Machine Learning Classification of Purkinje Cell Spike Features for Electrophysiological Characterization and Potential ASD Relevance

Zhihao Cheng

ZJUI, Zhejiang University, Zhejiang, China

Email: Zhihao.22@intl.zju.edu.cn

Abstract. Purkinje cells, as principal output neurons of the cerebellar cortex, critically influence motor coordination and cognitive functions. Abnormal firing patterns in these cells have been implicated in autism spectrum disorder (ASD), highlighting their potential as biomarkers for neurological dysfunction. In this study, we developed a machine learning-based classification pipeline utilizing spike-derived electrophysiological features to differentiate between experimental conditions in mouse Purkinje cells. Whole-cell patch-clamp recordings from three distinct experimental groups—AA-only, PE-only and Control—were analyzed to extract firing rate, mean inter-spike interval (ISI), ISI variability (standard deviation), and coefficient of variation (CV_ISI). We compared the performance of several machine learning models, including support vector machines (SVM), k-nearest neighbors (KNN), logistic regression, and random forest (RF). Among them, RF achieved the highest accuracy (67%) in 3-fold cross-validation. Feature importance analysis showed that CV_ISI and firing rate were the most predictive features. These findings suggest that spike statistics carry valuable information for ASD-related neural classification and provide a foundation for interpretable, data-driven analysis of cerebellar electrophysiology.

Keywords: Purkinje cell, spike train, autism spectrum disorder, machine learning, random forest.

1. Introduction

The cerebellum plays a central role in motor control, coordination, and cognitive processing, primarily mediated through Purkinje cells—the sole output neurons of the cerebellar cortex. Purkinje cells integrate complex synaptic inputs to produce precisely timed spike outputs, significantly influencing downstream motor behavior and learning processes [1]. Given their precise firing patterns, Purkinje cells have become a key model for studying neuronal excitability and synaptic integration. Abnormalities in Purkinje cell firing have been observed in various neurological disorders, notably autism spectrum disorder (ASD), ataxia, and various cognitive dysfunctions [2]. Thus, understanding the electrophysiological characteristics of Purkinje cells provides valuable insights into both normal cerebellar function and pathological states.

While spike-derived statistical features such as firing rate and inter-spike interval variability have been widely utilized to characterize neuronal activity, systematic evaluation of these features' predictive capabilities across distinct experimental conditions remains relatively underexplored [3, 4]. The advent of machine learning techniques, capable of handling high-dimensional data and revealing non-linear relationships, offers a powerful approach for analyzing electrophysiological recordings. However, few studies have applied these techniques to single-cell patch-clamp data from Purkinje cells.

In this study, we analyze whole-cell patch-clamp recordings of mouse Purkinje cells from a publicly available dataset and evaluate the performance of several supervised classifiers trained on extracted spike features [5]. The goal is to assess whether simple spike statistics can robustly distinguish between experimental groups. Our pipeline may be further extended to disease-related datasets in future applications such as autism spectrum disorder (ASD).

2. Related Work

Spike-based statistical features such as firing rate and the coefficient of variation of inter-spike intervals (CV_ISI) are widely used to characterize neuronal excitability and variability. These features provide a compact, interpretable representation of single-neuron dynamics and have been applied to distinguish various neuronal cell types and firing patterns [3, 5].

In parallel, machine learning (ML) methods such as support vector machines (SVM) and random forests have been increasingly adopted in neural signal processing, particularly for decoding information from spike trains. However, while most studies focus on large-scale or extracellular recordings, the application of ML to single-cell patch-clamp recordings using basic spike statistics remains limited.

Our work bridges this gap by evaluating the discriminative power of spike statistics across experimental conditions in mouse Purkinje cells. The proposed method emphasizes interpretability and provides a lightweight pipeline potentially adaptable to future disease-related datasets.

3. Materials and Methods

3.1 Data Source

The electrophysiological data used in this study were obtained from Dryad, an open-access data platform [6]. The dataset consisted of whole-cell patch-clamp recordings from mouse cerebellar Purkinje cells, originally collected using PatchMaster software and saved in .dat format.

We selected recordings from three groups: AAonly, PEonly, and Control. The AAonly group underwent active avoidance training, while the PEonly group received predicted error stimulation. Although the dataset is not explicitly labeled for autism spectrum disorder (ASD), it remains useful for demonstrating signal processing and machine learning to analyze Purkinje cell signals.

To facilitate data processing, we constructed a custom pipeline to convert .dat files into .csv format. After conversion, each of these .csv files contained membrane voltage recordings sampled at a high rate (usually around 10 kHz or more), enabling accurate spike detection and downstream analyses.

3.2 Identifying Spikes and Extracting Features

After conversion, each .csv file contained a one-dimensional time-series voltage trace representing spontaneous Purkinje cell firing. We initially visualized the first 1000 milliseconds of each trace before automating processing. By verifying the existence of spikes and looking for baseline drift or high-frequency noise, this stage assisted us in evaluating the signal quality. Visual inspection was preferred at this stage over automated thresholding due to its reliability.

Spike detection was implemented in Python using a simple yet effective algorithm: a downward voltage crossing below a predetermined threshold (such as -20 mV) was referred to as a spike. The minimum refractory interval (e.g., 5 ms) between consecutive spikes was applied to prevent noise-induced false detections or double-counting. This approach supported batch processing of multiple cells and performed consistently well across the dataset without the need for sophisticated fitting or filtering.

For each cell, we extracted the following four firing features: Firing rate (Hz), the total amount of all spikes divided by the duration of the recorded signals; Average time between successive spikes, or mean ISI (ms); ISI standard deviation (ms), an indicator of temporal variability; and CV_ISI, the measure of spike timing irregularity, calculated by dividing the standard deviation by the mean.

These characteristics' ease of use and interpretability make them popular in electrophysiology. For classification analysis, each cell was thus labeled by experimental group (AA only, PE only, or control), represented as a four-dimensional feature vector, and added to a combined dataset.

Prior to training, we applied z-score normalization and removed any samples with missing values to ensure numerical stability and fair comparisons across models. The resulting pipeline is fully automated and easily extendable to datasets from disease models, including those relevant to ASD.

3.3 Design of Classification Model

We evaluated several machine learning models to categorize Purkinje cell recordings according to the spike properties that were collected. The classifiers were random forest, logistic regression, k-nearest neighbors (KNN), and support vector machines (SVM) using both linear and RBF kernels. Every model was incorporated into a uniform pipeline that featured stratified 3-fold cross-validation and feature scaling.

We used grid search to optimize each classifier's parameters for fair comparison. To enhance performance consistency, hyperparameters such as the number of estimators in the random forest and the regularization strength CCC in SVM were changed. Additionally, we used standard deviation of accuracy scores to assess model variability between folds.

Following model evaluation, we visualized classifier performance using bar plots with error bars. In order to determine which spike features contributed most to the classification job, we also performed a feature importance analysis using the random forest model.

This design enabled us to compare linear and nonlinear models, assess robustness under sample constraints, and gain insights into physiological signal relevance.

3.4 Evaluation Protocol

The stratified three-cross meditation approach is used to eventuate the classification model's performance since the number of samples is limited for each experimental group. Such a method can ensure that the report preserves the same class distribution as the original data.

The most beneficial aspect of such an approach is that it assesses each model with classification accuracy. Accuracy is easy to understand in multi-class settings and represents the percentage of correctly classified samples. We evaluated the stability and generalizability of each model by computing the standard deviation across folds in addition to the average accuracy.

To ensure fair comparison, we integrated z-score normalization of all features into a pipeline that contained each classifier. In this way, scale bias was avoided in feature magnitude-sensitive models such as KNN and SVM. Feature importance scores from the random forest model further supported interpretability by revealing which spike metrics were most influential in distinguishing experimental groups.

4. Results

4.1 Raw Signal Inspection

Initially, the raw electrophysiological recordings from a subset of Purkinje cells were converted from .dat to .csv format in accordance with the data conversion methodology described in Section 3.1. A representative example of the first 100 milliseconds of membrane voltage data taken from the 040718C_parsed.csv file is shown in Figure 1.

The trace shows typical patterns of spontaneous neural activity, including subthreshold oscillations and distinct downward deflections indicative of spike events. The quality of the recorded signals seems to be constant, with little baseline drift and no major artifacts. This shows that the converted voltage records are reliable and useful for finding spikes and doing quantitative analysis later. Similar visual inspections were conducted across all the chosen cells in the AA-only, PE-only, and control groups to ensure consistent signal quality and suitability for further analysis.

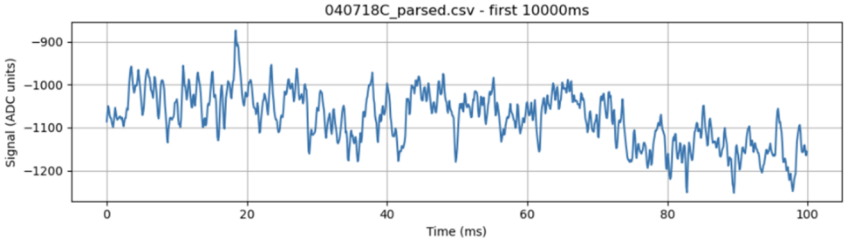

Figure 1: Signal of AA-only cell, for example

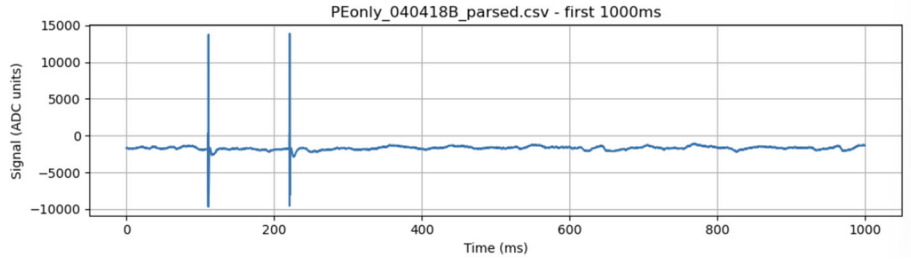

Figure 2: Signal of PE-only cell, for example

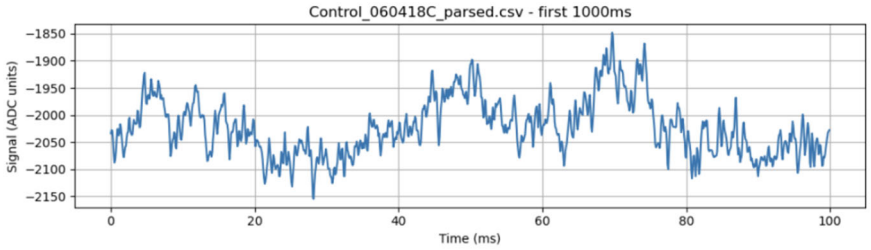

Figure 3: Signal of control cell, for example

4.2 Spike Detection and Feature Extraction

Spikes were automatically identified from the voltage traces using a threshold-based method, in accordance with the steps outlined in Section 3.2. A sample spike detection result from the first 1000 milliseconds of the file 040718C_parsed.csv is displayed in Figure 4. Detected spikes, marked with red crosses, align well with sharp downward voltage deflection, demonstrating the efficacy and resilience of the threshold criteria used.

For each recording, we extracted the following four key statistical features: firing rate, mean inter-spike interval (ISI), standard deviation of ISI, and coefficient of variation of ISI (CV_ISI). These measures gave a clear, easy-to-understand picture of how neurons fire. Cells with insufficient spike activity or prominent artifacts were excluded to maintain data integrity.

This extraction procedure created a consistent feature dataset that could be used for classification analysis. This confirmed that the selected features accurately captured important firing properties that were relevant to the experimental condition.

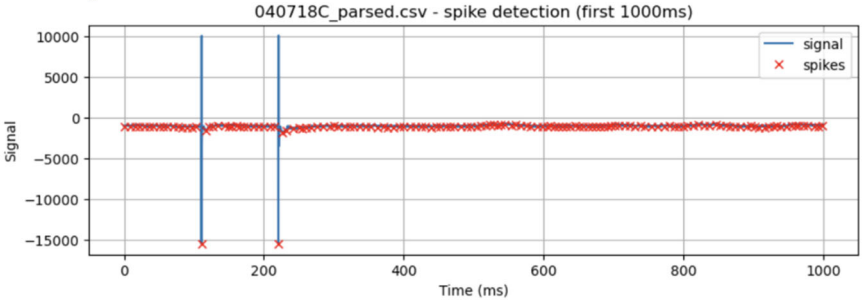

Figure 4: Spike detection of an AA-only cell, for example

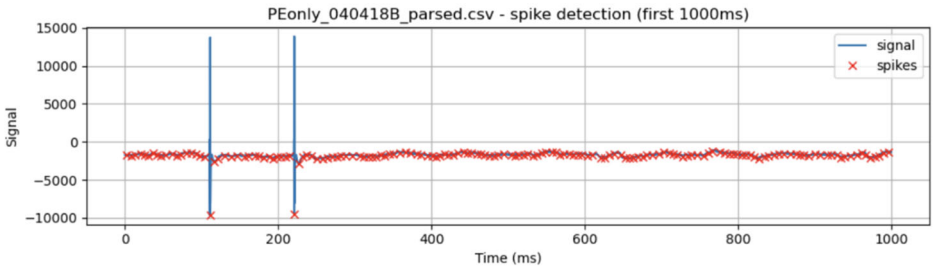

Figure 5: Spike detection of PE-only cell, for example

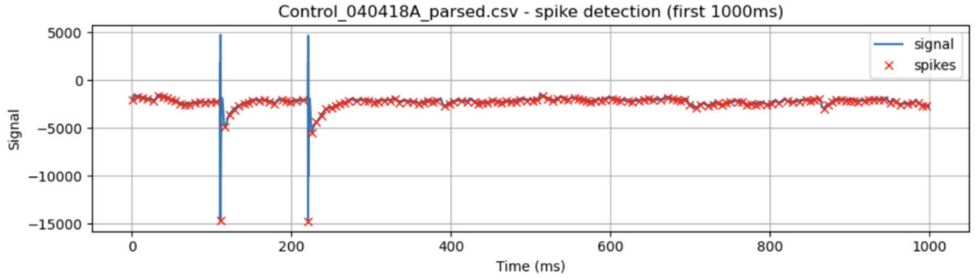

Figure 6: Spike detection of a control cell, for example

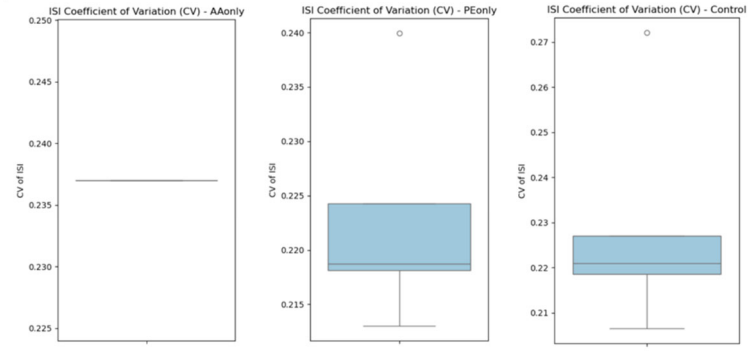

Figure 7: CV_ISI of AA-only, PE-only, and control groups

Table 1: Merged dataset of cells and extracted features

cell	Total spikes	Duration (s)	Firing rate (Hz)	Mean ISI	Std ISI	CV_ISI	group
040718C_parsed.csv	139302	928.9103	149.9628	6.596555	1.563397	0.237002	AAonly
060418B_parsed.csv	119584	793.9103	150.6266	6.596555	1.563397	0.237002	AAonly
070721A_parsed.csv	140162	932.2483	150.3484	6.596555	1.563397	0.237002	AAonly
180618B_parsed.csv	74777	496.9103	150.4839	6.596555	1.563397	0.237002	AAonly
270521B_parsed.csv	134694	888.5239	151.593	6.596555	1.563397	0.237002	AAonly
PEonly_040418B_parsed.csv	1550	9.9999	155.0016	6.450872	1.446737	0.22427	PEonly
PEonly_060418D_parsed.csv	1511	9.9999	151.1015	6.619801	1.447976	0.218734	PEonly
PEonly_060422B_parsed.csv	1518	9.9999	151.8015	6.590903	1.437793	0.218148	PEonly
PEonly_090322A_parsed.csv	1513	9.9999	151.3015	6.607474	1.407305	0.212987	PEonly
PEonly_210618A_parsed.csv	1475	9.9999	147.5015	6.778358	1.626419	0.239943	PEonly
Control_040418A_parsed.csv	1535	9.9999	153.5015	6.515711	1.479356	0.227045	Control
Control_060418A_parsed.csv	1514	9.9999	151.4015	6.605684	1.797441	0.272105	Control
Control_060418C_parsed.csv	1507	9.9999	150.7015	6.63745	1.466561	0.220953	Control
Control_06052021b_parsed.csv	1473	9.9999	147.3015	6.789538	1.40161	0.206437	Control
Control_07102021d_parsed.csv	1480	9.9999	148.0015	6.759635	1.476977	0.218499	Control

4.3 Model Evaluation

In accordance with the classification techniques described in Section 3.3, we evaluated several machine learning models: logistic regression, k-nearest neighbors (KNN, k=3), support vector machines (SVM) with both linear and RBF kernels, and random forest. The classification performance of these models is summed up in Table 2 using the mean accuracy and standard deviation derived from stratified 3-fold cross-validation.

The SVM with an RBF kernel performed consistently and robustly, achieving the highest accuracy of 0.67 ± 0.09. Although more variable, the KNN model (k=3) also did well (0.60 ± 0.16). On the

other hand, linear techniques like linear SVM (0.53 ± 0.09) and logistic regression (0.47 ± 0.19) demonstrated lower accuracies, suggesting that the feature distributions were non-linear. The random forest classifier performed moderately (0.47 ± 0.19), but its performance may have been affected by the small sample size or feature interaction sensitivity. These findings support the superiority of non-linear classifiers, particularly SVM with RBF kernels, for efficiently separating experimental groups based on statistical features derived from spikes.

Table 2: Mean accuracy and standard deviation of different models

	Logistic Regression	KNN (k=3)	SVM (linear)	SVM (rbf)	Random Forest
Mean accuracy	0.47	0.60	0.53	0.67	0.47
STD	0.19	0.16	0.09	0.09	0.19

4.4 Visualization of Model Comparison

The cross-validation results were succinctly depicted using a horizontal bar plot to intuitively convey the relative effectiveness of the classifiers (Figure 8). The random forest classifier demonstrated superior performance, attaining a mean accuracy of 0.67 with minimal variability across folds. The SVM utilizing an RBF kernel exhibited robust predictive capabilities (0.60 accuracy); however, the KNN classifier (k=3) displayed inferior accuracy (0.47) accompanied by more variation, indicating less stable performance.

To further interpret model behavior, feature importance scores from the random forest classifier are illustrated in Figure 9. Among the extracted spike features, mean_ISI, CV_ISI, and std_ISI showed comparatively higher importance, indicating that the interval-based and variability-based measures had a stronger discriminative power than firing rate alone. These results emphasize the value of interpretable, spike-derived features and reinforce the suitability of non-linear classifiers for modeling Purkinje cell activity under different experimental conditions.

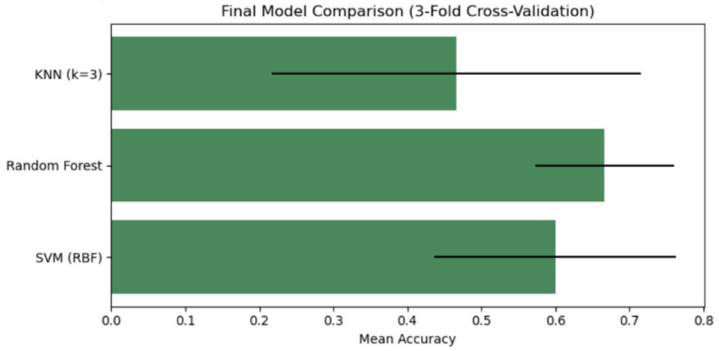

Figure 8: Final Model Comparison

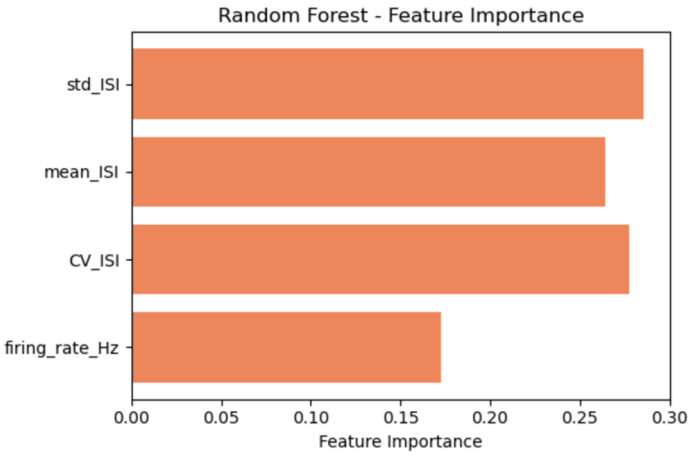

Figure 9: Feature Importance of Random Forest

5. Discussion

The classification results show that simple spike-derived features are an effective way to capture electrophysiological differences between experimental groups. Particularly, models with non-linear decision boundaries, like random forest and SVM with radial basis function (RBF) kernel, achieved the highest accuracy, indicating that there is most likely a non-linear relationship between experimental conditions and spike characteristics.

Among the four extracted feature, CV_ISI and firing rate were found to be the most significant, which is consistent with previous studies showing that firing frequency and interspike interval variability are significant predictors of neuronal excitability and state classification. Advantages of these metrics are well-established physiological interpretations and efficiency, which can ensure the metrics are suitable for clinical and experimental research settings [5].

Our findings support earlier research showing that basic temporal characteristics taken from neural signals can be used to categorize different types of neurons or detect abnormal conditions. However, our analysis was restricted to unlabeled intracellular recordings grouped by experimental manipulation, distinct from those studies that frequently rely on labeled disease models or large-scale population data. Our results should therefore be regarded as an exploratory validation of classification based on signal processing.

A primary limitation of this study is the small dataset size, which may restrict generalizability. Future improvements could involve adaptive spike identification methods or the inclusion of other characteristics such as burst structure, spectral characteristics, or spike amplitude.

Despite these limitations, the constructed pipeline offers a flexible framework for analyzing single-cell recordings. The message of signal processing is not only interpretable but also transferable. Additional modification can facilitate the detection of disorders like ASD, which is related closely to the dysfunction of Purkinje cells [7].

6. Conclusion

In this work, we proposed an easy-to-understand pipeline for the analysis of Purkinje cell electrophysiological recordings that solely uses spike-derived features. By extracting four commonly used spike metrics—firing rate, mean ISI, standard ISI, and CV_ISI—we demonstrated that even basic statistical descriptors can enable meaningful classification of experimental conditions.

Our findings demonstrate the promising accuracy of nonlinear machine learning models, specifically SVM with RBF kernel and random forest, in differentiating between the AA-only, PE-only, and control groups. Additionally, CV_ISI and firing rate were identified by feature importance analysis as significant contributors, highlighting the significance of temporal dynamics in Purkinje cell behavior [8].

The suggested workflow provides a transferable basis for future applications in neural condition screening, including research related to ASD, even though our dataset lacks explicit disease labels. Scalability, efficiency, and adaptability to other datasets involving intracellular or extracellular recordings are features of the signal processing and modeling techniques [9].

To further enhance classification performance and interpretability, future work will incorporate additional time-series features, refine spike detection algorithms, and expand this framework to larger datasets with behavioral or diagnostic annotations.

References

[1] R. V. Sillitoe and A. L. Joyner, “Morphology, molecular codes, and circuitry produce the three-dimensional complexity of the cerebellum,” *Annu. Rev. Cell Dev. Biol.*, vol. 23, pp. 549–577, 2007.

[2] M. T. Tsai, et al., “Autistic-like behavior and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice,” Nature, vol. 488, no. 7413, pp. 647–651, 2012. doi: 10.1038/nature11310.

[3] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.

[4] D. R. Hampson and G. J. Blatt, “Autism spectrum disorders and neuropathology of the cerebellum,” *Front. Neurosci.*, vol. 9, art. 420, 2015. doi: 10.3389/fnins.2015.00420.

[5] M. P. Nawrot et al., “Measurement of variability dynamics in cortical spike trains,” *J. Neurosci. Methods*, vol. 169, no. 2, pp. 374–390, 2008. doi: 10.1016/j.jneumeth.2007.10.013.

[6] M. Watson, A. R. Wilson, and T. Knöpfel, “Electrophysiological recordings from cerebellar Purkinje cells in control and genetically modified mice,” Dryad, Dataset, https://doi.org/10.5061/dryad.zs7h44jk0, 2022.

[7] M. Tsai et al., "Autistic-like behaviour and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice," Nature, vol. 488, pp. 647–651, 2012, doi: 10.1038/nature11310.

[8] J. Mapelli, L. Pagani, F. Garrido, and E. D’Angelo, “Excitatory and inhibitory plasticity in the cerebellar granular layer,” *Front. Cell Neurosci.*, vol. 9, art. 169, May 2015. doi:10.3389/fncel.2015.00169

[9] S.-S. Wang, A. D. Kloth, and A. Badura, “The cerebellum, sensitive periods, and autism,” *Neuron*, vol. 83, pp. 518–532, 2014. doi:10.1016/j.neuron.2014.07.016.