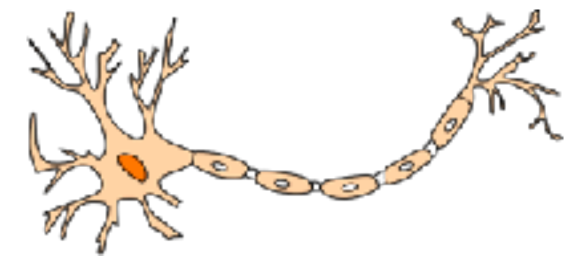
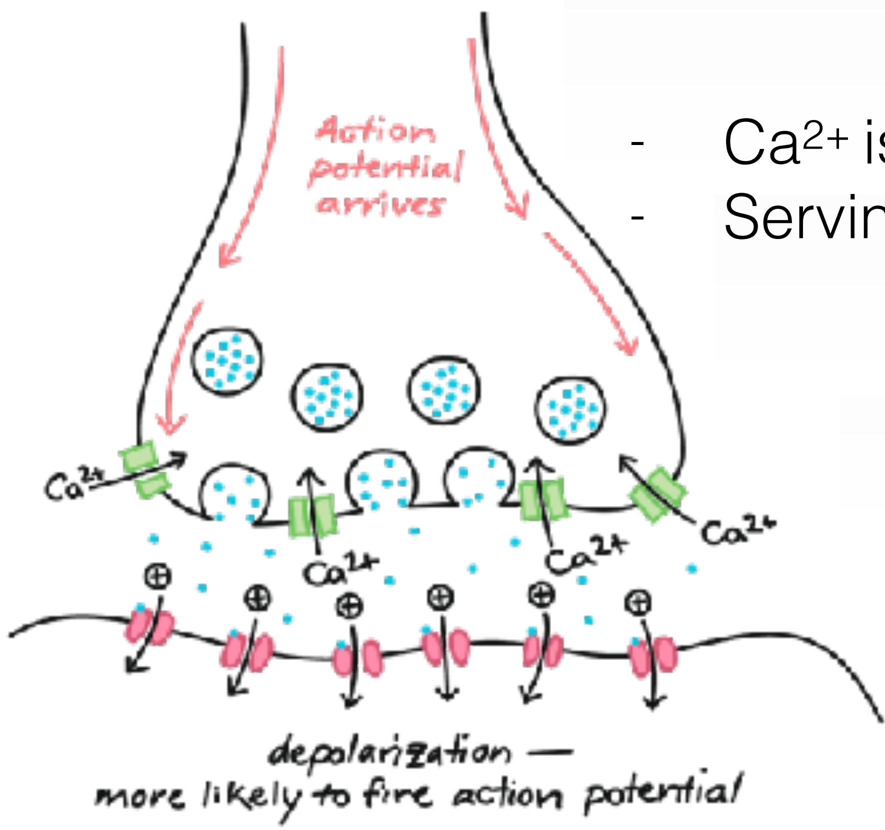
Exact spike train inference via ℓ_0 optimization[1]

Author: S. Jewell & D. Witten (AoAS 2018)

Presenter: Yiqun Chen

Calcium imaging data: what is it and why do we care?





- Ca^{2+} is essential for **information transmission** in the brain;
- Serving as a good **proxy of activity**

Calcium imaging data: what is it and why do we care?

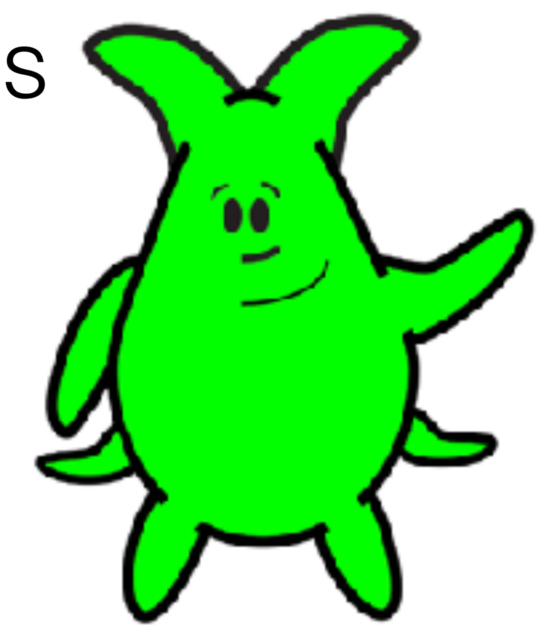
We genetically modify mice genome such that the neurons will be **fluorescent**..

When neuron 1 is active/spikes

Ca^{2+} Ca^{2+} Ca^{2+}

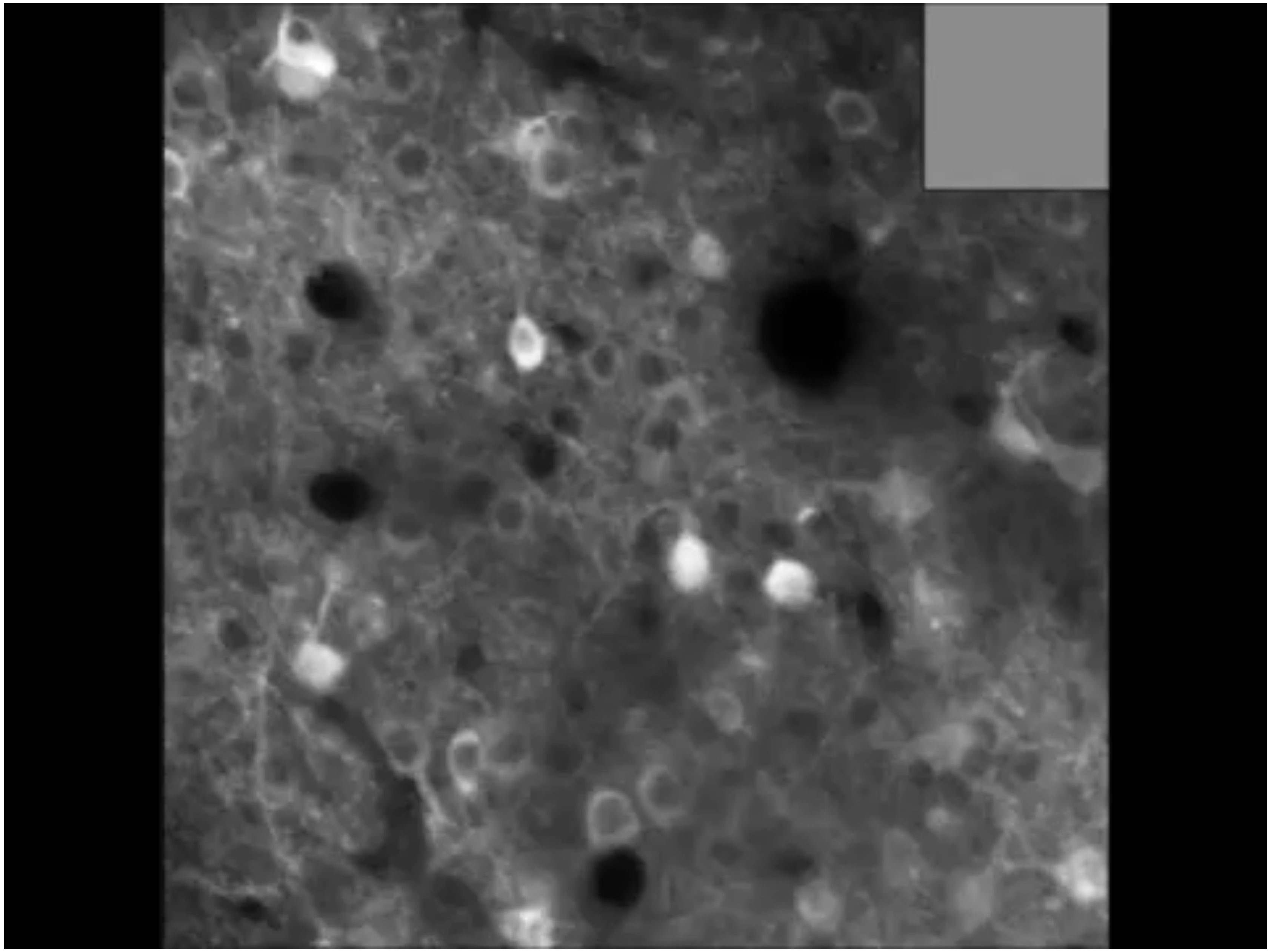
Ca^{2+} Ca^{2+} Ca^{2+}

Ca^{2+} Ca^{2+} Ca^{2+}



Neuron 1

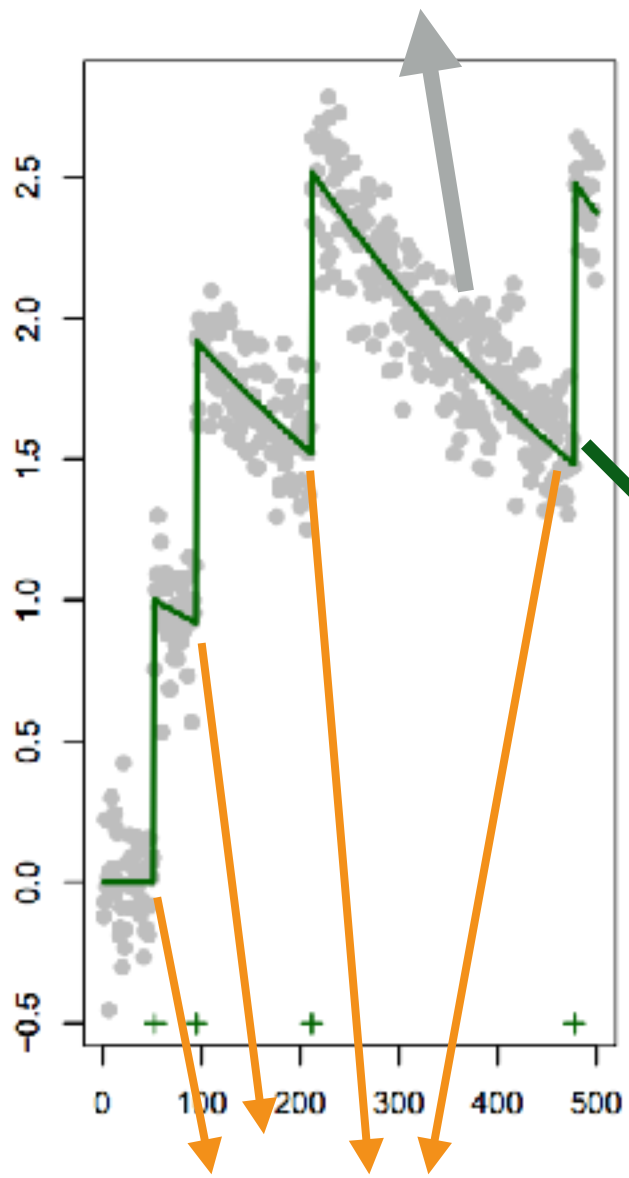
Neuron 1



Calcium imaging from Mouse V1
- Courtesy of the Sur Lab

A model for neuron spiking

Discrete-time measurements



Calcium concentration

Spikes

- Usually decays **exponentially**

$$C_t = \gamma C_{t-1} + s_t$$

Unless a **spike** drives it up

$$F_t = C_t + \epsilon_t$$

Observed fluorescence is a **noisy** version of the Ca2+ concentration

Optimization problem for neuron spiking

$$c_t = \gamma c_{t-1} + s_t$$

$$F_t = c_t + \epsilon_t$$

Non-negative tuning parameter

ℓ_0 penalty!

$$\text{minimize}_{c_1,\ldots,c_T} \sum_{t=1}^{T} (F_t - c_t)^2 + \boxed{\lambda \cdot \sum_{t=2}^{T} 1\{c_t \neq \gamma c_{t-1}\}}$$

Fit noisy fluorescence

But we shouldn't fit a ton of spikes

Ideally we wanna solve

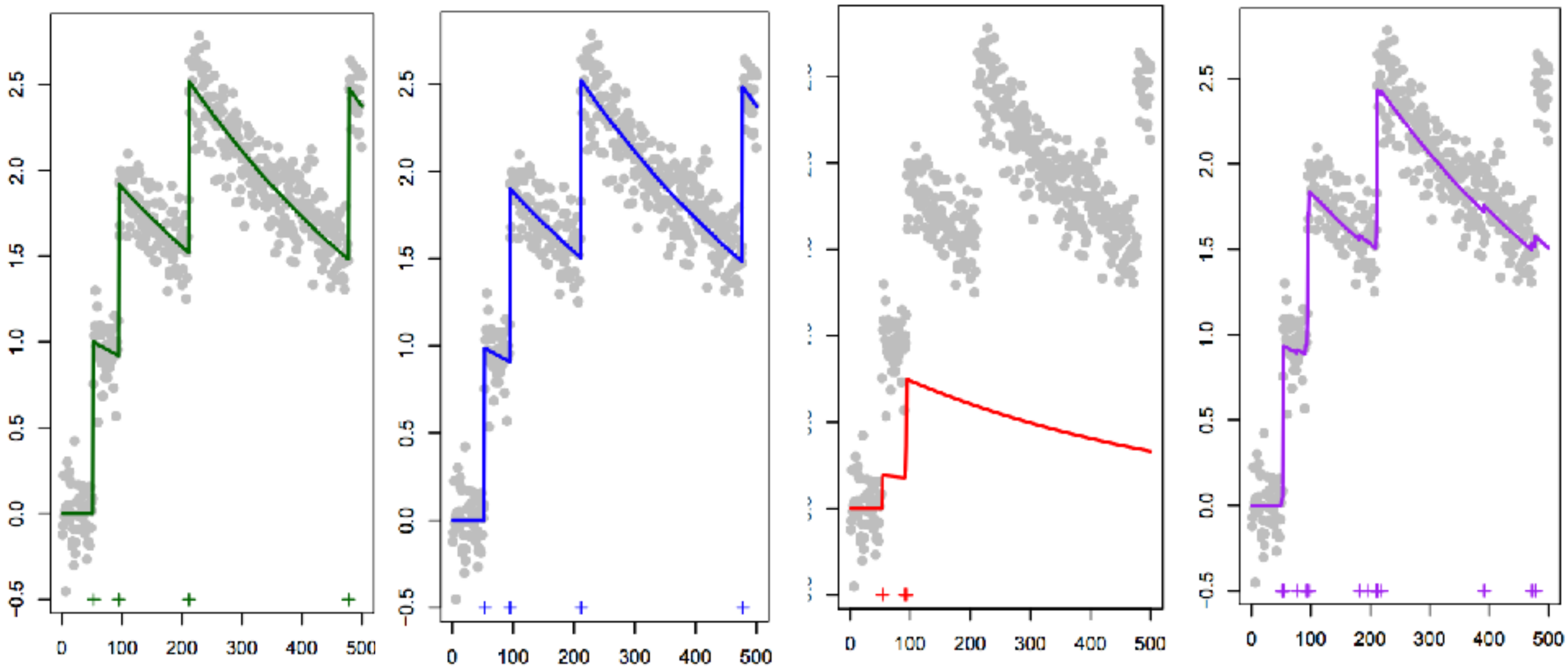
$$\text{minimize}_{c_1,\cdots,c_T} \sum_{t=1}^{T} (F_t - c_t)^2 + \lambda \cdot \underline{\sum_{t=2}^{T} 1\{c_t \neq \gamma c_{t-1}\}}$$

But *naively* this takes $O(2^T)$ operations

So let's instead try the following…

$$\text{minimize}_{c_1,\cdots,c_T;s_2,\cdots,s_T}\left\{\frac{1}{2}\sum_{t=1}^{T} (y_t - c_t)^2 + \lambda \underline{\sum_{t=2}^{T} | c_t - \gamma c_{t-1} |}\right\}$$

This now is computationally tractable!
But is it good?



True data generating mechanism

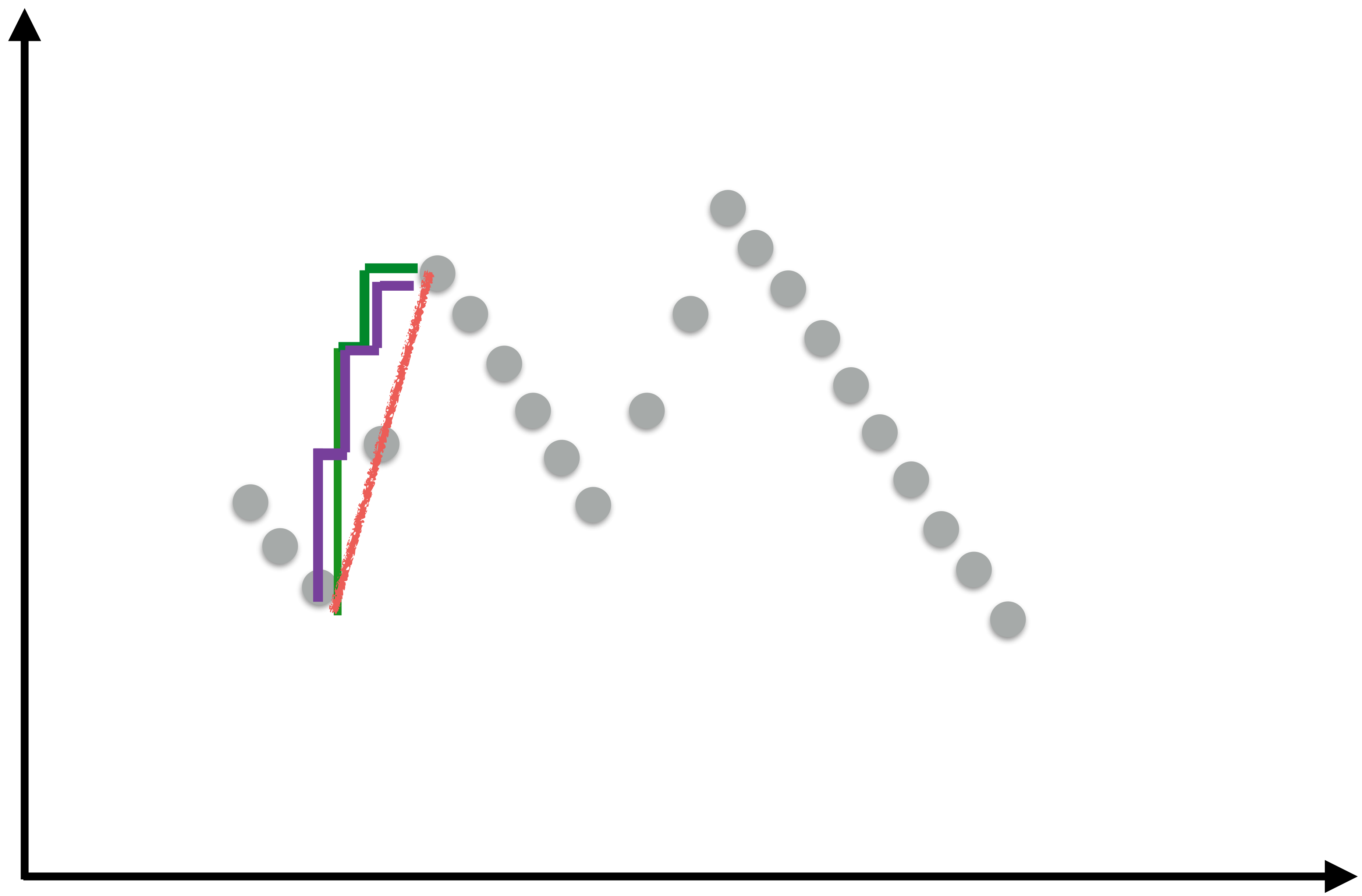
ℓ_0 with small λ

ℓ_1 with large λ

ℓ_1 with small λ

:(Fast computation seems to come with a cost!!

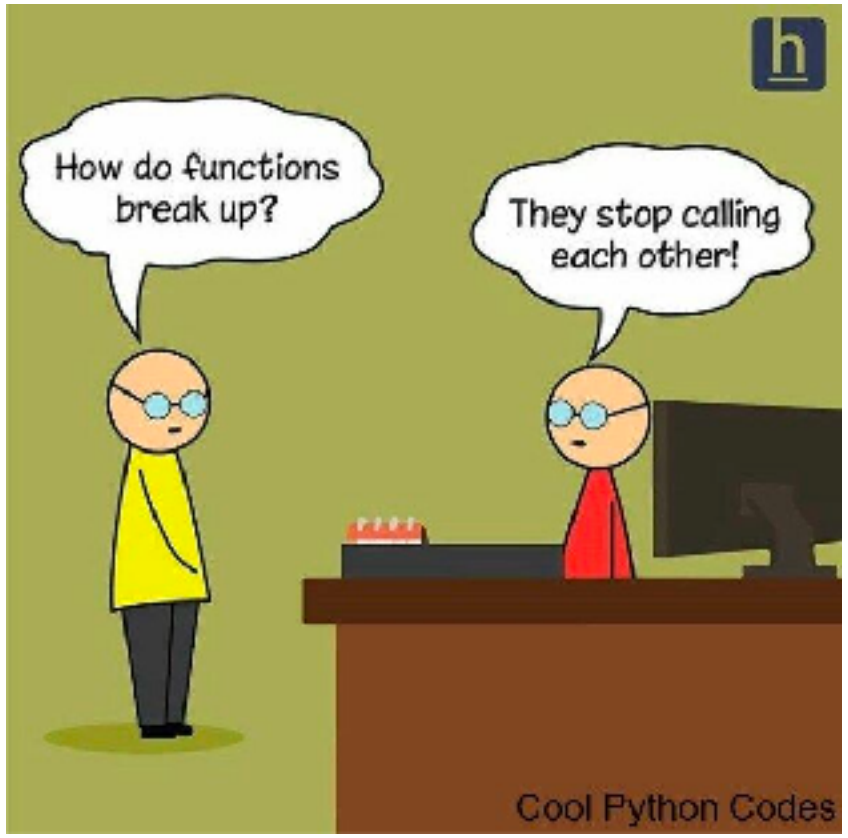
ℓ_0 versus ℓ_1



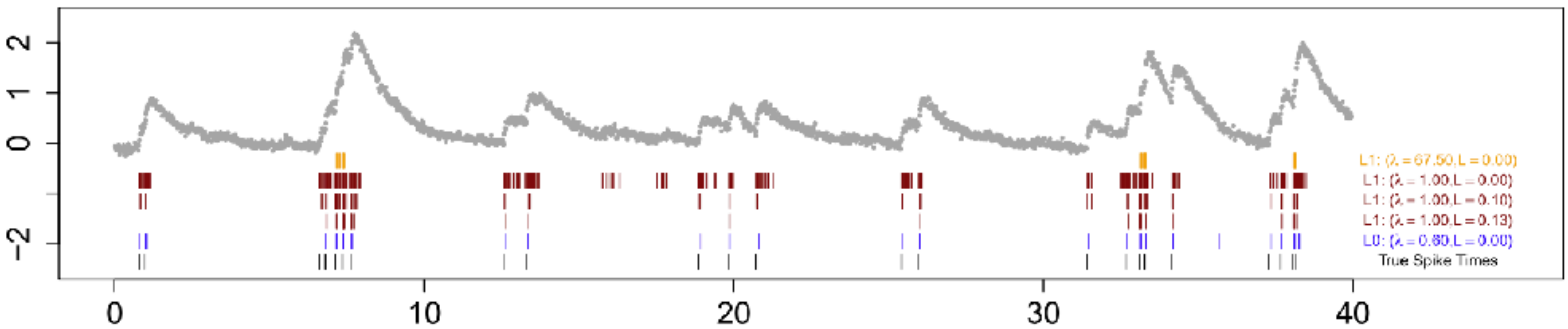
Contribution of this paper

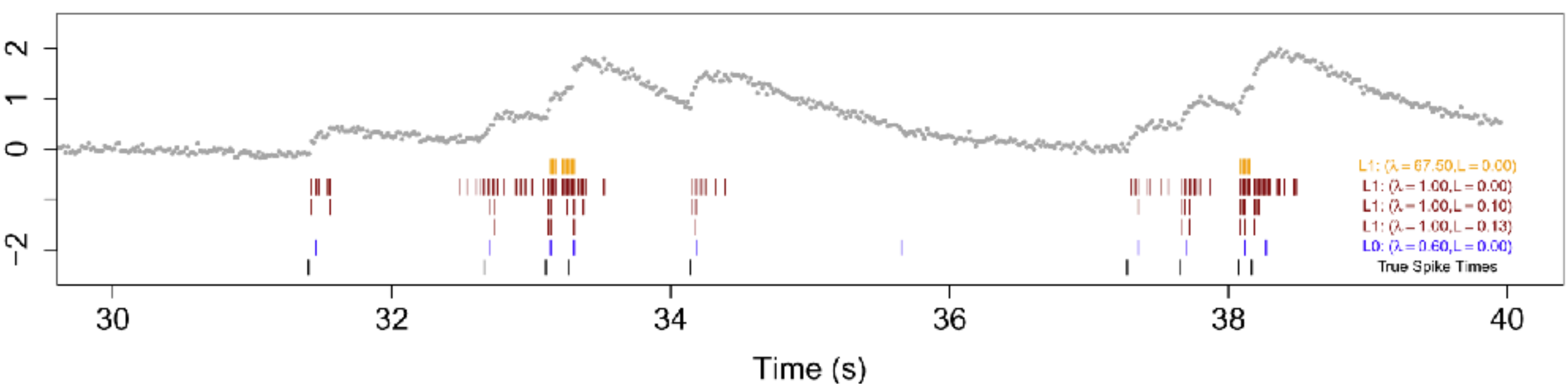
$$\text{minimize}_{c_1,\ldots,c_T} \sum_{t=1}^{T} (F_t - c_t)^2 + \lambda \cdot \sum_{t=2}^{T} 1\{c_t \neq \gamma c_{t-1}\}$$

Can be re-casted and solved **efficiently** in **$O(T^2)$!**
using
dynamic programming (aka smart recursion)



Application of the proposed algorithm to data in Chen et al. (2013)





Reference:

Chen, T.-W., et al. (2013). Ultra-sensitive fluorescent proteins for imaging neuronal activity. *Nature*

Friedrich, J., Zhou, P., & Paninski, L. (2017). Fast online deconvolution of calcium imaging data. *PLoS Computational Biology*, *13*(3), e1005423.

Jewell, S., & Witten, D. (2018). EXACT SPIKE TRAIN INFERENCE VIA $\ell 0$ OPTIMIZATION. *The Annals of Applied Statistics*, *12*(4), 2457–2482.

Vogelstein, J. T., et al. (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. *Journal of Neurophysiology*, *104*(6), 3691–3704.

Thank you for your attention!
Any questions?