Input:

Data: X_{ik}, sequence $i = 1, 2, \ldots, N$ and position $k = 1, 2, \ldots, L_i$

Motif width: W

λ

(F) Regime types: $I = (I_1, I_2, \ldots, I_W)$, $I_w = (1 \text{ or } 2)$

(V) Uni- or bi-modal motif

Initialization:

(V) Sample I from the prior on uni- or bi-modal change points

Sample $\mathcal{P}$ (motif matrix): each column p_w picked from prior according to regime I_w

Algorithm: Repeat E- and M-step until convergence of $\mathcal{P}'$

E-step

Update Y_{ik}: posterior probabilities of position j being start site in sequence i

(V) Update c_{st}: posterior probabilities of change points (s, t)

M-step (Update $\mathcal{P}'$)

For each motif position w

Calculate n_{wj}: expected counts (based on $\mathcal{Y}$) of base j at position w

Order n_{wj} such that $n_{w(1)} \geq n_{w(2)} \geq n_{w(3)} \geq n_{w(4)}$

(V) Calculate d_w: posterior probability that $I_w = 1$ for position w

Calculate γ: Lagrange multiplier

(F) L1: solves quadratic function

L2: solves nonlinear monotone decreasing function of λ & $n_{w(j)}$'s

(V) L1: solves cubic function

. L2: solves nonlinear monotone decreasing function of λ, $n_{w(j)}$'s & d_w

Plug γ into equations for $p_{w(j)}$

(F) L1: equations (8) & (10) or L2: equations (12) & (13)

(V) L1: equation (14) or L2: equation (17)

Output:

Estimated motif matrix $\mathcal{P}'$

Final update of posterior probabilities of start positions $\mathcal{Y}'$

(V) Final update of posterior probabilities of change points $\mathcal{C}'$