Labeled Optimal Partitioning

Toby Dylan Hocking
toby.hocking@nau.edu
joint work with Anuraag Srivastava

August 30, 2020

Introduction: supervised changepoint detection for cancer diagnosis with DNA copy number data

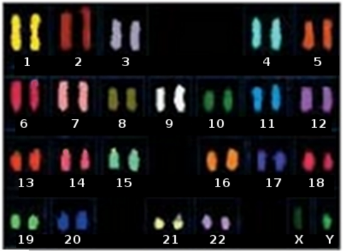
Labeled Optimal Partitioning (LOPART)

LOPART Demo

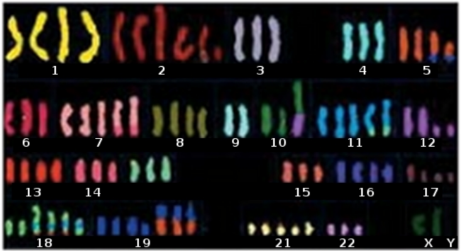
Results and Discussion

Cancer cells show chromosomal copy number alterations

Spectral karyotypes show the number of copies of the sex chromosomes (X,Y) and autosomes (1-22).
Source: Alberts *et al.* 2002.

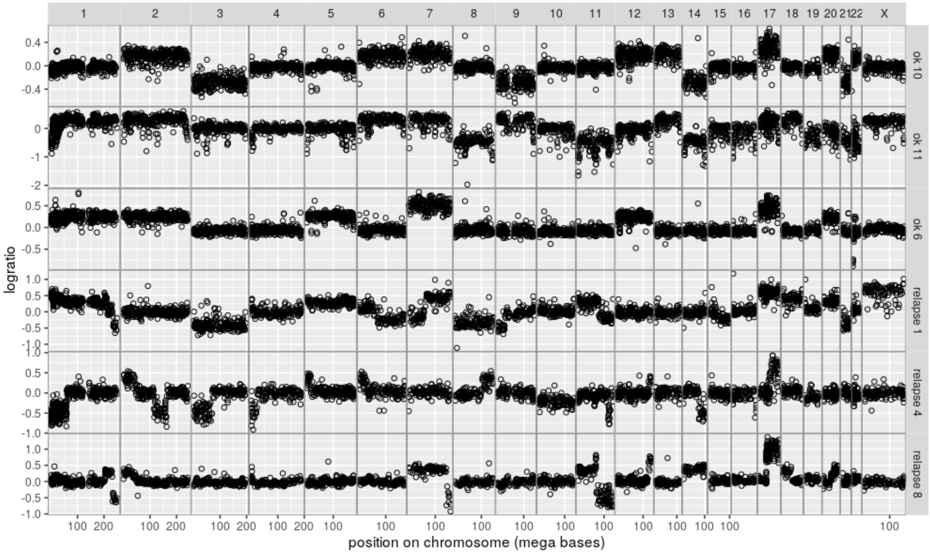


Normal cell with 2 copies of each autosome.

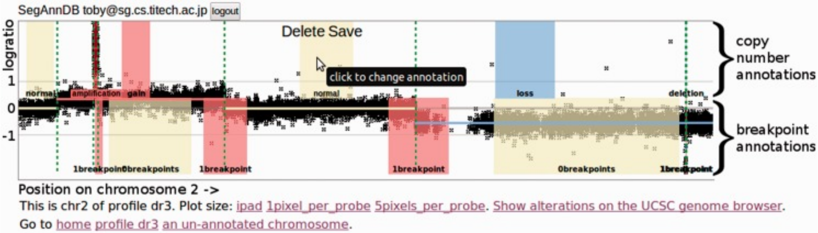


Cancer cell with many copy number alterations.

DNA copy number profiles from neuroblastoma patients with or without relapse



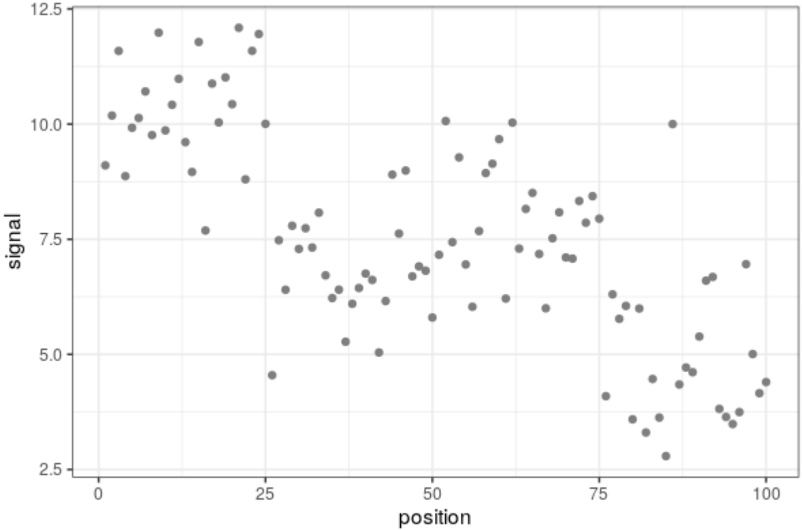
Previous work: SegAnnDB interactive machine learning system



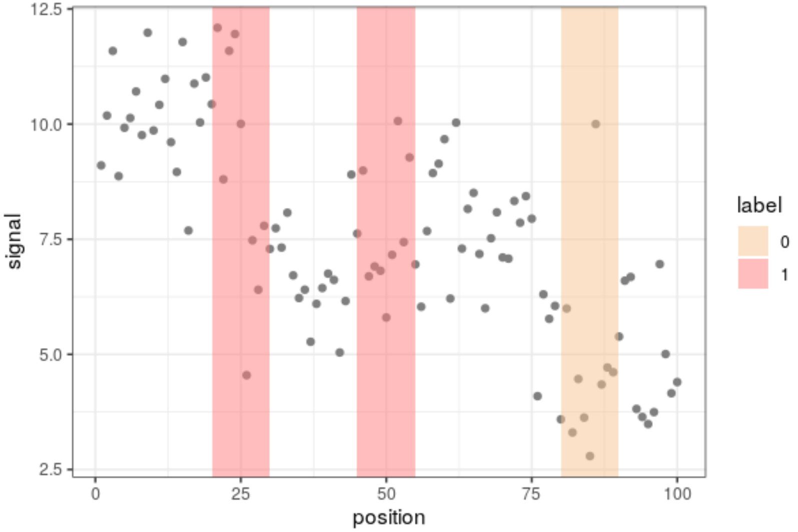
Hocking *et al.*, 2014.

- User uploads noisy data sets for machine learning analysis.
- User can provide labels which indicate presence(1) or absence(0) of changepoints in specific regions of data sets.
- Classic optimal changepoint model (max penalized Gaussian likelihood) used if it has zero label errors. OPART algorithm, Jackson *et al.*, 2005. FPOP algorithm, Maidstone *et al.*, 2016.
- Label-aware SegAnnot algorithm used otherwise (Hocking and Rigaill, 2012). Always zero train label errors, but never predicts any changepoints outside of positive(1) labels.

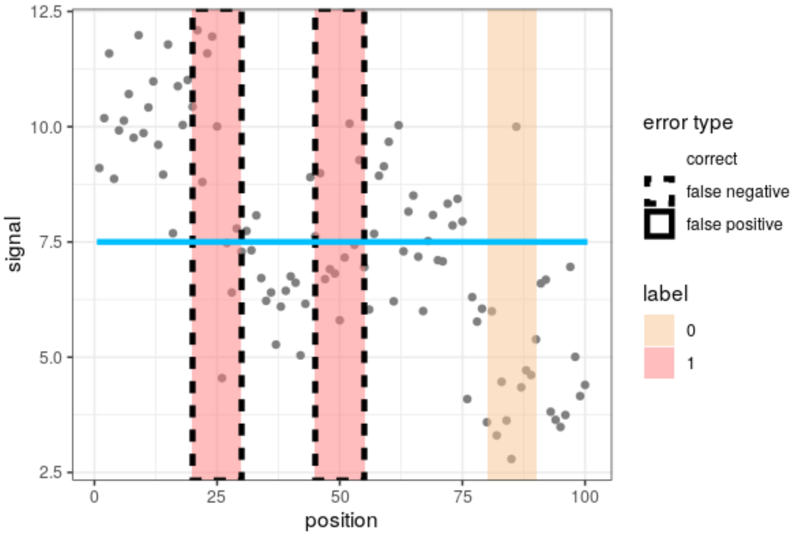
Example noisy data sequence



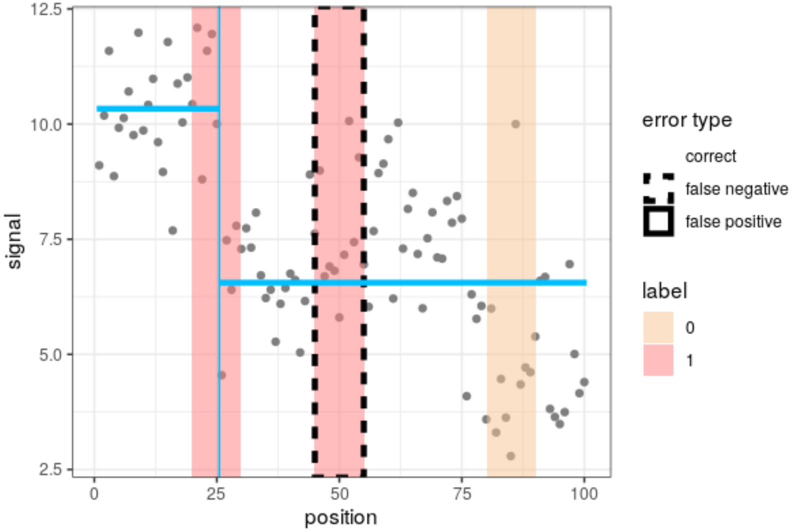
Example noisy data sequence with labels



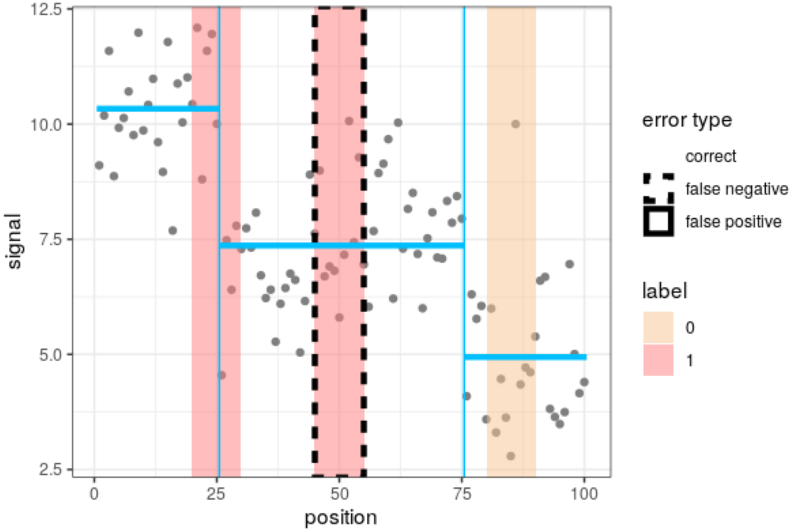
OPART with penalty $\lambda = 1000$ (ignores labels)



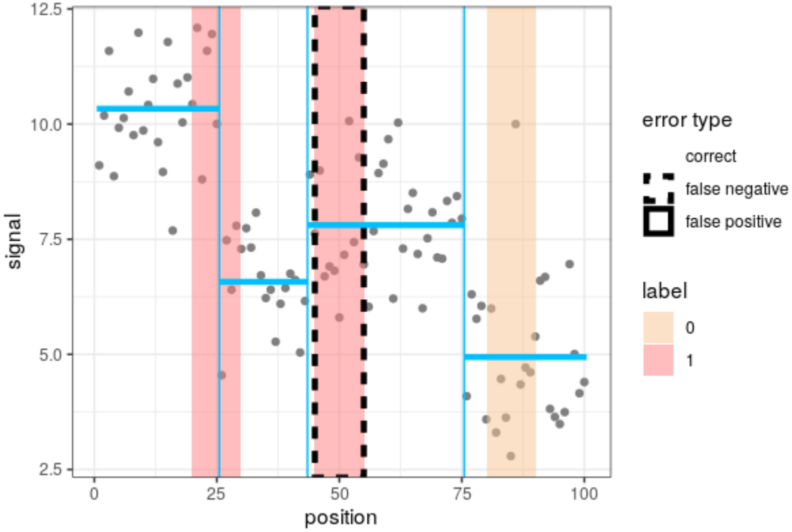
OPART with penalty $\lambda = 100$ (ignores labels)



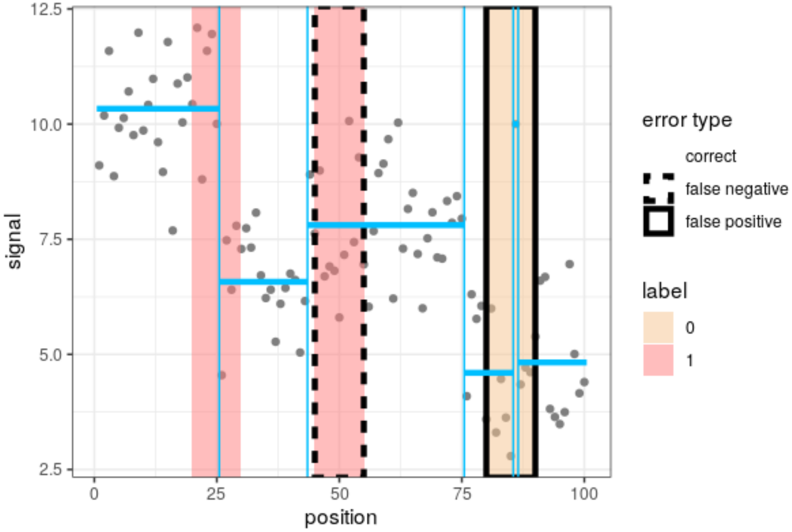
OPART with penalty $\lambda = 20$ (ignores labels)



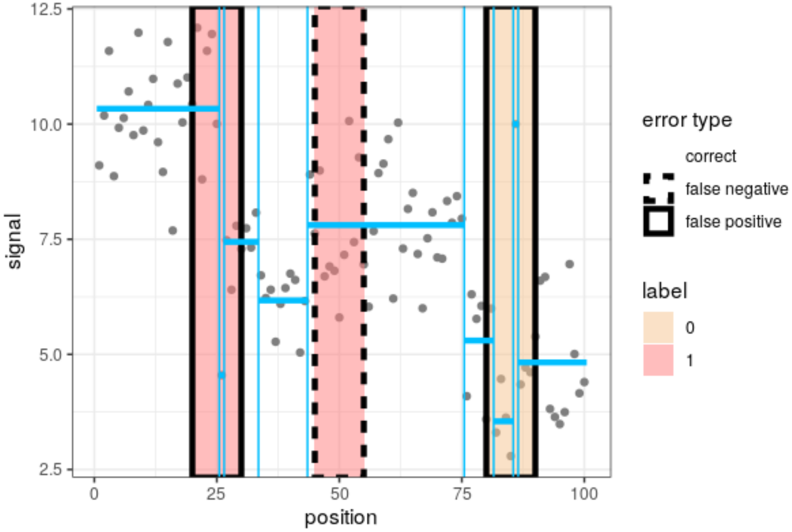
OPART with penalty $\lambda = 15$ (ignores labels)



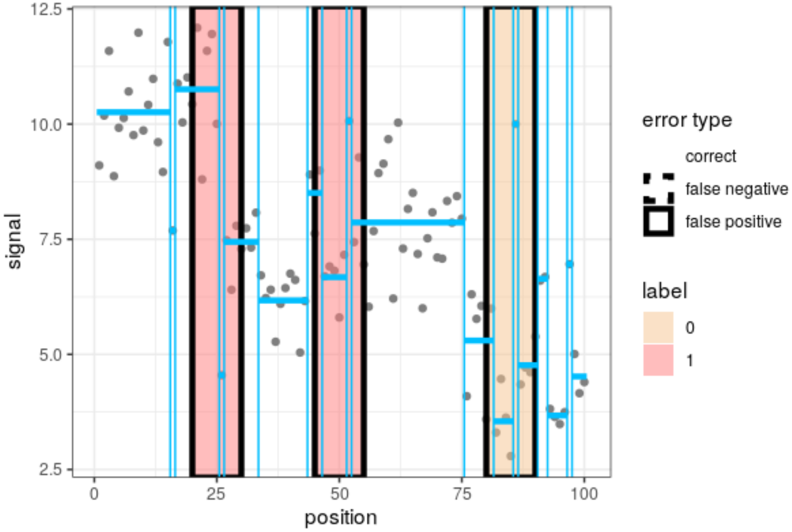
OPART with penalty $\lambda = 10$ (ignores labels)



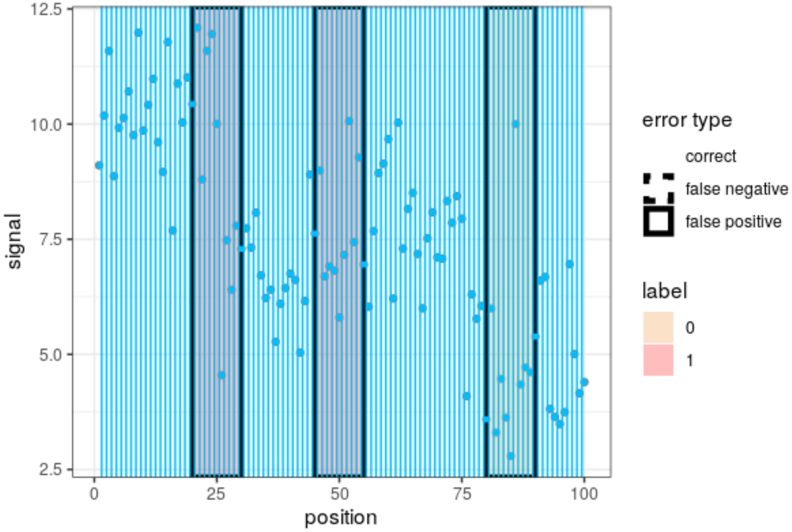
OPART with penalty $\lambda = 5$ (ignores labels)



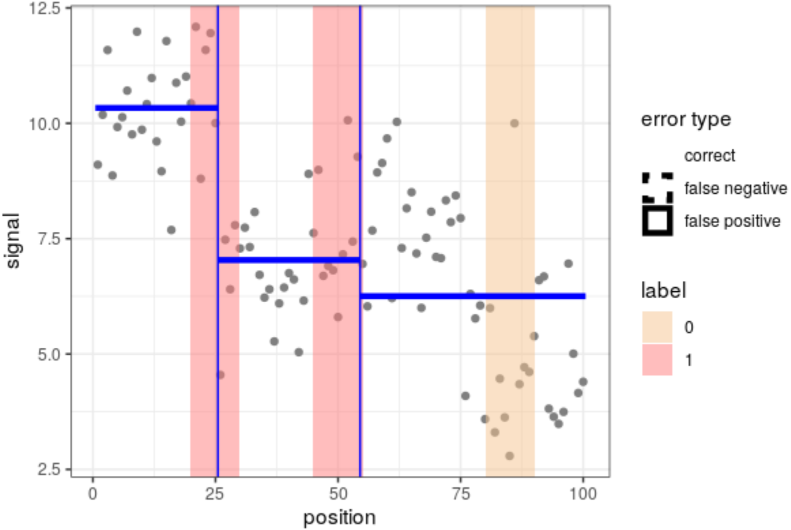
OPART with penalty $\lambda = 4$ (ignores labels)



OPART with penalty $\lambda = 0$ (ignores labels)



SegAnnot (no changes in unlabeled regions)



Introduction: supervised changepoint detection for cancer diagnosis with DNA copy number data

Labeled Optimal Partitioning (LOPART)

LOPART Demo

Results and Discussion

Baseline/previous OPART algorithm

Assume

- $\mathbf{x} = [x_1 \cdots x_N]$ is the sequence of N data,
- ℓ is a loss function (e.g. square loss),
- I the indicator function counts the number of changes,
- λ is a non-negative penalty (larger for fewer changes).

Then the problem and algorithm are

$$\begin{aligned} \hat{C}_N &= \min_{\mathbf{m}\in\mathbb{R}^N} \sum_{i=1}^{N} \ell(m_i, x_i) + \lambda \sum_{i=1}^{N-1} I[m_i \neq m_{i+1}]. \\ &= \min_{\tau\in\{0,1,\dots,N-1\}} \hat{C}_\tau + \lambda + L(\tau+1, N, \mathbf{x}). \end{aligned}$$

where

- τ is the last changepoint optimization variable,
- $\hat{C}_\tau$ is the optimal cost computed in previous iteration τ,
- L is the cost of the last segment.

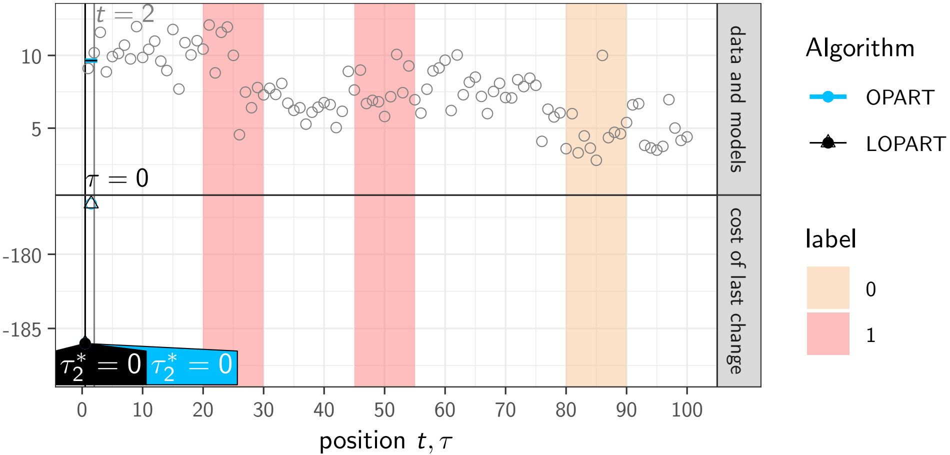
Introduction: supervised changepoint detection for cancer diagnosis with DNA copy number data

Labeled Optimal Partitioning (LOPART)

LOPART Demo

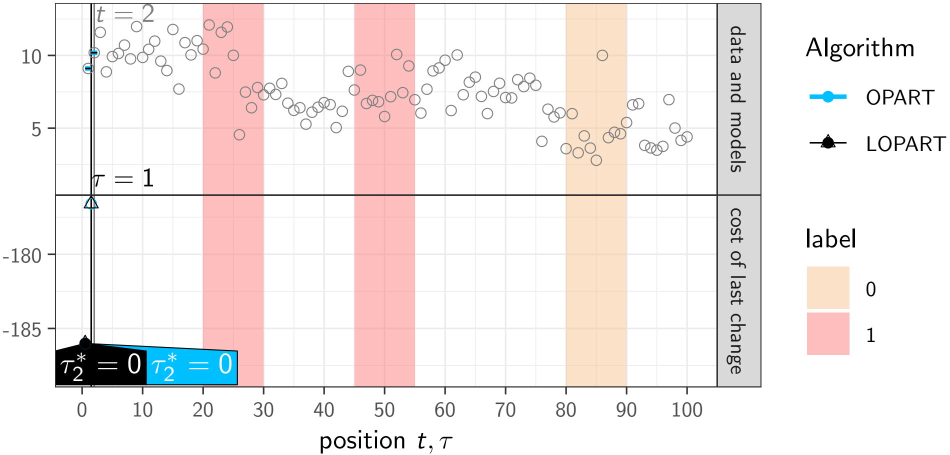
Results and Discussion

Demonstration of LOPART with penalty $\lambda = 10$



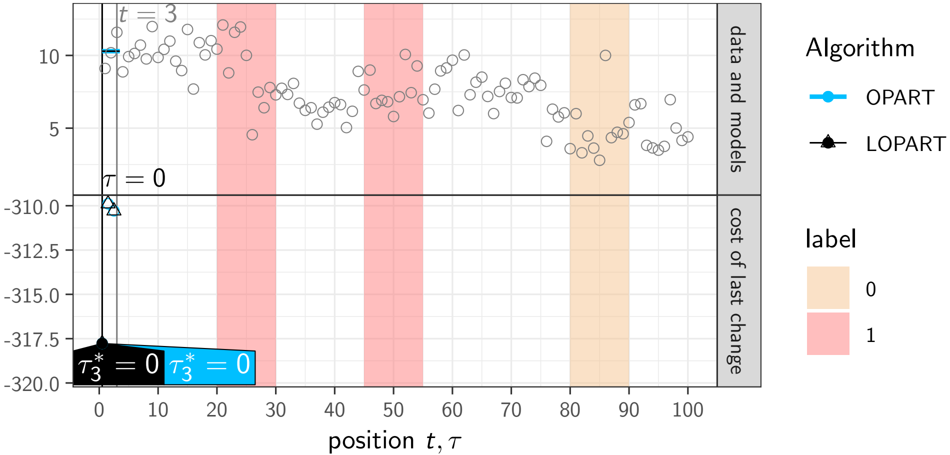
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



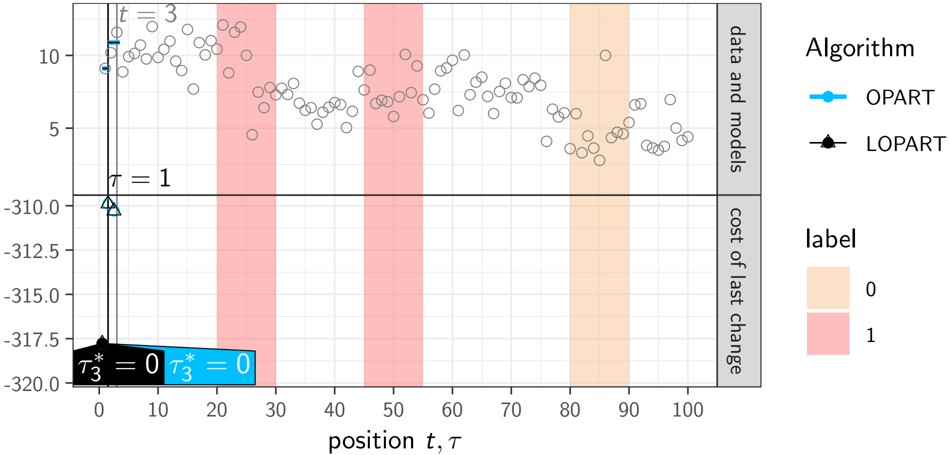
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



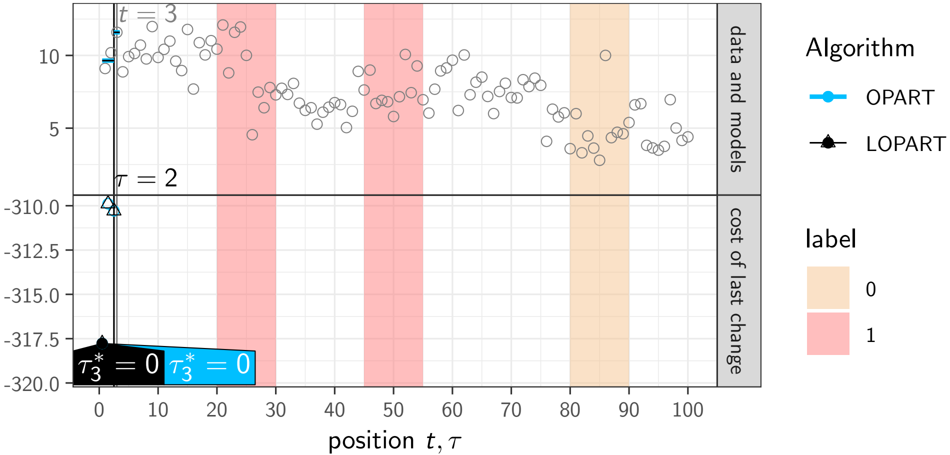
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



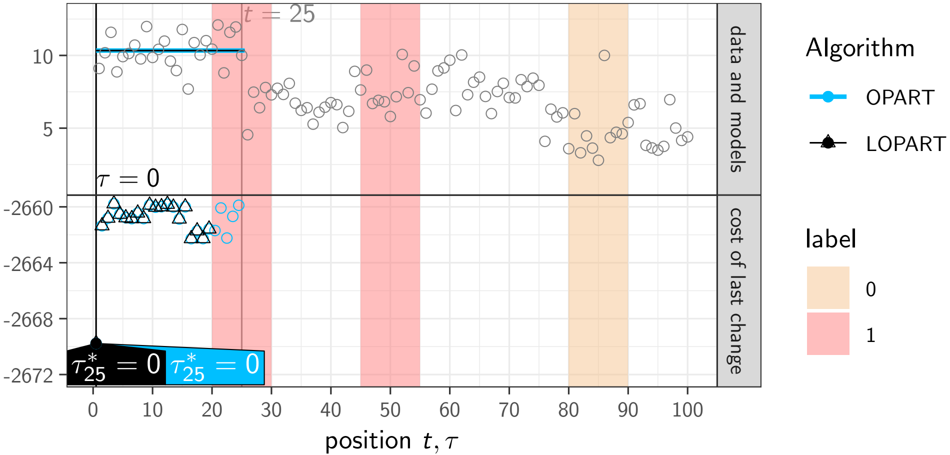
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



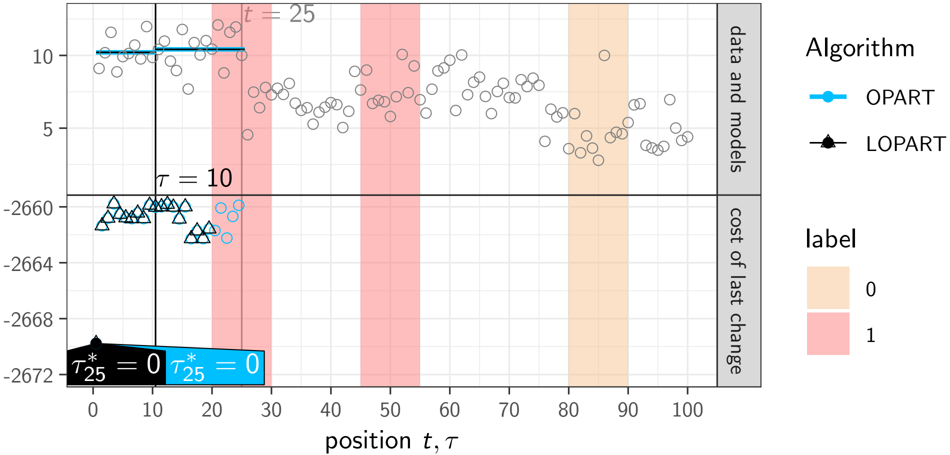
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



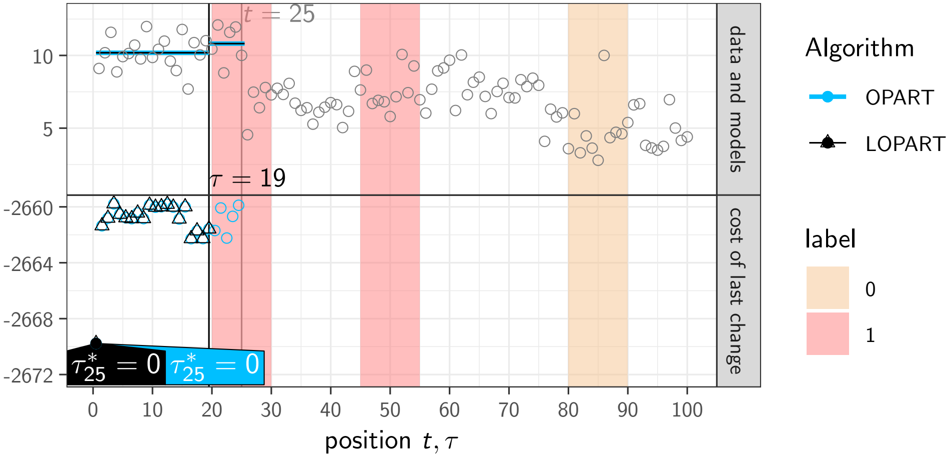
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



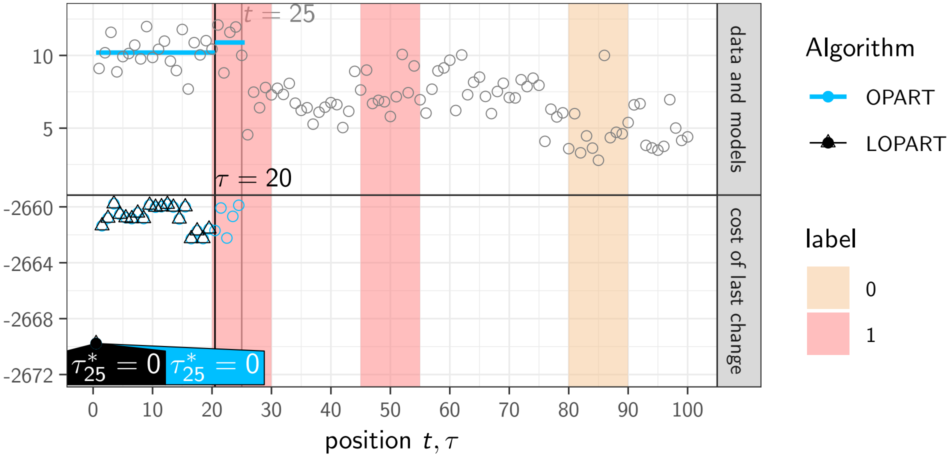
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



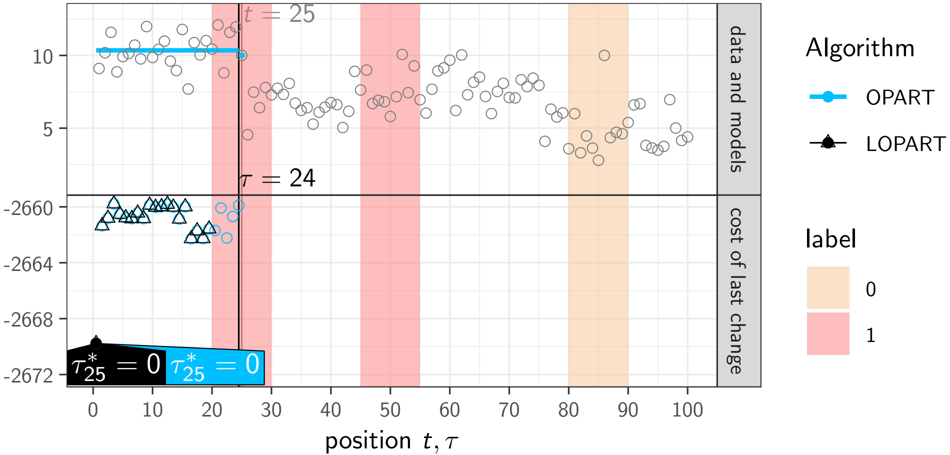
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



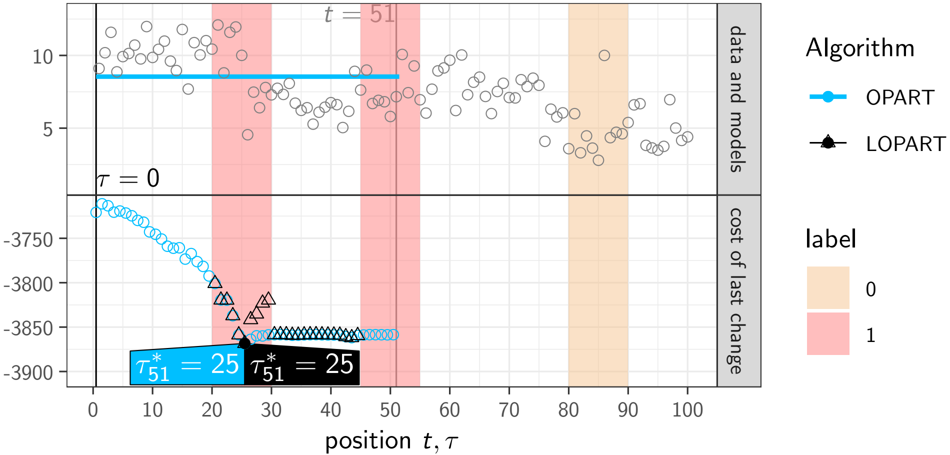
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



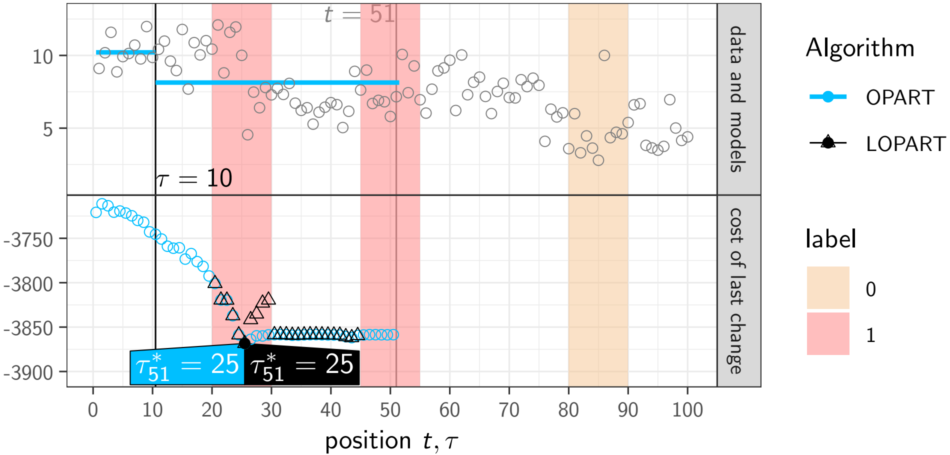
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



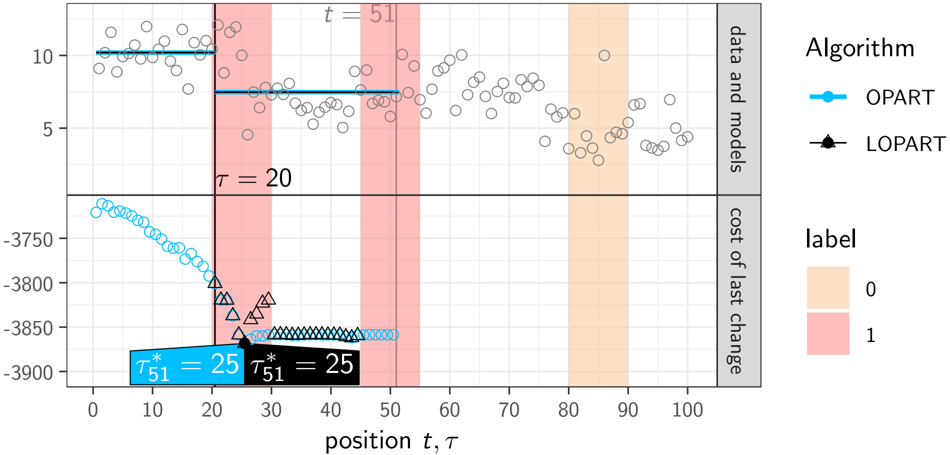
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



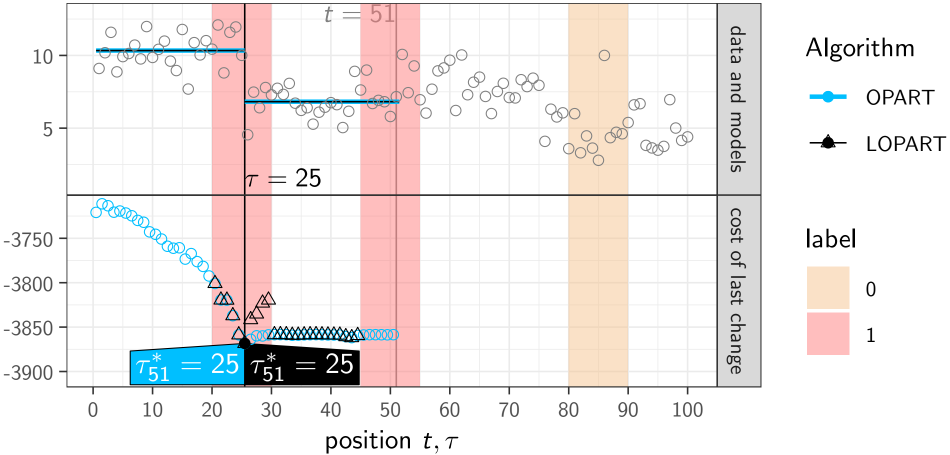
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



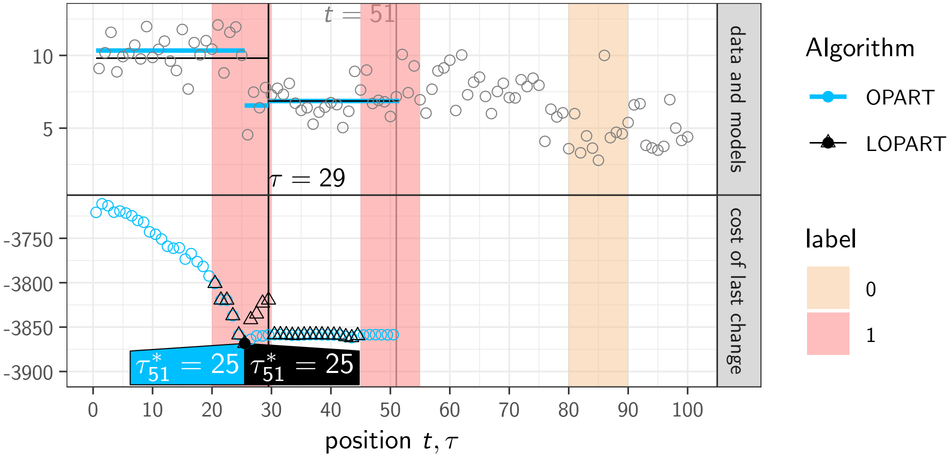
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



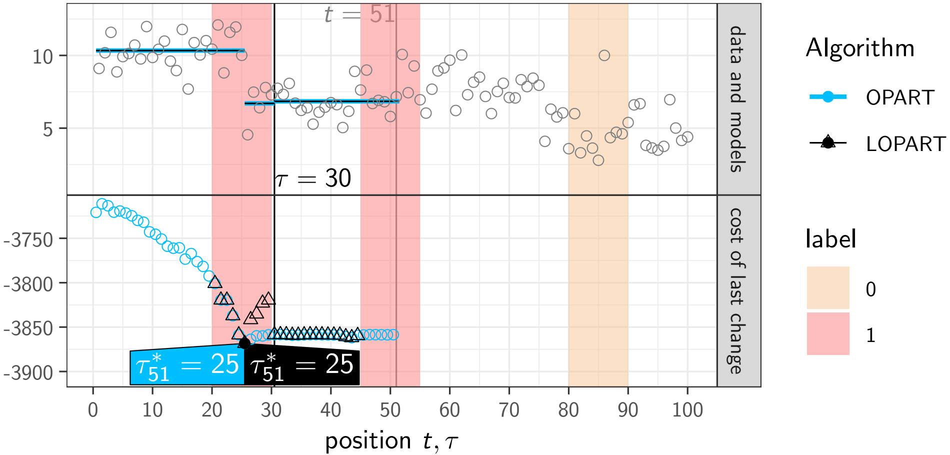
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



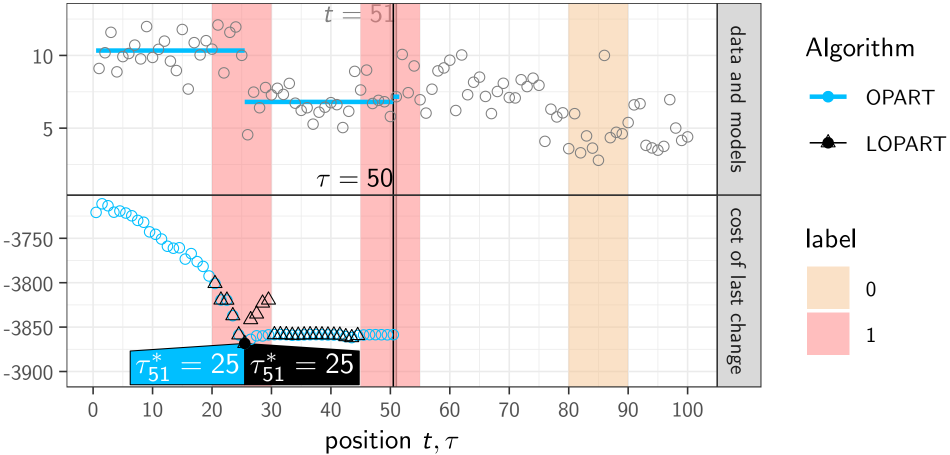
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



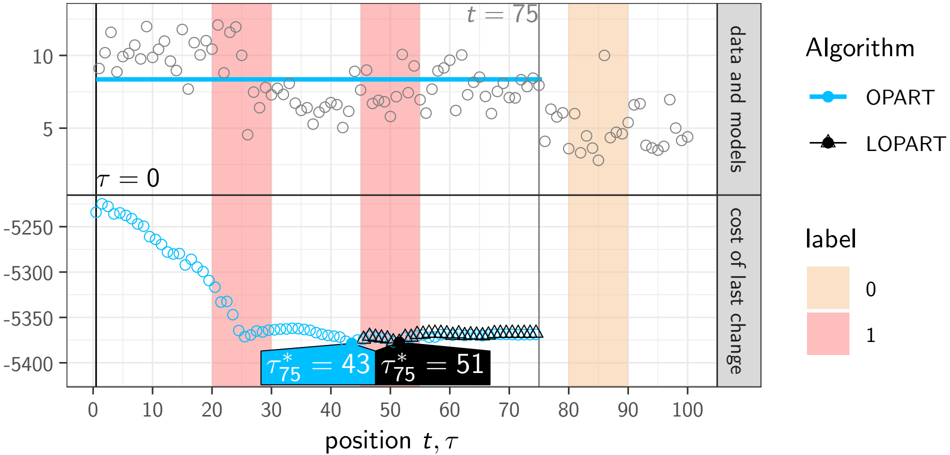
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



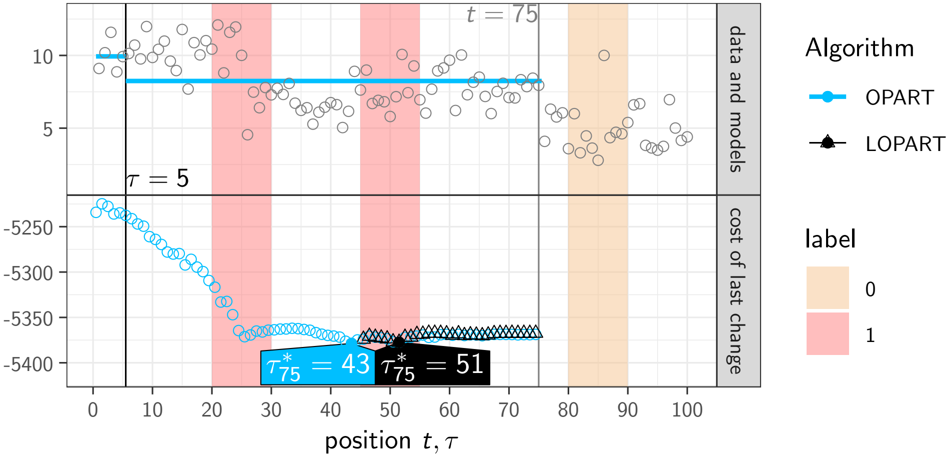
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



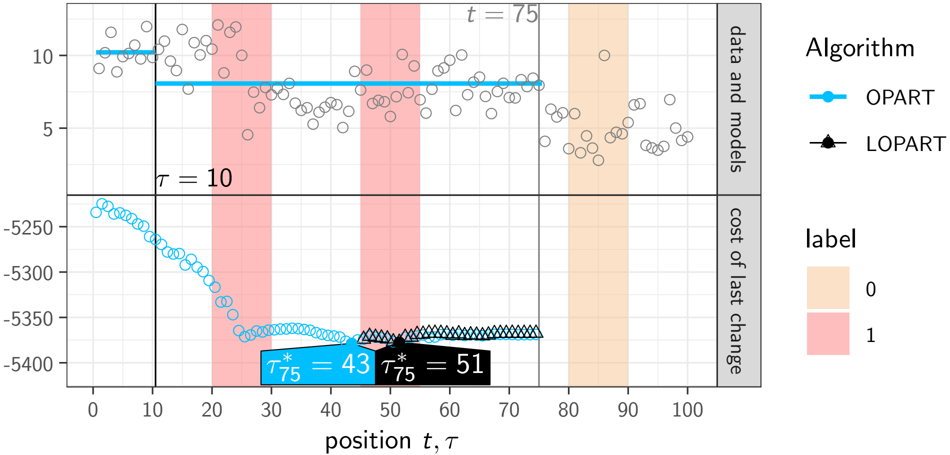
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



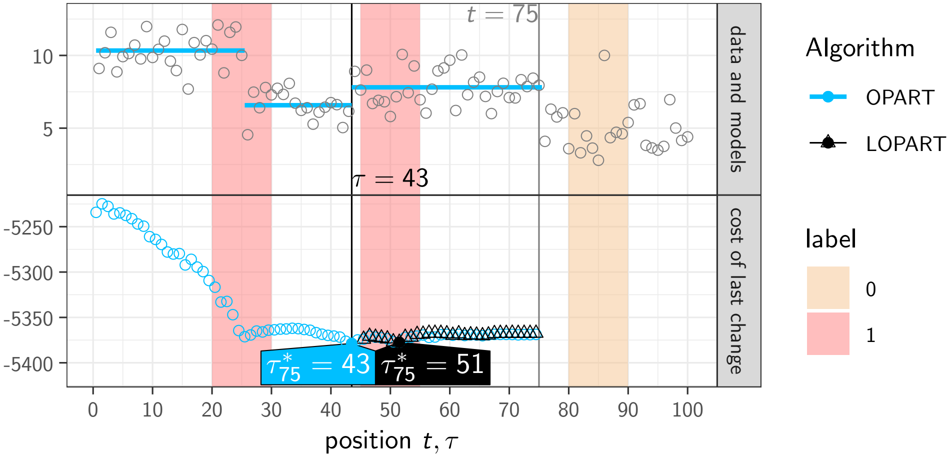
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



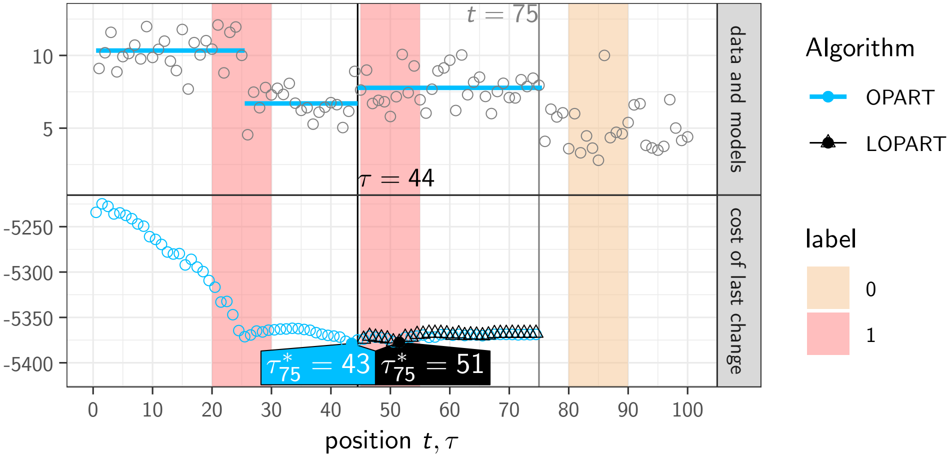
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



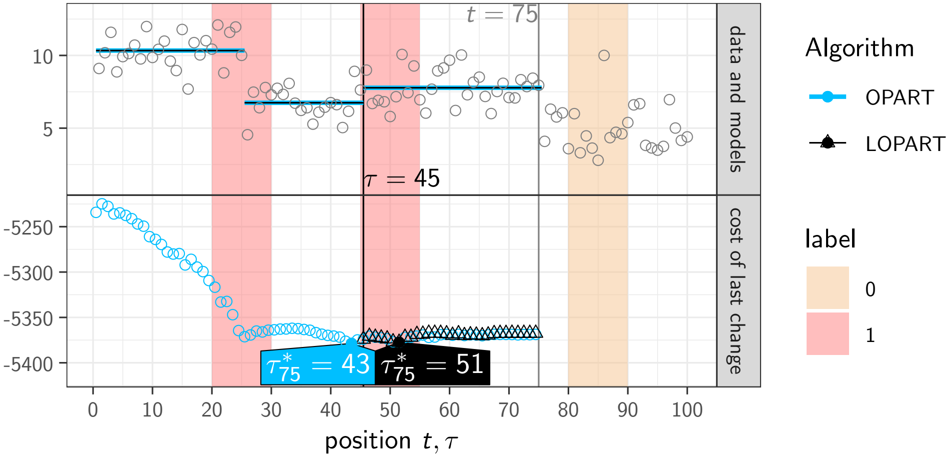
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



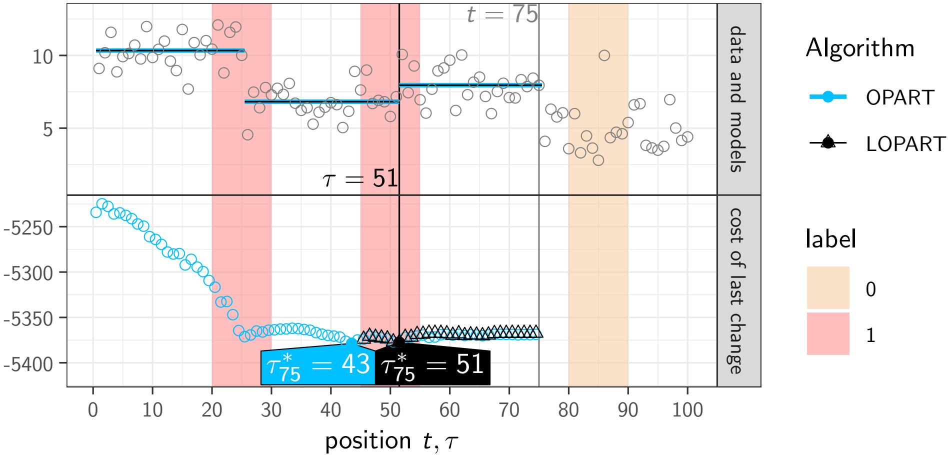
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



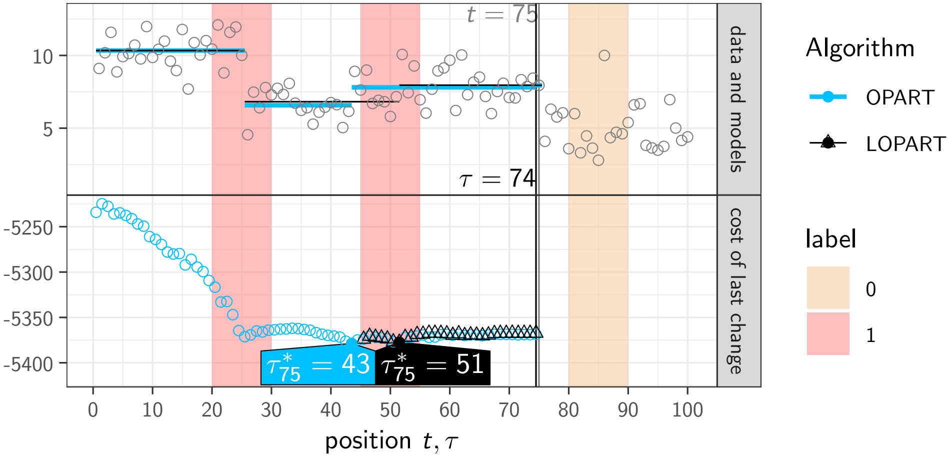
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



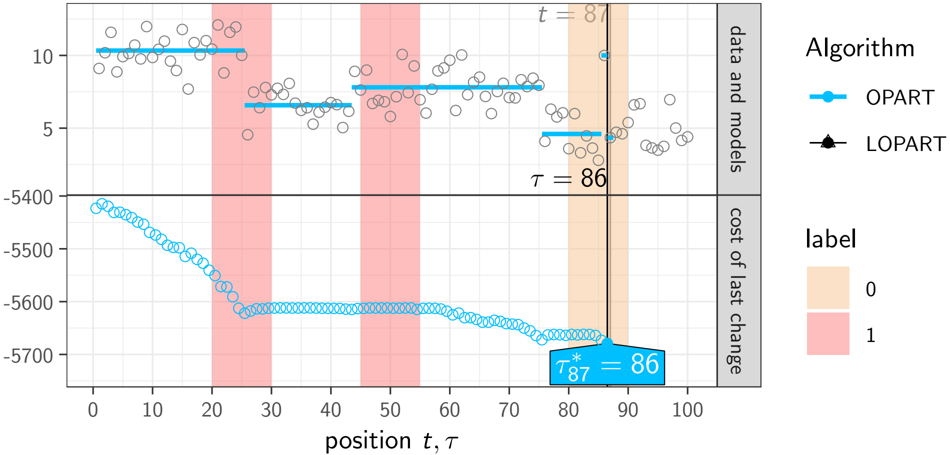
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



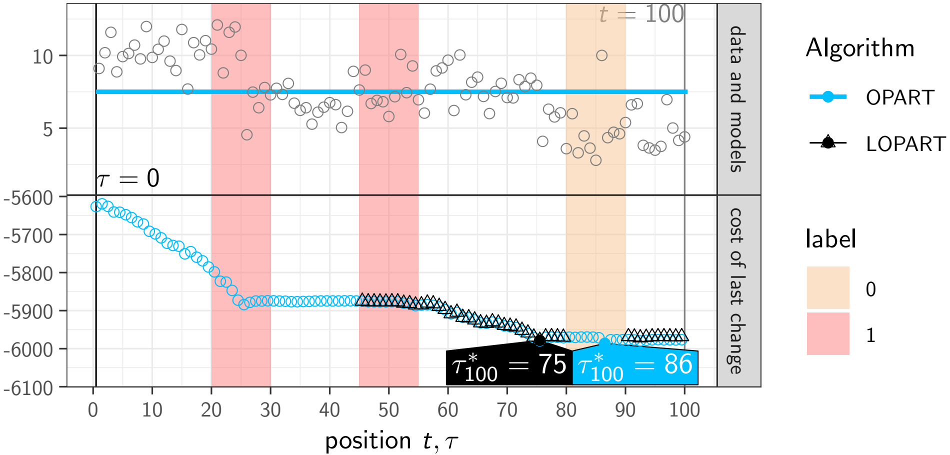
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



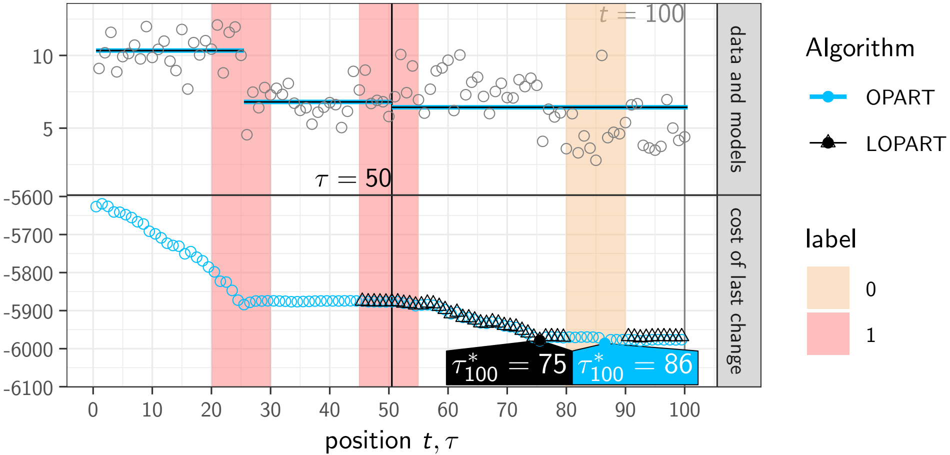
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



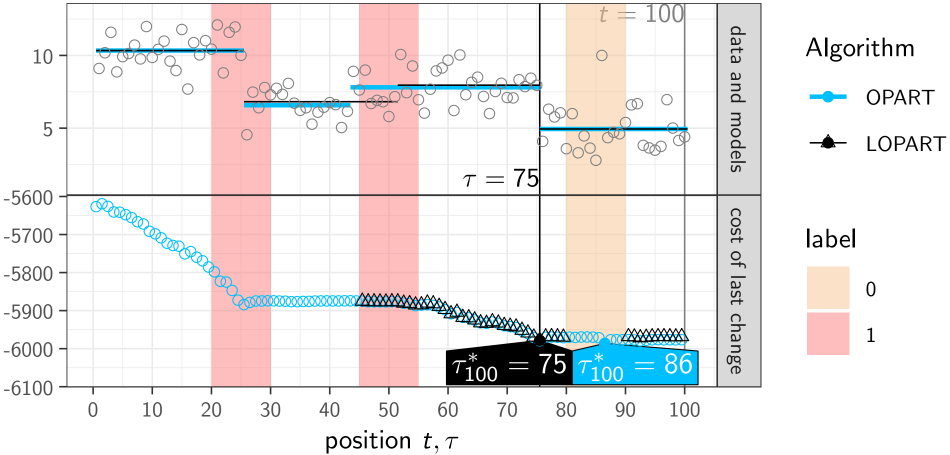
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



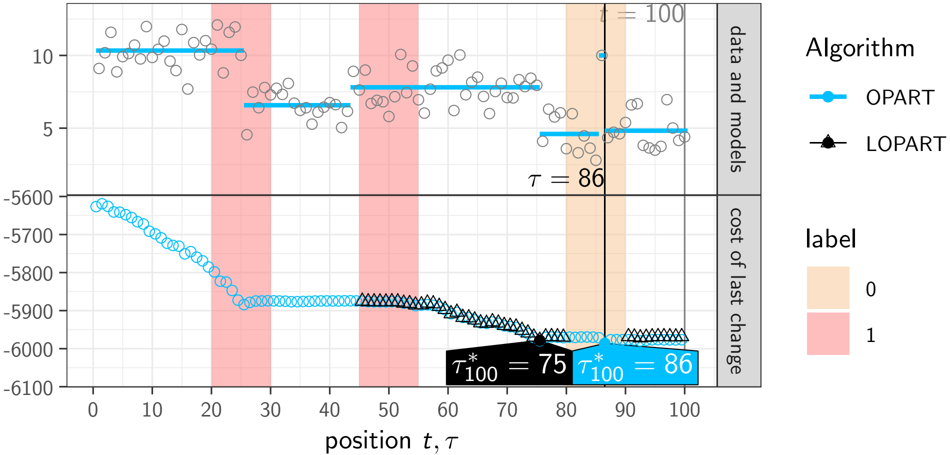
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau+1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



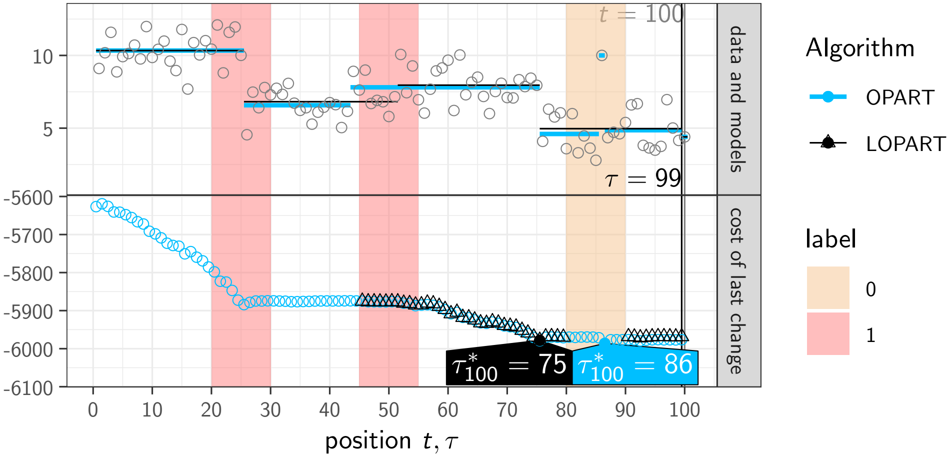
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$



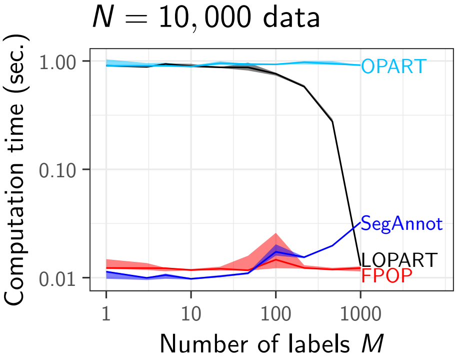
$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

Demonstration of LOPART with penalty $\lambda = 10$

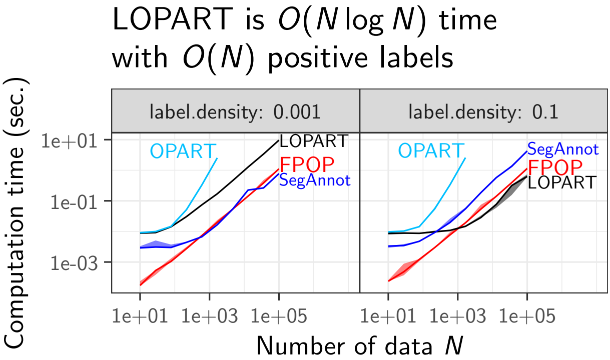


$$W_t = \min_{\tau \in T_t} \underbrace{W_\tau}_{\text{optimal cost up to } \tau} + \underbrace{\lambda}_{\text{penalty}} + \underbrace{L(\tau + 1, t, \mathbf{x})}_{\text{optimal cost of last segment}}$$

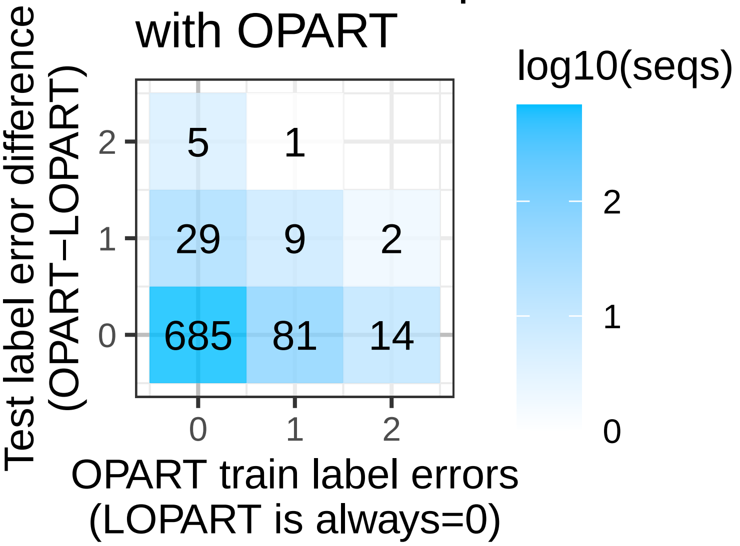
Empirical time complexity (labels)



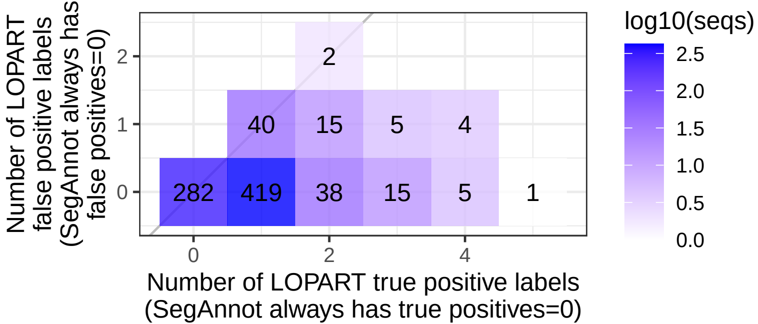
Empirical time complexity (data)



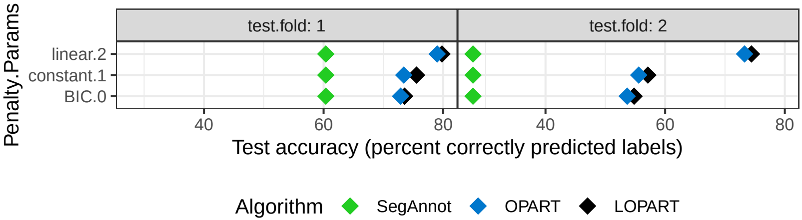
Best case comparison with OPART



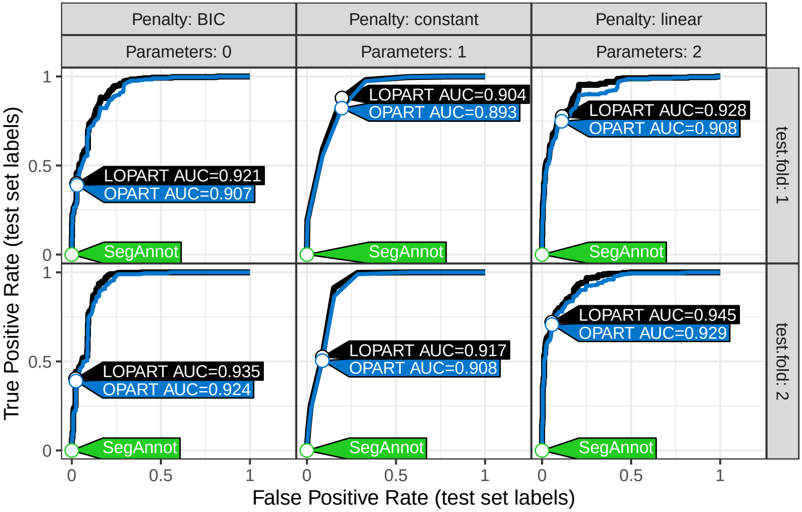
Best case comparison with SegAnnot



Test accuracy in cross-validation experiments



Test ROC curves in cross-validation experiments



Summary and Discussion

- Proposed algo fixes issues with two previous algorithms (better train AND test accuracy).
- Results demonstrate improved speed and accuracy.
- Future work: functional pruning algorithm, which can solve more complex constrained changepoint problems (e.g. change must be non-decreasing), and should be faster (log-linear instead of quadratic).