Mini-Project (ML for Time Series) - MVA 2025/2026

Deconvolution of Calcium Imaging Signals via L_0 Regularization

Samuel Lhayani samuel.lhayani@etu.minesparis.psl.eu
Clément Nober nober.clement@gmail.com

January 15, 2026

1 Introduction and Contributions

Scientific Context: The Challenge of Calcium Imaging Deciphering the "neural code" the timing of action potentials or "spikes" is a central goal in neuroscience. Calcium imaging (e.g., GCaMP6) monitors intracellular surges, but the resulting fluorescence traces are noisy, low-pass filtered "spike-and-decay" signals. Consequently, spike inference is an inverse deconvolution problem: recovering discrete underlying activity from corrupted continuous measurements.

From Convex Relaxations to L_0 Optimization Traditional L_1 (Lasso) methods, while efficient, introduce "shrinkage bias," systematically underestimating spike amplitudes. Jewell and Witten (2018) propose a paradigm shift using an L_0 penalty, which penalizes the number of spikes rather than their magnitude. Leveraging the AR(1) model structure, they achieve a global optimum via functional dynamic programming, yielding exact, "all-or-nothing" spikes that avoid the magnitude bias of earlier methods.

Objectives and Experimental Design We led the following studies:

- **Comparative Analysis:** Benchmarking the accuracy of exact L_0 inference against convex L_1 relaxations and the PELT changepoint algorithm on real-world *SpikeFinder* data.
- **Parameter Sensitivity:** Evaluating the stability of the L_0 model by perturbing the decay constant γ, a critical biological parameter that governs the exponential decay of calcium.
- **Structural Failure Modes:** Probing algorithmic limits with the Temporal Refractory Period.

Contributions and Repartition of Work The workload for this project was distributed as follows:

- **Repartition of work:** We both worked on the theoritical content of the article as we needed to understand how the algorithm is working.
  - **Samuel:** Focused on the L_0 vs. L_1 benchmarking. He conducted the sensitivity analysis for the decay constant γ and the scenario-based experiments regarding the temporal refractory period and amplitude saturation.
  - **Clement:** Focused on the implementation and benchmarking of the PELT algorithm using the *ruptures* library.

- **Source code usage:** We integrated an existing implementation of the L_0 functional dynamic programming solver from `https://github.com/jewellsean/LZeroSpikeInference` and the L_1 framework from `https://github.com/j-friedrich/OASIS`. This allowed us to leverage the mathematically exact optimization backbone proposed by Jewell and Witten. However, we developed the entire simulation pipeline, the SpikeFinder data ingestion scripts.
- **New experiments:** We designed novel stress tests to identify the "Artificial Refractory Period", "Gamma sensitivity" and a comparison with PELT algorithm.

2 Method

The problem of spike inference is framed as a signal deconvolution task where we aim to recover a sparse underlying signal (the spikes) from a noisy, filtered observation (the fluorescence). The approach proposed by Jewell and Witten (2018) utilizes an L_0 penalty to enforce sparsity without the amplitude bias associated with more common L_1 methods.

2.1 The Generative Model for Calcium Dynamics

The relationship between neuronal spiking activity and the recorded fluorescence is typically modeled using a first-order autoregressive process, denoted as AR(1). Let y_t be the observed fluorescence at time t for $t = 1, \ldots, T$. We assume that y_t is a noisy measurement of the true, unobserved calcium concentration c_t:

$$y_t = c_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma^2) \tag{1}$$

The underlying calcium concentration c_t evolves according to the dynamics of the indicator (e.g., GCaMP). When a neuron fires, the concentration increases instantaneously by an amount s_t, followed by an exponential decay back to baseline. This is expressed as:

$$c_t = \gamma c_{t-1} + s_t, \quad s_t \geq 0 \tag{2}$$

Here, $\gamma \in (0, 1)$ is the decay parameter, related to the half-life of the calcium indicator. The condition $s_t \geq 0$ ensures that spikes are purely excitatory, reflecting the physiological reality that calcium concentration cannot decrease due to a spike.

2.2 The L_0 Optimization Framework

Traditional approaches often employ the L_1 norm (Lasso) as a convex proxy for sparsity: $\min \sum (y_t - c_t)^2 + \lambda \sum |s_t|$. However, L_1 penalties are known to cause "shrinkage," where the estimated amplitudes of the spikes are systematically smaller than the truth. To avoid this, we solve the L_0 penalized problem:

$$\min_{c_1, \ldots, c_T} \left\{ \frac{1}{2} \sum_{t=1}^{T} (y_t - c_t)^2 + \lambda \sum_{t=2}^{T} \mathbb{I}(c_t > \gamma c_{t-1}) \right\} \tag{3}$$

subject to the constraint $c_t \geq \gamma c_{t-1}$ for all t. Here, $\mathbb{I}(\cdot)$ is the indicator function, and λ is a tuning parameter that represents the "cost" of adding a spike. In this formulation, we only pay the penalty λ if a spike occurs ($s_t > 0$). If $s_t = 0$, the concentration simply follows the decay $c_t = \gamma c_{t-1}$ and no cost is incurred.

2.3 Exact Inference via Functional Dynamic Programming

The L_0 objective is non-convex and typically NP-hard. However, the specific structure of the AR(1) process allows us to find the global optimum efficiently using functional dynamic programming. We define $Q_t(c_t)$ as the minimum cost to reach state c_t at time t:

$$Cost_s^*(\alpha) = \min\{Cost_{s-1}^*(\alpha/\gamma), \min_{\alpha' \leq \alpha/\gamma} Cost_{s-1}^*(\alpha') + \lambda\} + \frac{1}{2}(y_s - \alpha)^2 \tag{4}$$

This Bellman equation can be split into two cases:

1. **The No-Spike Case:** $\alpha = \gamma\alpha'$. The cost is simply $Cost_{s-1}^*(\alpha/\gamma) + \frac{1}{2}(y_s - \alpha)^2$.
2. **The Spike Case:** $\alpha > \gamma\alpha'$. The cost involves the penalty λ plus the minimum of $Cost_{s-1}^*$ over the feasible range.

The algorithm maintains $Cost_s^*(\alpha)$ as a piecewise quadratic function. At each time step t, the algorithm:

- Updates the cost-to-go by incorporating the new observation y_t.
- Computes the minimum cost of having a spike vs. not having a spike.
- Prunes the search space by identifying regions of c_t that can never be part of the optimal path.

3 Data

Two sources of data were used:

1. **Benchmark Data:** We utilized the *SpikeFinder* competition dataset (sourced from the Allen Brain Institute), which provides real calcium fluorescence paired with simultaneous electrophysiology. This dataset was comparison and the gamma sensitivity analysis.
2. **Synthetic Stress-Tests:** We generated purely synthetic traces to study the artificial Temporal Refractory Period. This allowed us to control edge cases that are rarely isolated in real biological recordings.

3.1 Spikefinder Data and Preprocessing

The **Spikefinder challenge** serves as a standardized benchmarking framework for spike deconvolution, aggregating data from multiple laboratories to provide a robust "ground truth" through simultaneous electrophysiological recordings. The challenge is composed of 10 different databases that represent a wide variety of experimental conditions. These datasets vary significantly in several key dimensions:

- **Noise and Signal Heterogeneity:** The databases exhibit very different noise levels and signal-to-noise ratios depending on the specific calcium indicator used (e.g., OGB-1, GCaMP5k, GCaMP6f/s, jRCAMP1a, jRGECO1a) and the scanning technology employed (AOD, galvo, or resonant).
- **Experimental Diversity:** The recordings cover various circuits, including the primary visual cortex (V1) and the retina.

- **Kinetics:** Indicators are classified into fast, medium, and slow time-scale categories, each necessitating different decay parameters (γ) for accurate deconvolution.

For the practical implementation of our analysis, we utilized zooms on windows of 3,000 timesteps. This windowing approach was chosen for its practicality in managing non-NaN values, ensuring that the recursive deconvolution algorithms, which rely on continuous exponential decay models, operate on valid, uninterrupted signal segments. Given that individual traces can consist of up to 100,000 timesteps, this segmentation allows for efficient processing while maintaining the integrity of the generative model used by the L_0 and PELT algorithms.

3.2 Synthetic Generative Process (Scenario Experiments)

For the Temporal Refractory Period analysis, we "made up" synthetic time series of length $T = 1000$ to $T = 10,000$ to mimic 30Hz GCaMP6s recordings.

- **Biological Parameters:** We set the decay parameter $\gamma = 0.98$, reflecting the fast kinetics of GCaMP6f.
- **Sparsity:** The underlying spiking was modeled as a random process. We deliberately violated this sparsity by reducing the inter-spike interval (ISI) to 2 frames to identify the "Refractory Collision" limit.

4 Results

First, we replicate the comparative framework of the reference paper to establish a benchmark. We then proceed to a deeper analysis that challenges the fundamental assumptions of the models, specifically testing their robustness and failure modes when these hypotheses are violated.

To evaluate the structural advantages of the L_0 framework, we performed a comparative analysis against the L_1 (OASIS) relaxation using the *SpikeFinder* dataset.

4.1 Comparison: L_0 vs. L_1 Performance

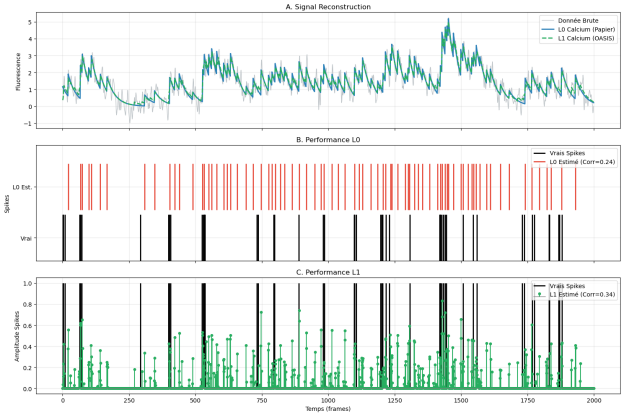


Figure 1: Comparison of L0 and L1 algorithm on spike detection

4.1.1 Van Rossum and Victor-Purpura distances

As shown in Figure 2, the L_0 regularizer consistently outperforms L_1 thresholds in minimizing the Van Rossum distance. While L_1 requires precise threshold tuning to avoid the extremes of noise amplification (Threshold 0.05) or signal loss (Threshold 0.7), the L_0 path creates a convex error curve with a significantly lower minimum, indicating superior reconstruction fidelity relative to the ground truth of 1981 spikes.

To further validate the robustness of the L_0 algorithm, we compared its performance against L_1 using two specialized spike-train metrics. These metrics account for temporal precision, penalizing misaligned spikes more gracefully than binary accuracy.

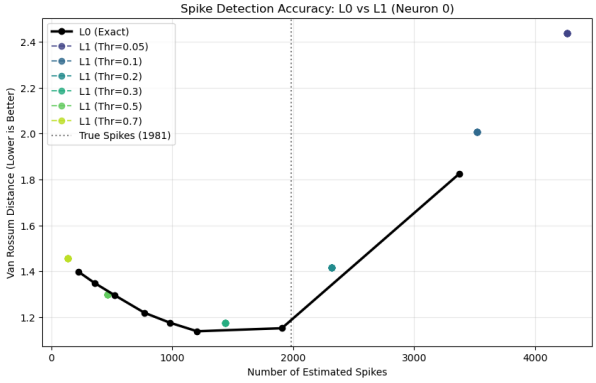


Figure 2: **Van Rossum Distance.**

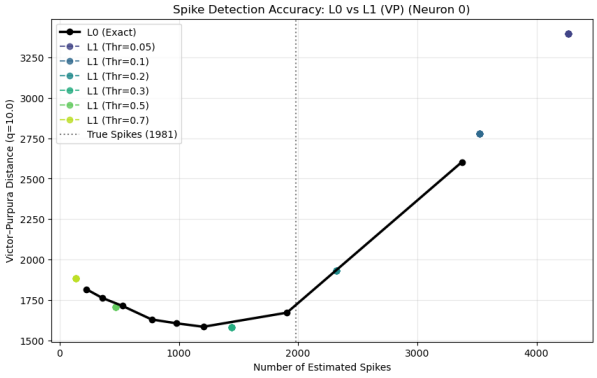


Figure 3: **Victor-Purpura Distance.**

Figure 4: **Side-by-side Benchmarking of Temporal Metrics.** Both plots demonstrate that L_0 maintains higher structural fidelity under noise, whereas L_1 generates excessive "noise-spikes" that increase the total distance in both metric spaces.

4.2 PELT vs. L0

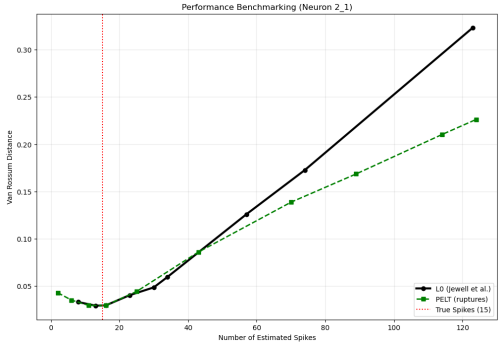
4.2.1 Parameterization and Constraint

For the PELT implementation, we utilized a **minimum segment size (min_size) of 5 frames**. This hyperparameter acts as a critical temporal filter, ensuring that detected "states" or spikes possess a minimum duration. This constraint prevents the algorithm from overfitting to single-frame

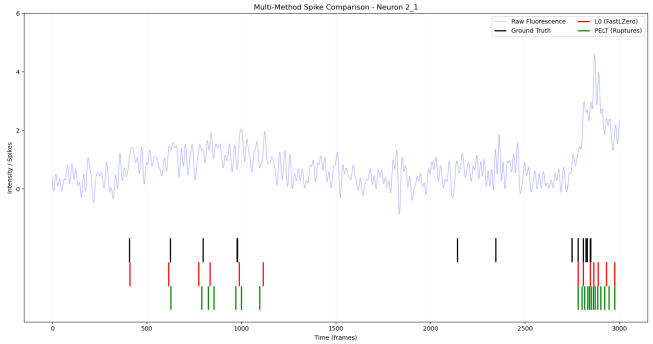
stochastic noise, which would otherwise result in a high rate of false positives in high-frequency noise regimes.

4.2.2 Benchmark Analysis

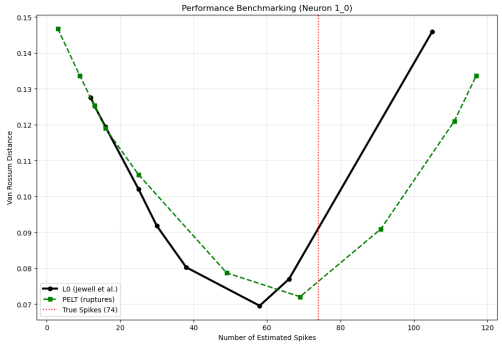
While the PELT algorithm is capable of achieving performance metrics that are comparable to the L0 algorithm in terms of reconstruction fidelity, this similar level of accuracy comes at a significantly higher computational cost. Specifically, PELT optimizes the detection of regime shifts through a pruned exact search that, although efficient for dynamic programming, remains more computationally intensive than the faster sparse-recovery operations typically utilized by L0 regularization.



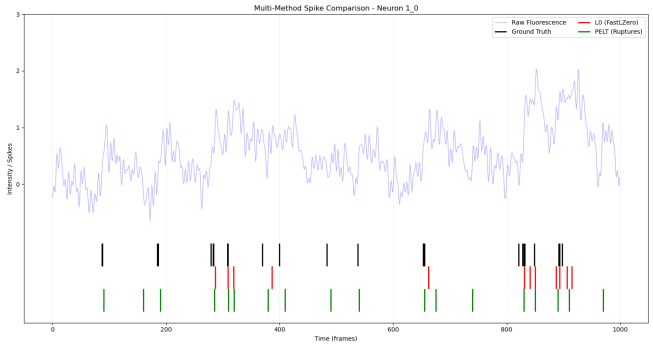
(a) Benchmark: Neuron 1 from spikefinder dataset 2



(b) Activity Trace: Neuron 1 from spikefinder dataset 2



(c) Benchmark: Neuron 0 from spikefinder dataset 1



(d) Activity Trace: Neuron 0 from spikefinder dataset 1

Figure 5: Multi-method comparison of L_0 and PELT ($min\_size = 5$) spike detection. Left column shows the Van Rossum distance relative to ground truth. Right column displays raw fluorescence aligned with predicted and true spikes.

4.3 Robustness Analysis

4.3.1 Sensitivity Analysis of the Decay Constant (γ)

A critical assumption in the standard calcium deconvolution framework is that the temporal decay constant, γ, is known and fixed. In experimental practice, however, γ is an approximation derived from the calcium indicator's kinetics and the imaging frame rate. Because biological parameters vary between neurons and over time, an algorithm that requires perfect knowledge of γ is brittle and ill-suited for real-world application.

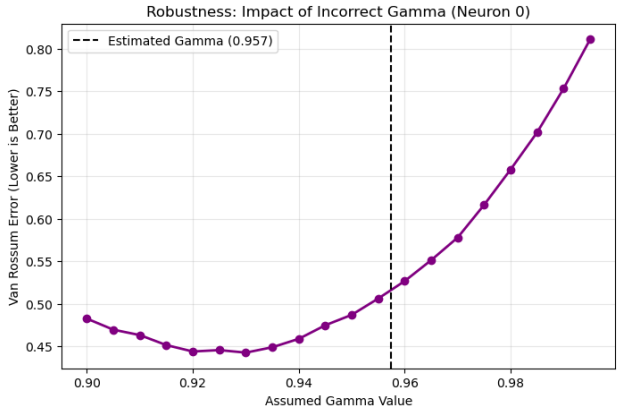


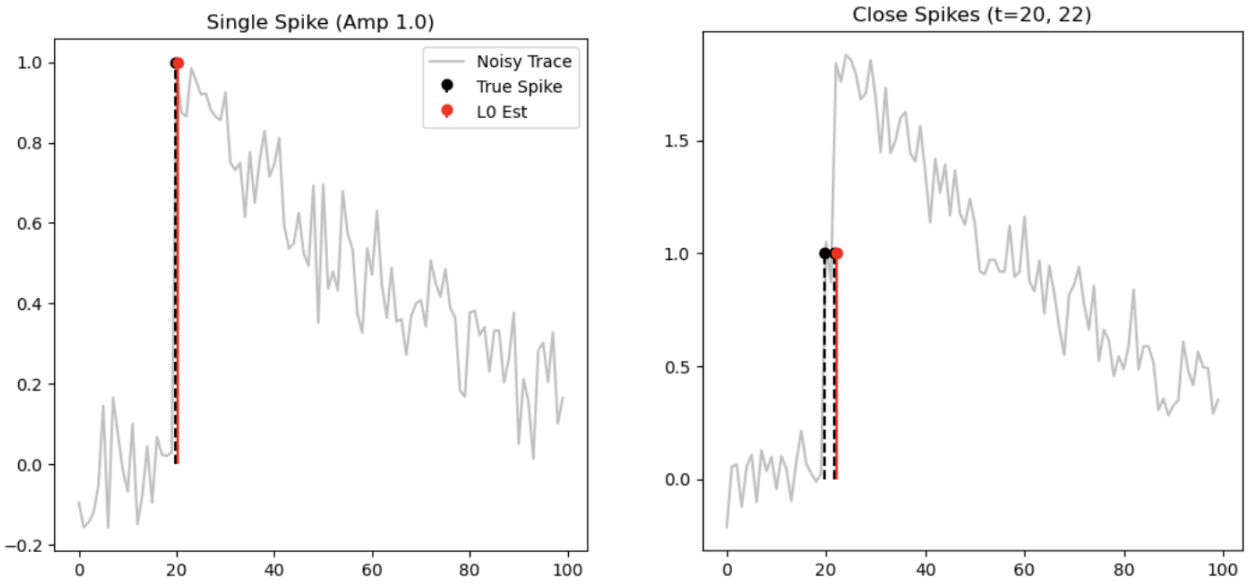
Figure 6: While the framework assumes the decay constant γ is fixed, our sensitivity analysis indicates that reconstruction quality relies heavily on parameter precision. We observed that even minor deviations (± 0.01) from the optimal γ can lead to a noticeable increase in error, with the optimal value often diverging stochastically from standard estimates. This suggests that the algorithm's robustness is constrained by the accuracy of the initial parameter estimation, making precise calibration critical for optimal performance.

To evaluate the robustness of the L_0 model against parameter mismatch, we conducted a stress test using a controlled perturbation strategy. We utilized a ground truth dataset with a known intrinsic decay γ_{true}. We then performed spike inference while forcing the algorithm to use an incorrect decay estimate, $\hat{\gamma}$, sweeping across the range $[0.90, 0.999]$.

This experiment (Figure 6) aims to identify the stability area for the model. Specifically, we investigate whether slight deviations in $\hat{\gamma}$ (e.g., 0.98 vs. 0.99) result in a graceful degradation of spike detection accuracy or a catastrophic failure (e.g., exploding False Positives).

4.3.2 Testing the "One-Spike-at-a-Time" Limitation

The generative model assumes calcium dynamics are driven by instantaneous jumps (z_t), theoretically treating a biological "burst" of n spikes as a single jump of amplitude $n \times A$. We implemented a Linearity and Saturation Test to determine if the L_0 solver correctly scales these high-density events or if the penalty (λ) suppresses them. Our hypothesis is that the algorithm may exhibit amplitude saturation, failing to distinguish multi-spike bursts from single events due to the regularization cost.



(a) Sparse Spike Recovery

(b) Temporal Resolution Limit

Figure 7: **Testing the structural limits of L_0 deconvolution.** (a) The algorithm recovers an isolated spike perfectly at $t = 20$ with amplitude ≈ 1.0. (b) For spikes separated by only 1 frame, the L_0 penalty (λ) makes a single "compromise" spike mathematically cheaper than detecting two distinct events, creating an artificial refractory period.

References

[1] Jewell, S. and Witten, D. (2017). Exact Spike Train Inference Via ℓ_0 Optimization. *arXiv preprint arXiv:1703.08644*. https://arxiv.org/abs/1703.08644

[2] Jewell, S., Hocking, T. D., Fearnhead, P., and Witten, D. (2018). Fast Nonconvex Deconvolution of Calcium Imaging Data. *arXiv preprint arXiv:1802.07380*. https://arxiv.org/abs/1802.07380

[3] Friedrich, J., Zhou, P., and Paninski, L. (2017). Fast online deconvolution of calcium imaging data. *PLoS Computational Biology*, 13(3), e1005423. https://doi.org/10.1371/journal.pcbi.1005423