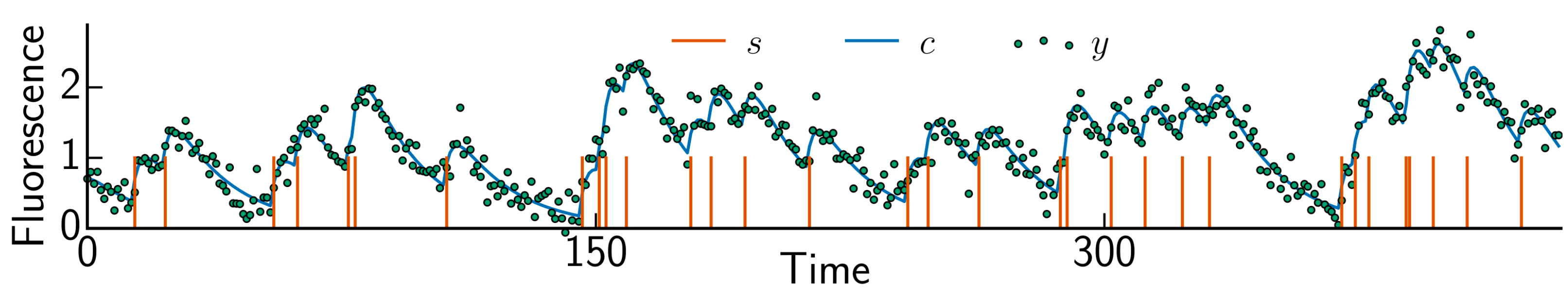
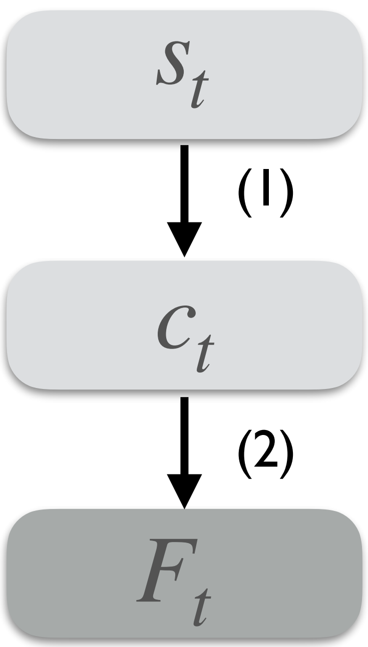
Exact spike train inference via ℓ_0 optimization

Yiqun Chen,
572 methodology presentation
05/07/2019



What you missed from the last season of my presentation…



Spike count

Calcium concentration

Observed fluorescence traces

$$\forall t \in [T]$$

$$s_t \overset{i.i.d.}{\sim} Poisson(\theta)$$

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

$$F_t = c_t + \epsilon_t, \epsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0,\sigma^2) \qquad (2)$$

Wait - how did we cast this as an optimization problem?

MAP estimator with **Poisson** prior for the spikes:

$$\hat{s}_t = \text{argmax}_{s_t \in \mathbb{N}} P(s_1, \cdots, s_T | F)$$

$$= \text{argmax}_{s_t \in \mathbb{N}, c_{1:T}} \sum_{t=1}^{T} -\frac{1}{2\sigma^2}(F_t - c_t)^2 + s_t \log \theta - \log s_t!$$

MAP estimator with **Exponential** prior for the spikes:

$$\hat{s}_t = \text{argmax}_{s_t > 0} \sum_{t=1}^{T} -\frac{1}{2\sigma^2}(F_t - c_t)^2 - s_t \cdot \theta$$

$$\overset{equiv}{=} \text{argmax}_{c_1, \cdots, c_T; c_t - \gamma c_{t-1} > 0} \sum_{t=1}^{T} -\frac{1}{2\sigma^2}(F_t - c_t)^2 + (c_t - \gamma c_{t-1})$$

$$\approx \text{argmin}_{c_1, \cdots, c_T} \frac{1}{2} \sum_{t=1}^{T} (F_t - c_t)^2 + \underline{\lambda | c_t - \gamma c_{t-1} |}$$

Some **post-thresholding** is usually performed

Wait - how did we cast this as an optimization problem?

MAP estimator with **Poisson** prior but **indicator** s_t:

$$\hat{s}_t = \text{argmax}_{s_t \in \{0,1\}} P(s_1, \cdots, s_T | F)$$

$$= \text{argmax}_{s_t \in \{0,1\}} \sum_{t=1}^{T} -\frac{1}{2\sigma^2}(F_t - c_t)^2 + s_t \log\theta - \log s_t!$$

$$= \text{argmin}_{s_t \in \{0,1\}} \frac{1}{2} \sum_{t=1}^{T} (F_t - c_t)^2 + \lambda \cdot s_t$$

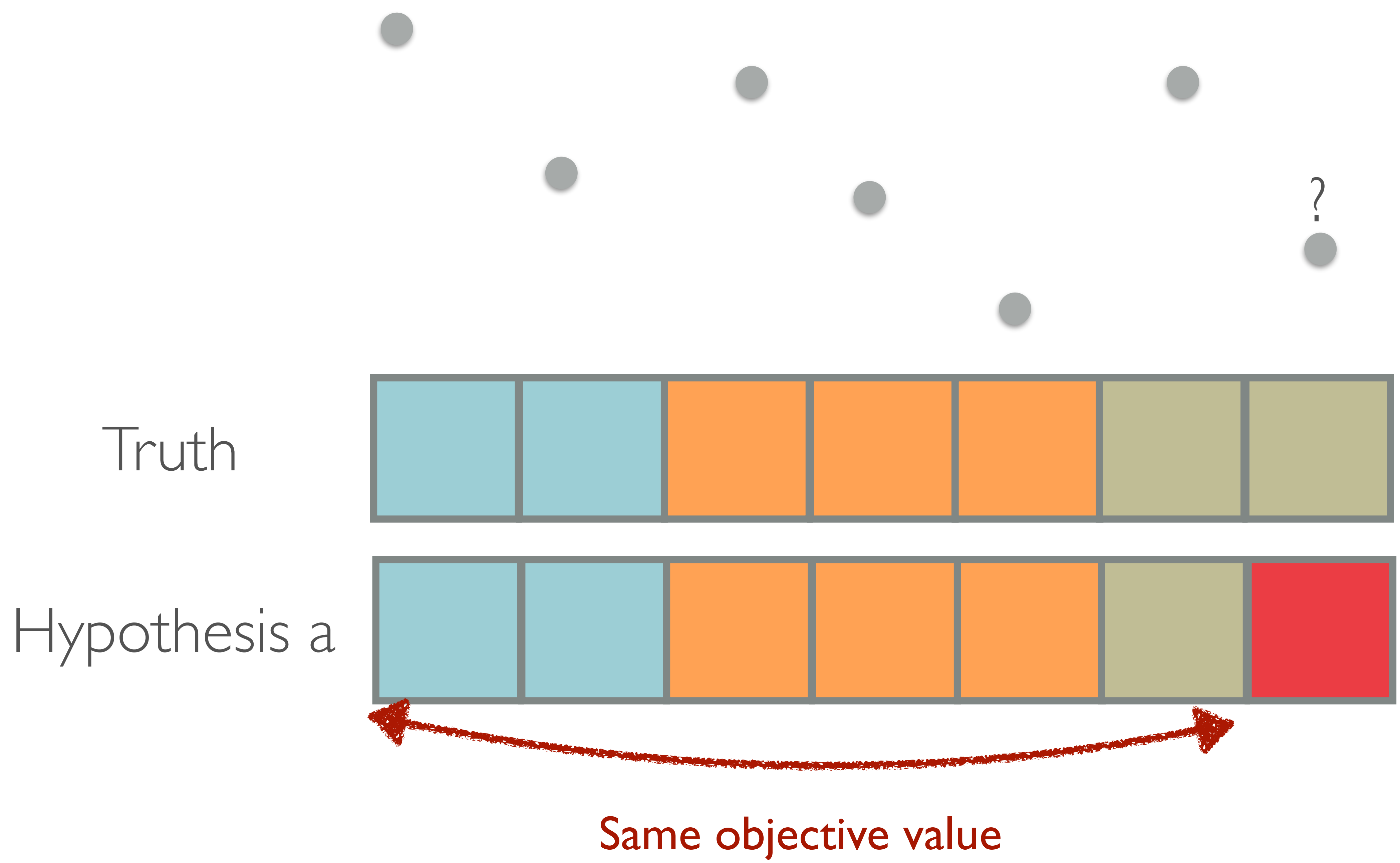
$$\hat{s}_t = \hat{c}_t - \gamma\hat{c}_{t-1} \overset{equiv}{=} \text{argmin}_{\underline{c_1, \cdots, c_T; c_t - \gamma c_{t-1} > 0}} \frac{1}{2} \sum_{t=1}^{T} (F_t - c_t)^2 + \lambda \cdot \mathbf{1}\{c_t \neq \gamma c_{t-1}\}$$

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

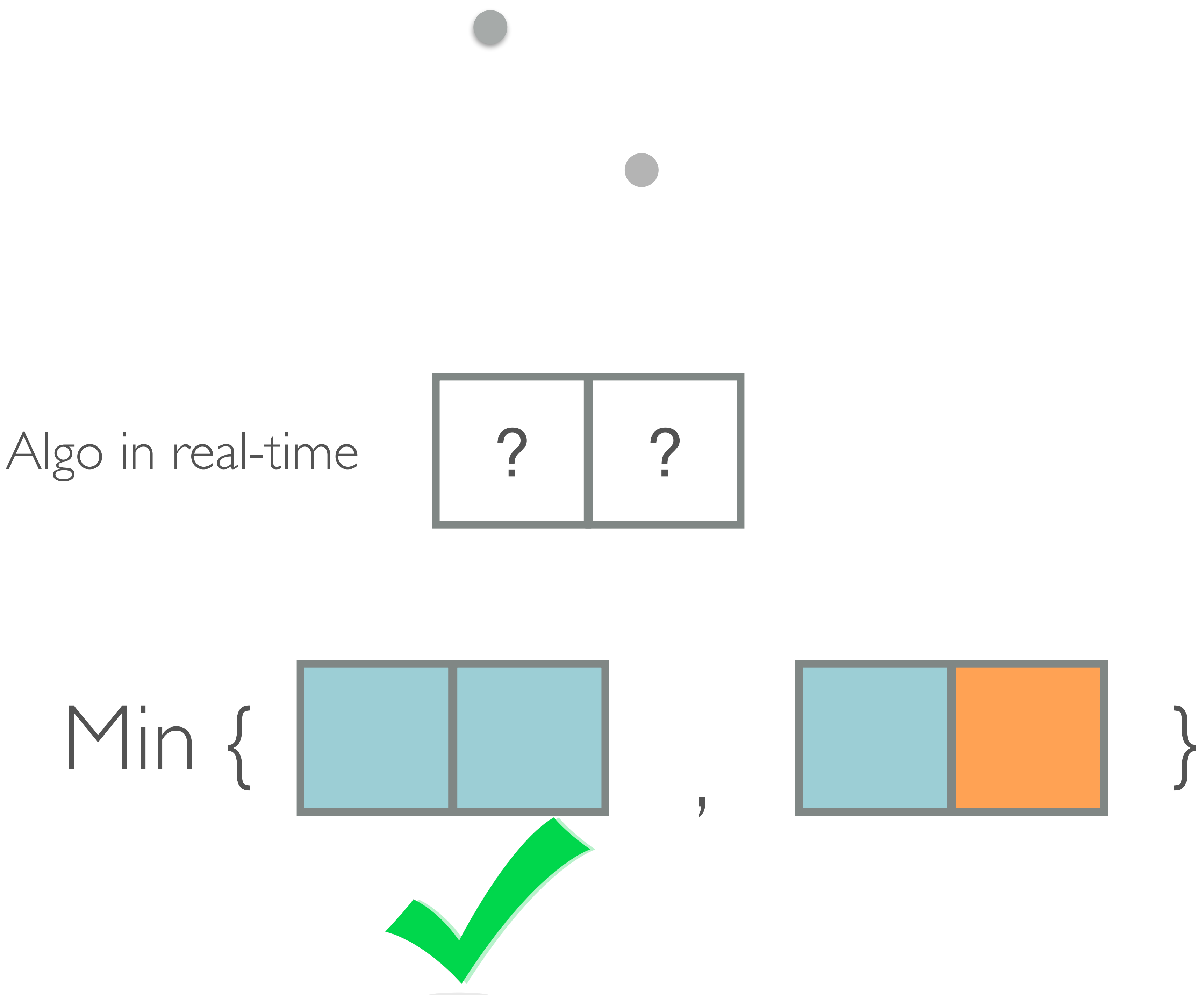
Presence of spikes should only **increase** the Ca^{2+} concentration

Observation: spikes **partition** the observed fluorescence

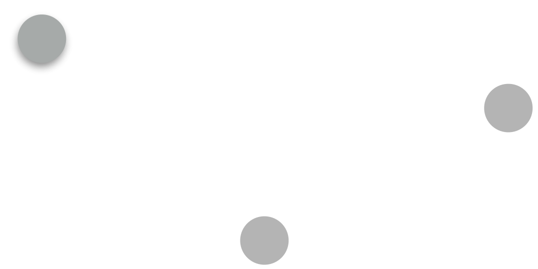
Given the occurrence of most recent spike,
what **happened after does not affect what happened before** the spike



Algo in real-time

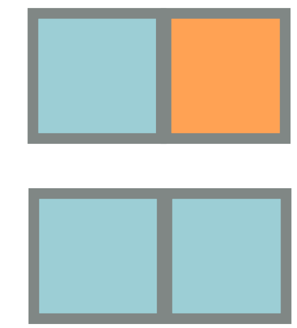


Algo in real-time
?
?
Min {
,
}



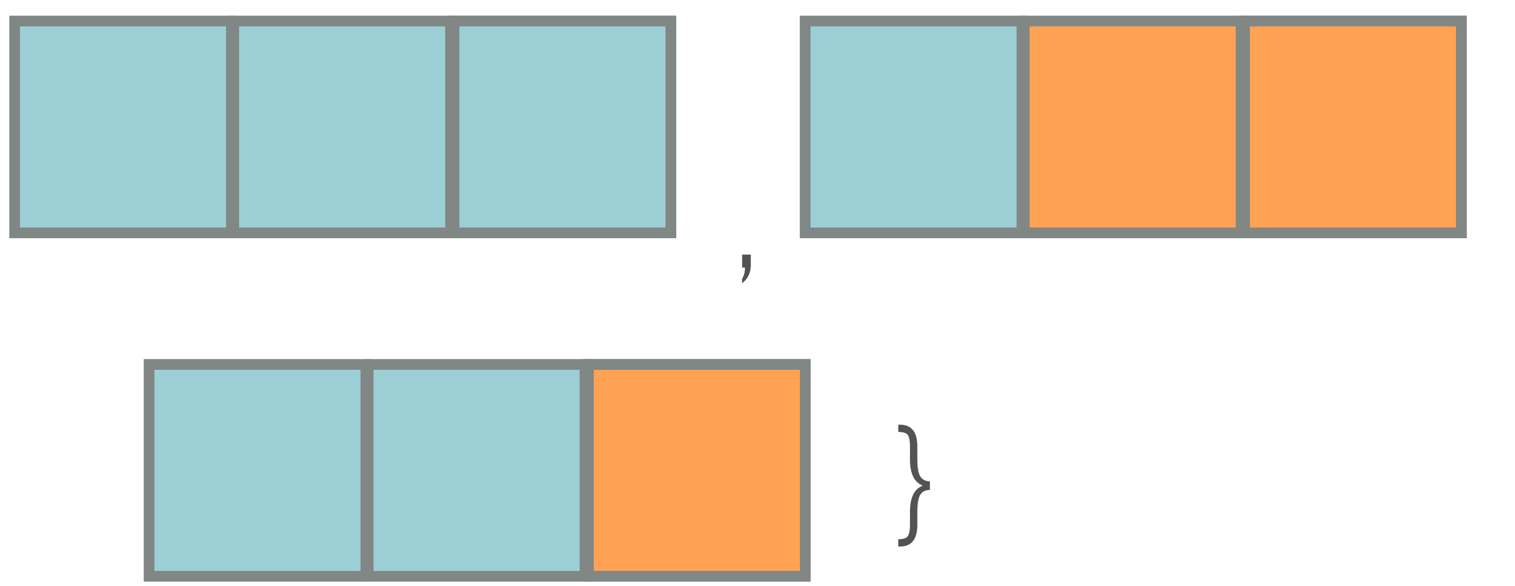
Algo in real-time ? ? ?

We computed



at time point 2!

Min {



}

Dynamic programming:

$$\min_{c_1,\cdots,c_T} \frac{1}{2}\sum_{t=1}^{T}(F_t - c_t)^2 + \lambda \sum_{t=2}^{T} \mathbf{1}\{c_t \neq \gamma c_{t-1}\}$$

$$= \min_{|\mathscr{P}|=K\in\mathbb{N}} \min_{\forall t\in[T]\setminus\mathscr{P}, c_t=\gamma c_{t-1}} \frac{1}{2}\sum_{t=1}^{T}(F_t - c_t)^2 + \lambda \sum_{t=2}^{T} \mathbf{1}\{c_t \neq \gamma c_{t-1}\}$$

$$= \min_{K\in\mathbb{N}} \min_{0<\tau_0<\cdots<\tau_K<\tau_{K+1}=T, c_{\tau_1}:c_{\tau_K}} \frac{1}{2}\sum_{j=0}^{K}\left[\underline{(F_{\tau_j+1} - c_{\tau_j+1})^2} + \underline{\sum_{l=1}^{\tau_{j+1}-\tau_j}(F_{\tau_{j+1}+l} - \gamma^l \cdot c_{\tau_{j+1}})^2}\right] + \lambda \cdot K$$

Right after a spike the fluorescence can **increase by an unspecified** amount

Otherwise we have **exponential decay**…

Dynamic programming:

$$G(T) = \min_{0<\tau_0<\cdots<\tau_K<\tau_{K+1}=T, K\in\mathbb{N}} \sum_{j=0}^{K} \left\{ \underline{\frac{1}{2} \min_{c_{\tau_j}+1} \left[\sum_{t=\tau_j+1}^{\tau_{j+1}} (y_t - \gamma^{t-(\tau_j+1)} \cdot c_{\tau_j+1})^2 \right]} \right\} + \lambda \cdot K$$

Too much effort to type - let's abbreviate it as $\mathbf{D}(y_{\tau_j} : y_{\tau_{j+1}})$

$$G(T) = \min_{0<\tau_0<\cdots<\tau_K<\tau_{K+1}=T, K\in\mathbb{N}} \sum_{j=0}^{K} \left\{ \mathbf{D}(y_{\tau_j} : y_{\tau_{j+1}}) + \lambda \cdot K \right\}$$

$$= \min_{s<T} \left\{ \min_{0<\tau_0<\cdots<\tau_K=s<T, K-1\in\mathbb{N}} \left[\sum_{j=0}^{K-1} \mathbf{D}(y_{\tau_j} : y_{\tau_{j+1}}) + \lambda \cdot (K-1) \right] + D(y_{\tau_K} : T) + \lambda \right\}$$

$$= \min_{s<T} \left\{ G(s) + \mathbf{D}(s : T) + \lambda \right\}$$

Obs:

1. Min is taking over **T** different things
2. G(s) values can be **memoized**/stored
3. D(s:T) can be computed in **O(1)**

$$G(t) = \min_{s<t} \left\{ G(s) + \mathbf{D}(s:t) + \lambda \right\}$$ takes O(t) steps to compute

Unroll this for loop (aka DP):

$$G(0) = -\lambda$$

$$G(1) = G(0) + \mathbf{D}(y_1) + \lambda$$

$$G(2) = \min\{G(0) + \mathbf{D}(y_{1:2}) + \lambda, G(1) + \mathbf{D}(y_2) + \lambda\}$$

Time complexity: $\sum_{t=1}^{T} t = O(T^2)$

Pruning of the active sets

$$G(t) = \min_{s<t} \left\{ G(s) + \mathbf{D}(s : t) + \lambda \right\}$$

Maybe we don't need to do it for all t possible candidates…

$$\mathscr{E}_{s+1} = \{\tau \in \{\mathscr{E}_s \cup s\} : G(\tau) + \mathbf{D}(y_{\tau+1:s}) < G(s)\}$$

It's **mathematically impossible** to have the most recent change point to have occurred before: **lower cost with not putting a spike there!**

In practice we can often eliminate some s: under some mild assumptions; **expected** time complexity is $\tilde{O}(T)$

$$c_t = \gamma c_{t-1} + s_t$$
$$F_t = c_t + \epsilon_t, \epsilon_t \overset{i.i.d}{\sim} \mathcal{N}(0,\sigma^2)$$

1. How to estimate γ?

   a. Estimate from a segment of exponentially decaying data (eye-ball test)
   b. (Follow-up paper) $\gamma = 1 - \frac{\Delta_t}{\phi}$

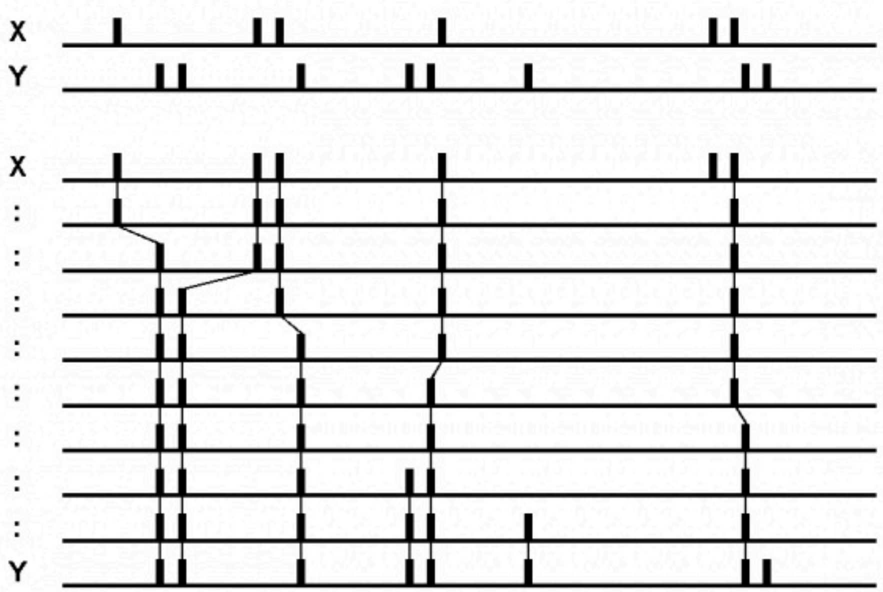
2. How to choose λ?

   2-fold CV to minimize MSE of estimated Ca^{2+}, e.g., $\hat{c}_2 = \frac{1}{2}(\hat{c}_1 + \hat{c}_3)$

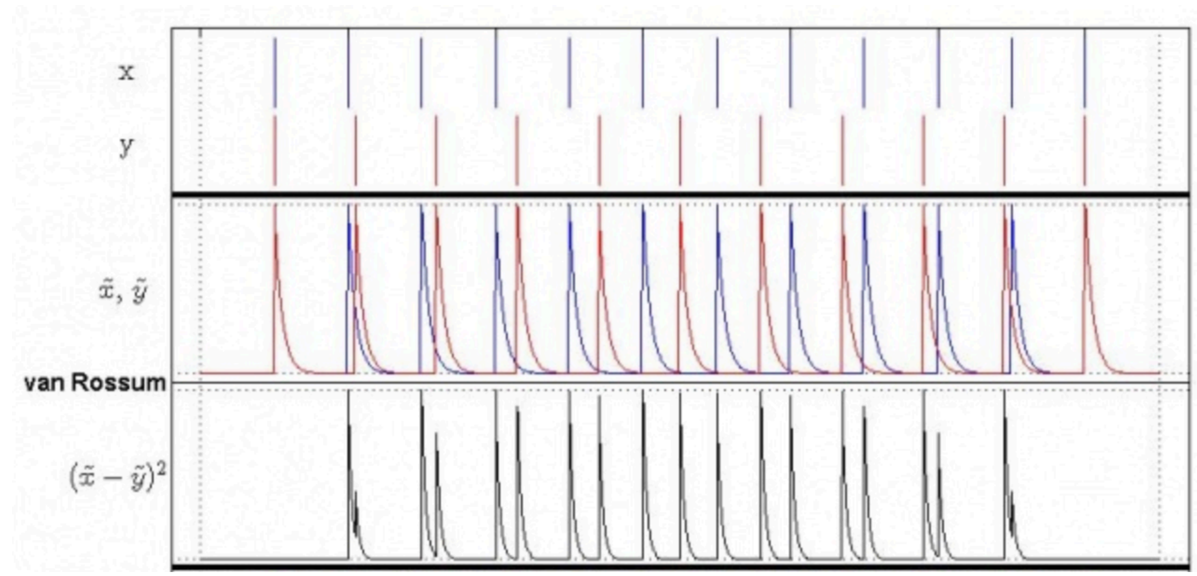
3. In reality, $F_t = \beta_1 c_t + \beta_0 + \epsilon_t$ or $F_t = c_t + \beta_{0,t} + \epsilon_t$

   Try a grid of intercept values via CV and pick the best fit

Almost there - metrics for evaluating spike sequences



Victor-Purpura distance



Van Rossum distance

Thanks!!!

Next steps:

1. Real data application
2. Wrap up the report