An Automatic Feature Extraction Method for Noisy Electrophysiology Data

Eric Johnson[1] and William L. Kath[1,2]

1. Engineering Sciences and Applied Mathematics; 2. Neurobiology and Physiology

NORTHWESTERN UNIVERSITY

1. Background: Feature Extraction

Automatic feature extraction is increasingly necessary for quantitative analyses of electrophysiology experiments.

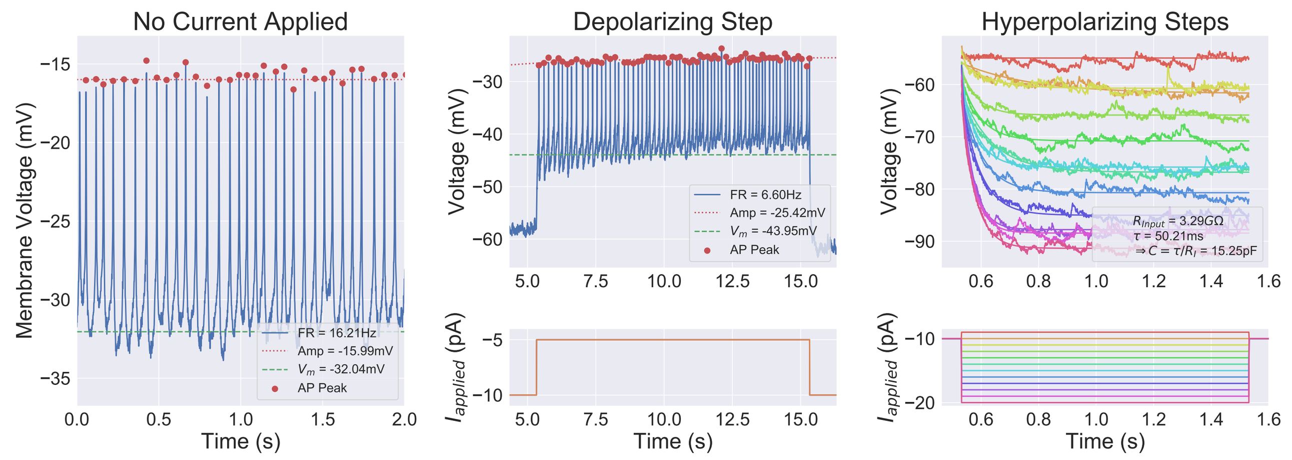
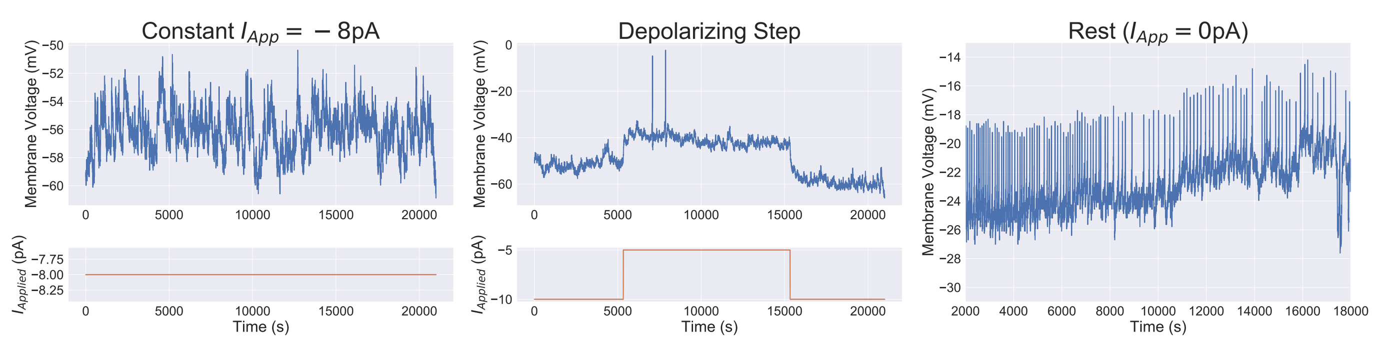


Figure: Examples of electrophysiology recordings of DN1p neurons in *Drosophila* under various stimulus protocols.

- Features such as Firing Rate (FR), Resting Membrane Potential (RMP, V_m), Spike Amplitude, Input Resistance (R_I), Cell Capacitance (C), F-I curves, activation and inactivation time constants, and action potential (AP) shape are variously useful for modeling or characterizing neurons. (Flourakis et al. 2015, Abouzeid et al. 2015)
- As datasets become larger, automation is necessary for processing.
- Automation also creates consistency and can improve accuracy of feature measurements.

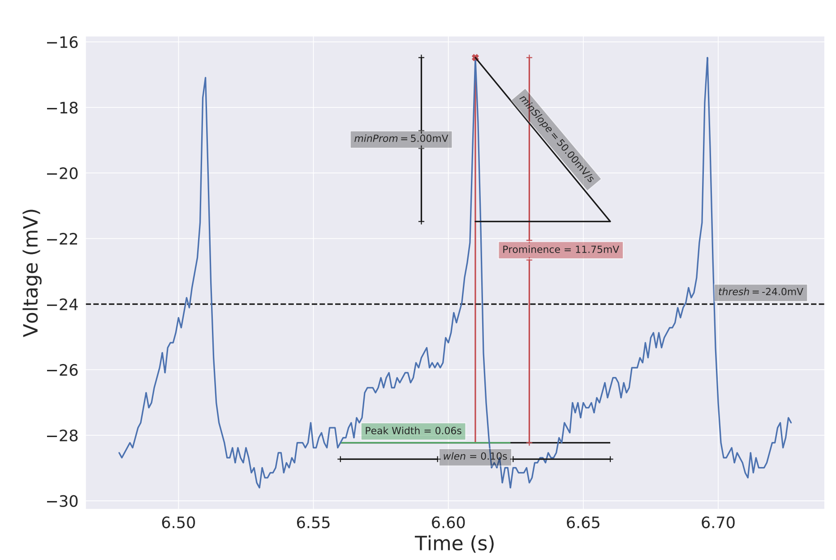
2. DN1p Neurons in *Drosophila melanogaster*

The DN1p neurons are "clock neurons": part of the *Drosophila* brain that controls circadian behaviors (Allada & Chung 2010). The Allada lab undertook to characterize these neurons and their circadian behavior (Flourakis et al. 2015). Even the most up-to-date tools generally rely on simple threshold setting for peak detection (Siegle et al. 2017).



- *Drosophila* neurons are small, making electrophysiological measurements relatively noisy.
- The data set consists of hundreds recordings of various protocols, making a quantitative analysis possible, and automation necessary.
- These neurons exhibit behaviors such as small and stubby spikes, irregular firing, and slow drifting of the RMP, making typical approaches to feature extraction very difficult.
- The data were recorded at a low sampling rate (1kHz), compounding the difficulty of feature extraction.

3. Peak Detection Parameters



- `thresh` is the value a peak must surmount to be counted as a spike.
- `prominence` is the distance between a peak and the (largest) minimum within `wLen` data points.
- `Prominence` and `wLen` together give a minimum slope for a spike.

4. Objective

Develop an algorithm that can automatically extract features

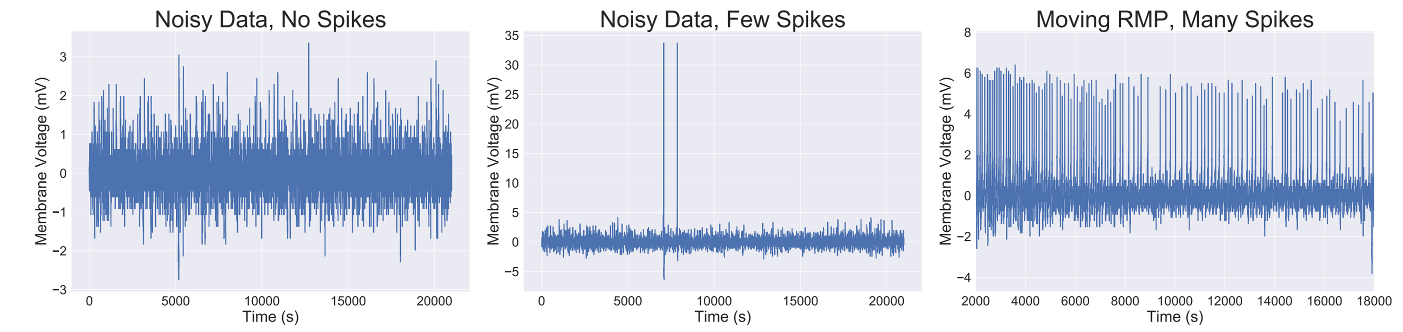
- Applicable to noisy data
- Requires minimal user input
- Can detect spikes of different shapes and sizes
- Robust to moving RMP and low sampling rates.

5. Outline of Spike Detection Method

- Subtract median "background" from signal.
- Use statistics of recording to guess algorithm parameters `thresh`, `prominence`, and `wLen`.
- Sweep reasonable ranges of parameters to find where the number of spikes detected remains stable.

5.a. Median Subtraction

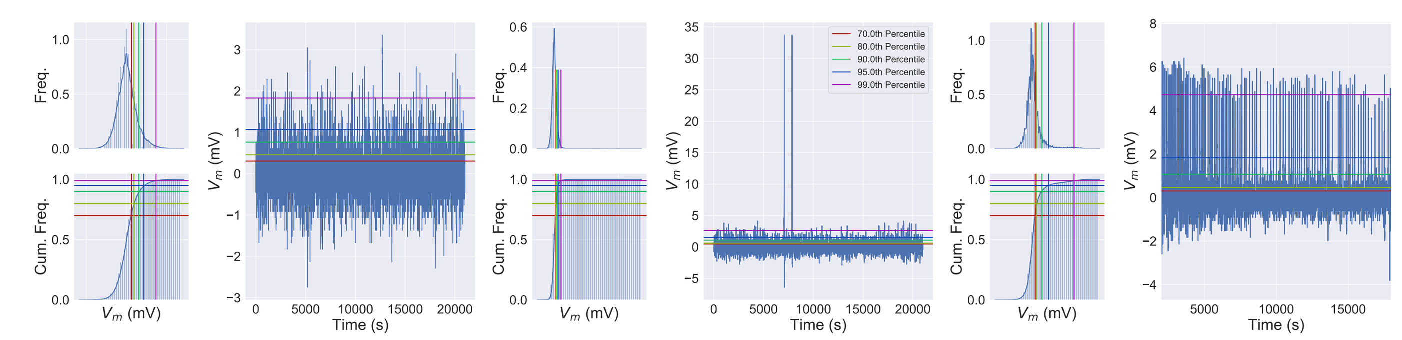
Median background subtraction removes changing resting membrane voltage as well as some noise.



- Window size for median filter must be specified, but results are not sensitive to this choice.
- The window size is reduced at the edges of the data to remove filter artifacts.

4.b. Parameter Initialization

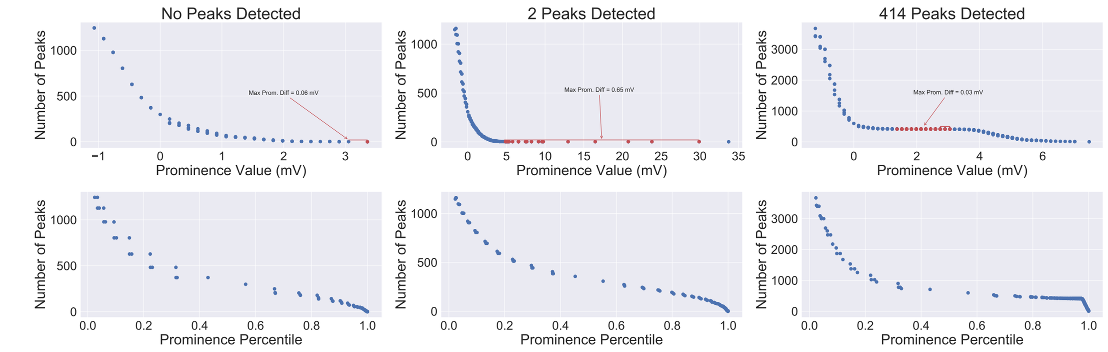
Distribution of voltages in median-subtracted data can be used to set `thresh` and `prominence`.



- Algorithm is relatively robust to choice of `thresh`, but setting it at the 50th-80th percentile is reasonable.
- The empirical CDF gives an array of possible values for `prominence`.

4.c. Prominence Sweeping for Determination of Stable Number of APs

Sweeping over `prominence`, we look for the spikes that are detected over the largest range of `prominence`.



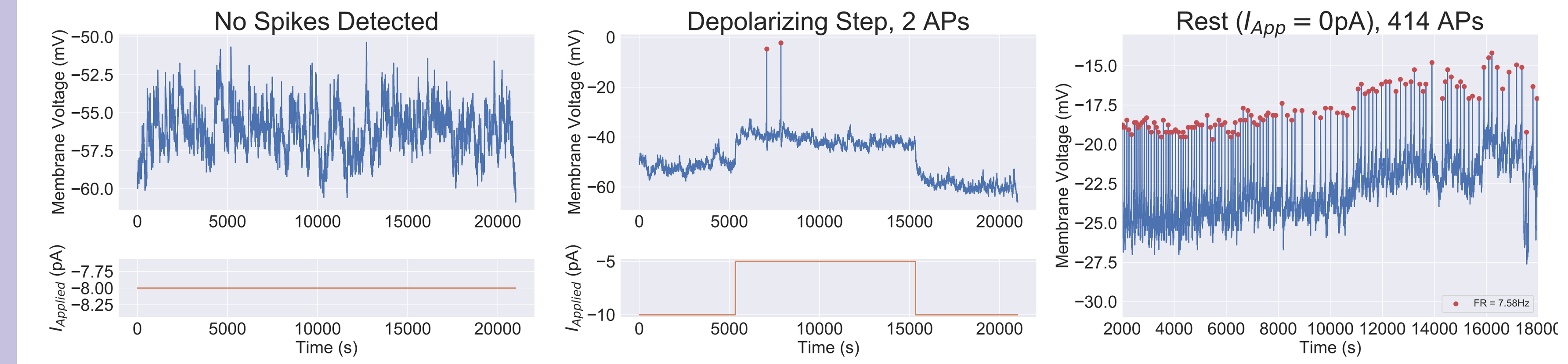
- The user must supply the minimum slope that an AP can take, so that `wLen` can be calculated:

$$wLen = Prom/Slope_{min}/dt,$$

where dt is the resolution of the data.

6. Results

Applying the method to the data, we can see that the correct number of spikes, and their positions, are found.



This holds up consistently if we compare the algorithm's results to those gathered "by-hand".

7. Conclusions

- A method for automatically detecting spikes from noisy electrophysiology has been proposed and verified on real data.
- Combination of this algorithm with other automated feature extraction methods allows for new techniques to be applied, such as multi-objective optimization of non-linear neuron models.
- Python code available on Github soon!

9. References

Abouzeid, A., Spruston, N., Kath, W. (2015). Fully-automated multi-objective optimization for fitting a neuronal model with real morphology. BMC Neuroscience, 16(Suppl 1), P117.

Allada, R., Chung, B. Y. (2010). Circadian Organization of Behavior and Physiology in Drosophila. Annual Review of Physiology, 72(1), 605â624.

Flourakis, M., Kula-Eversole, E., Hutchison, A. L., Han, T. H., Aranda, K., Moose, D. L., â Allada, R. (2015). A Conserved Bicycle Model for Circadian Clock Control of Membrane Excitability. Cell, 162(4), 836â848.

Siegle JH, Cuevas Lopez A, Patel YA, Abramov K, Ohayon S, Voigts J (2017) Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology. J Neural Eng 14: 045003

10. Acknowledgements

This work was supported by the ARO Grant W911NF1610584

The authors would like to thank Joshua Levy for helpful discussions concerning the implementation of the algorithm, as well as the Allada group at Northwestern for helping with the data.

kath@northwestern.edu

ericjohnson1.2015@u.northwestern.edu