directly demonstrate that individual cells move from one regime to another under experimentally controlled plastic changes. The framework developed here is state-sensitive, but it is not yet longitudinal. Plus, *boundary neurons*, that is neurons that lie near the boundaries between electrophysiological regimes, are difficult to classify with static descriptors alone. Future work should therefore focus more explicitly on identifying and characterizing these cells in order to understand whether they are simply noisy intermediate cases or represent biologically meaningful transitional regimes.

Finally, the study was restricted to mouse visual-cortex data and, for the whole-signal analysis, to ramp stimulation. This gave a clean experimental setting and enabled direct comparison across datasets, but it also limits the generality of the conclusions as highlighted by [14]. Whether the same electrophysiological regimes are preserved across brain regions, cell populations, stimulation paradigms, or species remains an open question. When made available the recently published human Patch-seq electrophysiology dataset [70] will provide an immediate opportunity to ask whether the excitability regimes identified in mouse visual cortex are preserved in human interneurons, and whether the relationship between electrophysiological state and plasticity-related transcription is conserved across these species.

5.5 Final remarks

The central result of this thesis is that intrinsic neuronal excitability can be organized through a set of recurrent electrophysiological regimes extracted directly from patch-clamp recordings, and that these regimes are related to, but not determined by, transcriptomic subtype or plasticity-related gene expression. First-spike analysis captured a compact and highly interpretable near-threshold organization, while whole-signal analysis revealed broader dynamical structure unfolding over the full course of ramp stimulation. Together, these approaches show that intrinsic excitability is not adequately described either by isolated waveform measurements

or by transcriptomic identity alone.

More generally, the work argues for a shift in perspective. Rather than treating excitability as a fixed label attached to a neuron, it is more useful to treat it as a physiological dimension of responsiveness that can be expressed through different electrophysiological regimes under stimulation. In this sense, the excitability landscape developed here is not a replacement for cell-type taxonomy, but a complementary functional layer. It describes how neurons *respond* under stimulation, while transcriptomics describes, more broadly, what they *are* and what molecular machinery they express.

For this reason, the most useful outcome of the thesis is perhaps not a single clustering solution, but the methodological framework itself: a reproducible way to move from raw intracellular recordings to interpretable dynamical regime representations, and from there to a joint comparison with molecular organization. That framework can now be extended, challenged, and refined in future work, but the present results already support the idea that intrinsic excitability is best understood as a dynamic functional dimension whose expression under stimulation cuts across, and interacts with, molecular cell identity.

Bibliography

[1] Bruce P. Bean. «The action potential in mammalian central neurons». In: *Nature Reviews Neuroscience* 8.6 (2007), pp. 451–465 (cit. on pp. 1, 4–6, 8).

[2] M Biel, C Wahl-Schott, S Michalakis, and X Zong. «Hyperpolarization-activated cation channels: from genes to function». In: *Physiological Reviews* 89.3 (2009), pp. 847–885. DOI: 10.1152/physrev.00029.2008 (cit. on pp. 1, 4).

[3] A L. Hodgkin and A F. Huxley. «A quantitative description of membrane current and its application to conduction and excitation in nerve». In: *The Journal of Physiology* 117.4 (1952), pp. 500–544 (cit. on pp. 2, 9).

[4] C Morris and H Lecar. «Voltage oscillations in the barnacle giant muscle fiber». In: *Biophysical Journal* 35.1 (1981), pp. 193–213 (cit. on pp. 2, 9).

[5] F Takens. «Detecting Strange Attractors in Turbulence». In: *Dynamical Systems and Turbulence*. Ed. by D A. Rand and L-S Young. Vol. 898. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer, 1981, pp. 366–381. DOI: 10.1007/BFb0091924 (cit. on pp. 2, 10, 32, 36, 38, 40, 97).

[6] H Kantz and T Schreiber. *Nonlinear Time Series Analysis*. Vol. 7. Cambridge Nonlinear Science Series. Cambridge, UK: Cambridge University Press, 2004. DOI: 10.1017/CBO9780511755798 (cit. on pp. 2, 10, 20, 32, 36, 38, 97).

[7] H.D.I. Abarbanel. *Analysis of Observed Chaotic Data*. New York: Springer-Verlag, 1996. DOI: 10.1007/978-1-4612-0763-4 (cit. on pp. 2, 10, 32, 36, 97).

[8] Dominique Debanne, Yann Inglebert, and Marion Russier. «Plasticity of intrinsic neuronal excitability». In: *Nature Reviews Neuroscience* 20.1 (2019), pp. 51–62. DOI: 10.1038/s41583-018-0118-5 (cit. on pp. 2, 4, 8, 9, 23).

[9] C Hansel and R Yuste. «Neural ensembles: role of intrinsic excitability and its plasticity». In: *Frontiers in Cellular Neuroscience* 18 (2024). ISSN: 1662-5102. DOI: 10.3389/fncel.2024.1440588 (cit. on p. 2).

[10] Gina Turrigiano. «Homeostatic synaptic plasticity: local and global mechanisms for stabilizing neuronal function». In: *Cold Spring Harbor Perspectives in Biology* 4.1 (2012), a005736. DOI: 10.1101/cshperspect.a005736 (cit. on pp. 2, 4).

[11] Shreejoy J. et al. Tripathy. «Transcriptomic correlates of neuron electrophysiological diversity». In: *Neuron* 94.5 (2017), pp. 1–15. DOI: 10.1016/j.neuron.2017.05.007 (cit. on pp. 2, 6–8, 96).

[12] C. Bomkamp, Shreejoy J. Tripathy, et al. «Transcriptomic correlates of electrophysiological and morphological diversity within and across neuron types». In: *PLOS Computational Biology* 15.8 (2019), e1007113. DOI: 10.1371/journal.pcbi.1007113 (cit. on pp. 2, 6–8, 96).

[13] Eve Marder and Jean-Marc Goaillard. «Variability, compensation and homeostasis in neuron and network function». In: *Nature Reviews Neuroscience* 7 (2006), pp. 563–574. DOI: 10.1038/nrn1949 (cit. on p. 4).

[14] Á Szabó, K Schlett, and A Szücs. «Conventional measures of intrinsic excitability are poor estimators of neuronal activity under realistic synaptic inputs». In: *PLoS Computational Biology* 17.9 (2021), e1009378. DOI: 10.1371/journal.pcbi.1009378 (cit. on pp. 4, 11, 62, 99, 100).

[15] S. J. Moore et al. «Of mice and intrinsic excitability: genetic background affects the size of the postburst afterhyperpolarization in CA1 pyramidal neurons». In: *Journal of Neurophysiology* 106.1 (2011), pp. 157–168. DOI: 10.1152/jn.00257.2011 (cit. on pp. 5, 9).

[16] David B. Jaffe and R. Brenner. «A computational model for how the fast afterhyperpolarization paradoxically increases gain in regularly firing neurons». In: *Journal of Neurophysiology* 119.4 (2018), pp. 1506–1520. DOI: 10.1152/jn.00385.2017 (cit. on p. 5).

[17] L Martini, G Amprimo, S Di Carlo, G Olmo, C Ferraris, A Savino, and R Bardini. «Neuronal Spike Shapes (NSS): A straightforward approach to investigate heterogeneity in neuronal excitability states». In: *Computers in Biology and Medicine* 168 (2024), p. 107783. DOI: 10.1016/j.compbiomed.2023.107783 (cit. on pp. 5, 8, 95, 97).

[18] Bert Sakmann and Erwin Neher. *Single-Channel Recording*. 2nd. Springer, 1995 (cit. on p. 5).

[19] Eugene M. Izhikevich. «Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting». In: *MIT Press* (2007) (cit. on pp. 6, 9, 98).

[20] D. B. Stockton and F. Santamaria. «Integrating the Allen Brain Institute Cell Types Database into Automated Neuroscience Workflow». In: *Neuroinformatics* 15.4 (2017), pp. 333–342. DOI: 10.1007/s12021-017-9337-x (cit. on p. 6).

[21] Nathan W. Gouwens, Staci A. Sorensen, Jeremy Berg, et al. «Classification of electrophysiological and morphological neuron types in the mouse visual cortex». In: *Nature Neuroscience* 22 (2019), pp. 1182–1195 (cit. on pp. 6, 8, 13, 53).

[22] Nathan W. Gouwens and Sorensen et al. «Integrated morphoelectric and transcriptomic classification of cortical GABAergic cells». In: *Cell* 183.4 (2020), pp. 935–953. DOI: 10.1016/j.cell.2020.09.057 (cit. on pp. 7, 13, 47, 53, 96).

[23] C. R. Cadwell et al. «Multimodal profiling of single-cell morphology, electrophysiology, and gene expression using Patch-seq». In: *Nature Biotechnology* 34 (2016), pp. 199–203. DOI: 10.1038/nbt.3445 (cit. on p. 7).

[24] Ethan Bahl, Snehajyoti Chatterjee, et al. «Using deep learning to quantify neuronal activation from single-cell and spatial transcriptomic data». In: *Nature Communications* 14 (2023), p. 8234. DOI: 10.1038/s41467-023-44503-5 (cit. on pp. 7, 98).

[25] Corinne Teeter, Ramakrishnan Iyer, Vilas Menon, and et al. «Generalized leaky integrate-and-fire models classify multiple neuron types». In: *Cell Reports* 23.12 (2018), pp. 1–14. DOI: 10.1016/j.celrep.2018.04.021 (cit. on pp. 8, 14, 17, 22).

[26] D. Golomb, J. R. Donner, B. Shacham-Diamand, A. Shlosberg, Y. Amitai, and I. Segev. «Mechanisms of Firing Patterns in Fast-Spiking Cortical Interneurons». In: *PLoS Computational Biology* 3.8 (2007), e156. DOI: 10.1371/journal.pcbi.0030156 (cit. on pp. 9, 98).

[27] Shimon Marom and Eve Marder. «A biophysical perspective on the resilience of neuronal excitability across timescales». In: *Nature Reviews Neuroscience* 24.10 (2023), pp. 640–652. DOI: 10.1038/s41583-023-00730-9 (cit. on pp. 9, 98).

[28] Michael Brennan, Marimuthu Palaniswami, and Peter Kamen. «Poincaré plot interpretation using a physiological model of HRV based on a network of oscillators». In: *American Journal of Physiology-Heart and Circulatory Physiology* 283.5 (2002), H1873–H1886. DOI: 10.1152/ajpheart.00405.2000 (cit. on p. 10).

[29] Che-Hao Hsu and Ming-Ya et al. Tsai. «Poincaré plot indexes of heart rate variability detect dynamic autonomic modulation during general anesthesia induction». In: *Acta Anaesthesiologica Taiwanica* 50.1 (2012), pp. 12–18. DOI: 10.1016/j.aat.2012.02.002 (cit. on p. 10).

[30] M. P. Tulppo, T. H. Mäkikallio, T E. S. Takala, T Seppänen, and H V. Huikuri. «Quantitative beat-to-beat analysis of heart rate dynamics during exercise». In: *American Journal of Physiology* 271 (1996), H244–H252 (cit. on p. 10).

[31] M. Brennan et al. *Beyond traditional Poincaré Analysis: Second Order Plots Reveal Respiratory Effects in HRV*. Manuscript / Preprint. Preprint. 2025 (cit. on p. 10).

[32] F. Amato, V. Cesarini, L. Pietrosanti, G. Costantini, G. Olmo, and G. Saggio. «Hallmarks of Parkinson's disease progression determined by temporal evolution of speech attractors in the reconstructed phase-space». In: *Proc. IEEE International Workshop on Metrology for Industry 4.0 & IoT*. 2023. DOI: 10.1109/MetroInd4.0IoT57462.2023.10180199 (cit. on p. 11).

[33] J.L. Teeters et al. «Neurodata Without Borders: Creating a Common Data Format for Neurophysiology». In: *Neuron* 88.4 (2015), pp. 629–634. DOI: 10.1016/j.neuron.2015.10.025 (cit. on p. 14).

[34] Jan Benda and Andreas V. M. Herz. «A universal model for spike-frequency adaptation». In: *Neural Computation* 15.11 (2003), pp. 2523–2564. DOI: 10.1162/089976603322385063 (cit. on pp. 23, 36, 39, 41).

[35] I. T. Jolliffe and J. Cadima. «Principal component analysis: a review and recent developments». In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2065 (2016), p. 20150202. DOI: 10.1098/rsta.2015.0202 (cit. on p. 29).

[36] P. J. Rousseeuw. «Silhouettes: a graphical aid to the interpretation and validation of cluster analysis». In: *Journal of Computational and Applied Mathematics* 20 (1987), pp. 53–65 (cit. on p. 29).

[37] T. Calinski and J. Harabasz. «A dendrite method for cluster analysis». In: *Communications in Statistics* 3.1 (1974), pp. 1–27 (cit. on p. 29).

[38] David L. Davies and D. W. Bouldin. «A cluster separation measure». In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1.2 (1979), pp. 224–227 (cit. on p. 30).

[39] Steven H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. 2nd ed. Boca Raton, FL: CRC Press, 2015. DOI: 10.1201/9780429492563 (cit. on p. 31).

[40] A. Savitzky and M. J. E. Golay. «Smoothing and Differentiation of Data by Simplified Least Squares Procedures». In: *Analytical Chemistry* 36 (1964), pp. 1627–1639 (cit. on p. 32).

[41] Herbert Edelsbrunner, David Kirkpatrick, and Raimund Seidel. «On the Shape of a Set of Points in the Plane». In: *IEEE Transactions on Information Theory* 29.4 (1983), pp. 551–559. DOI: 10.1109/TIT.1983.1056714 (cit. on p. 34).

[42] Rony Azouz and Charles M. Gray. «Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons». In: *Proceedings of the National Academy of Sciences* 97.14 (2000), pp. 8110–8115. DOI: 10.1073/pnas.130200797 (cit. on pp. 36, 39, 41).

[43] M. Rabinovich, P. Varona, and A. Selverston. «Dynamical principles in neuroscience». In: *Reviews of Modern Physics* 78 (2006), pp. 1213–1265 (cit. on p. 38).

[44] M. Brennan, M. Palaniswami, and P. Kamen. «Do existing measures of Poincaré plot geometry reflect nonlinear features of heart rate variability?» In: *IEEE Transactions on Biomedical Engineering* 48.11 (2001), pp. 1342–1347. DOI: 10.1109/10.959330 (cit. on pp. 38, 40).

[45] Jack J. Jiang, Yu Zhang, and Carol McGilligan. «Chaos in voice, from modeling to measurement». In: *Journal of Voice* 20.1 (2006), pp. 2–17. DOI: 10.1016/j.jvoice.2005.01.001 (cit. on pp. 39, 41).

[46] B. Wang, D. Liu, X. Gao, and Y. Luo. «Three-Dimensional Poincaré Plot Analysis for Heart Rate Variability». In: *Complexity* 2022 (2022), p. 3880047. DOI: 10.1155/2022/3880047 (cit. on pp. 39, 41).

[47] T. Pizzorusso, P. Medini, N. Berardi, S. Chierzi, J. W. Fawcett, and L. Maffei. «Reactivation of ocular dominance plasticity in the adult visual cortex». In: *Science* 298.5596 (2002), pp. 1248–1251. DOI: 10.1126/science.1072699 (cit. on pp. 49, 80).

[48] B. A. Sorg, S. Berretta, J. M. Blacktop, J. W. Fawcett, H. Kitagawa, J. C. F. Kwok, and M. Miquel. «Casting a wide net: Role of perineuronal nets in neural plasticity». In: *Journal of Neuroscience* 36.45 (2016), pp. 11459–11468 (cit. on pp. 49, 80, 82, 97).

[49] S. Sugiyama, A. A. Di Nardo, S. Aizawa, I. Matsuo, M. Volovitch, A. Prochiantz, and T. K. Hensch. «Experience-dependent transfer of OTX2 homeoprotein into the visual cortex activates postnatal plasticity». In: *Cell* 134.3 (2008), pp. 508–520 (cit. on pp. 49, 80).

[50] E. J. Weeber, U. Beffert, C. Jones, J. M. Christian, E. Förster, J. D. Sweatt, and J. Herz. «Reelin and ApoE receptors cooperate to enhance hippocampal synaptic plasticity and learning». In: *Journal of Neuroscience* 22.23 (2002), pp. 9995–10003 (cit. on pp. 49, 80).

[51] L. Mei and K.-A. Nave. «Neuregulin-ERBB signaling in the nervous system and neuropsychiatric diseases». In: *Nature Reviews Neuroscience* 15.6 (2014), pp. 365–378 (cit. on pp. 49, 80).

[52] P. Fazzari, A. V. Paternain, M. Valiente, R. Pla, R. Luján, K. Lloyd, J. Lerma, O. Marín, and B. Rico. «Control of cortical GABA circuitry development by ErbB4». In: *Neuron* 67.1 (2010), pp. 1–15 (cit. on pp. 49, 80).

[53] B. Rudy, G. Fishell, S. Lee, and J. Hjerling-Leffler. «Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons». In: *Developmental Neurobiology* 71.1 (2011), pp. 45–61 (cit. on pp. 49, 80).

[54] I. Ogiwara et al. «Nav1.1 localizes to axons of parvalbumin-positive interneurons». In: *Journal of Neuroscience* 27.22 (2007), pp. 5903–5914 (cit. on pp. 49, 80).

[55] R. E. Dolmetsch, U. Pajvani, K. Fife, J. Spotts, and M. E. Greenberg. «Signaling to the nucleus by an L-type calcium channel-calmodulin complex». In: *Science* 294.5541 (2001), pp. 333–339 (cit. on pp. 49, 80).

[56] Z. J. Huang, A. Kirkwood, T. Pizzorusso, V. Porciatti, B. Morales, M. F. Bear, L. Maffei, and S. Tonegawa. «BDNF regulates inhibitory synaptic transmission in the developing visual cortex». In: *Cell* 98.6 (1999), pp. 739–755 (cit. on pp. 49, 80).

[57] J. O. Lipton and M. Sahin. «The neurology of mTOR». In: *Neuron* 84.2 (2014), pp. 275–291 (cit. on pp. 49, 80).

[58] D. M. Kullmann, A. W. Moreau, Y. Bakiri, and E. Nicholson. «Plasticity of inhibition». In: *Neuron* 75.6 (2012), pp. 951–962. DOI: 10.1016/j.neuron.2012.07.030 (cit. on pp. 49, 80).

[59] D. A. Brown and G. M. Passmore. «Neural KCNQ (Kv7) channels». In: *British Journal of Pharmacology* 156 (2009), pp. 1185–1195 (cit. on pp. 49, 80).

[60] Vincent A. Traag, Ludo Waltman, and Nees Jan van Eck. «From Louvain to Leiden: guaranteeing well-connected communities». In: *Scientific Reports* 9 (2019), p. 5233. DOI: 10.1038/s41598-019-41695-z (cit. on pp. 50, 83).

[61] Vincent D. Blondel, Jean-Loup Guillaume, and Renaud Lambiotte. «Fast unfolding of communities in large networks». In: *Journal of Statistical Mechanics: Theory and Experiment* 2008.10 (2008), P10008. DOI: 10.1088/1742-5468/2008/10/P10008 (cit. on p. 50).

[62] Leland McInnes, John Healy, and James Melville. «UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction». In: *arXiv preprint arXiv:1802.03426* (2018) (cit. on p. 50).

[63] Lawrence Hubert and Phipps Arabie. «Comparing partitions». In: *Journal of Classification* 2 (1985), pp. 193–218 (cit. on p. 51).

[64] Alexander Strehl and Joydeep Ghosh. «Cluster ensembles - a knowledge reuse framework for combining multiple partitions». In: *Journal of Machine Learning Research* 3 (2002), pp. 583–617 (cit. on p. 51).

[65] Malte D. Luecken and Fabian J. Theis. «Current best practices in single-cell RNA-seq analysis: a tutorial». In: *Molecular Systems Biology* 15.6 (2019), e8746. DOI: 10.15252/msb.20188746 (cit. on p. 52).

[66] Valentine Svensson, K. N. Natarajan, Lam-Han Ly, Ricardo J. Miragaia, C. Labalette, Iain C. Macaulay, Ana Cvejic, and Sarah A. Teichmann. «Power analysis of single-cell RNA-sequencing experiments». In: *Nature Methods* 14 (2017), pp. 381–387 (cit. on p. 52).

[67] Steven W. Flavell and Michael E. Greenberg. «Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system». In: *Annual Review of Neuroscience* 31 (2008), pp. 563–590. DOI: 10.1146/annurev.neuro.31.060407.125631 (cit. on p. 52).

[68] Anne E. West and Michael E. Greenberg. «Neuronal activity–regulated gene transcription in synapse development and cognitive function». In: *Cold Spring Harbor Perspectives in Biology* 3.6 (2011), a005744. DOI: 10.1101/cshperspect.a005744 (cit. on p. 52).

[69] Y. Huang, H. Chang, X. Chen, J. Meng, M. Han, T. Huang, L. Yuan, and G. Zhang. «A cell marker-based clustering strategy (cmCluster) for precise cell type identification of scRNA-seq data». In: *Quantitative Biology* 11.2 (2023), pp. 163–174. DOI: 10.15302/J-QB-022-0311 (cit. on p. 100).

[70] B. Lee, R. Dalley, T. Chartrand, B. Kalmbach, and E. Lein. *Human interneuron Patch-seq electrophysiology*. Version 0.241120.0510. DANDI Archive, 2024. DOI: 10.48324/dandi.000636/0.241120.0510. URL: https://doi.org/10.48324/dandi.000636/0.241120.0510 (cit. on p. 100).