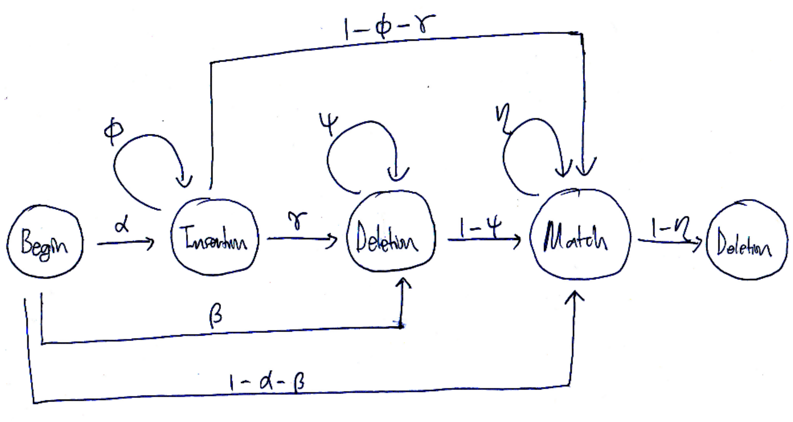
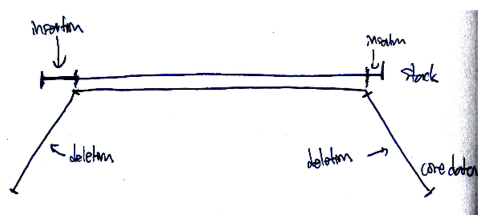
Supplementary Note: Detailed Computations

April 11th 2017

Taehee Lee

- K: number of cores
- T: number of discretized age indices in the stack
- N_i: number of discretized depth indices in the core i
- $\{A_t\}_{t=1}^{T}$: ages assigned to age indices in the stack
- $\{\mu_t\}_{t=1}^{T}$: means of $\delta^{18}O$ concentrations assigned to age indices in the stack
- $\{\sigma_t\}_{t=1}^{T}$: standard deviations of $\delta^{18}O$ concentrations assigned to age indices in the stack
- τ_i: shifts of $\delta^{18}O$ concentrations assigned to the core i
- $\left\{d_n^{(i)}\right\}_{n=1}^{N_i}$: depths assigned to depth indices in the core i
- $\left\{v_n^{(i)}\right\}_{n=1}^{N_i}$: $\delta^{18}O$ assigned to depth indices in the core i
- $\left\{\left(a_n^{(i)}, b_n^{(i)}\right)\right\}_{n=1}^{N_i}$: confidence intervals of ages inferred by radiocarbon data assigned to depth indices in the core i
- $\left\{C_n^{(i)}\right\}_{n=1}^{N_i}$: (hidden) ages assigned to depth indices in the core i
- R_i: average sedimentation rate in the core i
- $\alpha_i, \beta_i, \gamma_i, \phi_i, \psi_i, \eta_i, \delta_i$: transition parameters for the core i, where





1. Forward Algorithm:

1.1. Initialization:

1.1.1. For each $s \in \{1,2,\cdots,\mathrm{T}\}$,

$$\mathbf{F}_{\boldsymbol{i}}(\mathbf{2},\mathbf{1},\boldsymbol{s}) \triangleq p\left(v_{1:2}^{(i)}, \mathrm{C}_1^{(i)}=\mathrm{A}_1, \mathrm{C}_2^{(i)}=\mathrm{A}_s, 1 \text{ was not skipped} \,\middle|\, d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)$$

$$= p\left(v_2^{(i)} \middle| \mathrm{C}_2^{(i)}=\mathrm{A}_s\right) p\left(\mathrm{C}_2^{(i)}=\mathrm{A}_s \middle| \mathrm{C}_1^{(i)}=\mathrm{A}_1, d_1^{(i)}, d_2^{(i)}, a_2^{(i)}, b_2^{(i)}\right) p\left(v_1^{(i)} \middle| \mathrm{C}_1^{(i)}=\mathrm{A}_1\right) p\left(\mathrm{C}_1^{(i)}=\mathrm{A}_1 \middle| a_1^{(i)}, b_1^{(i)}\right) p(1 \text{ was not skipped})$$

$$= \begin{cases} \mathbf{N}\left(\boldsymbol{v}_2^{(i)} \middle| \boldsymbol{\tau}^{(i)}+\boldsymbol{\mu}_s, \boldsymbol{\sigma}_s\right)\cdot \mathbf{N}\left(\boldsymbol{v}_1^{(i)} \middle| \boldsymbol{\tau}^{(i)}+\boldsymbol{\mu}_1, \boldsymbol{\sigma}_1\right)\cdot \mathbf{1}_{\left\{\mathbf{A}_s \in \left(\boldsymbol{a}_2^{(i)}, \boldsymbol{b}_2^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathbf{A}_1 \in \left(\boldsymbol{a}_1^{(i)}, \boldsymbol{b}_1^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathbf{0.25} \le \frac{\mathbf{A}_s - \mathbf{A}_1}{\mathbf{R}_i\left(\boldsymbol{d}_2^{(i)} - \boldsymbol{d}_1^{(i)}\right)} \le \mathbf{4}\right\}} \cdot \boldsymbol{\rho}\left(\dfrac{\mathbf{A}_s - \mathbf{A}_1}{\mathbf{R}_i\left(\boldsymbol{d}_2^{(i)} - \boldsymbol{d}_1^{(i)}\right)}\right) \cdot (\mathbf{1} - \boldsymbol{\alpha}_i - \boldsymbol{\beta}_i), & \boldsymbol{s} > \mathbf{1} \\ \mathbf{N}\left(\boldsymbol{v}_2^{(i)} \middle| \boldsymbol{\tau}^{(i)}+\boldsymbol{\mu}_1, \boldsymbol{\sigma}_1\right)\cdot \mathbf{N}\left(\boldsymbol{v}_1^{(i)} \middle| \boldsymbol{\tau}^{(i)}+\boldsymbol{\mu}_1, \boldsymbol{\sigma}_1\right)\cdot \mathbf{1}_{\left\{\mathbf{A}_1 \in \left(\boldsymbol{a}_2^{(i)}, \boldsymbol{b}_2^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathbf{A}_1 \in \left(\boldsymbol{a}_1^{(i)}, \boldsymbol{b}_1^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathbf{0.125} \le \mathbf{R}_i\left(\boldsymbol{d}_2^{(i)} - \boldsymbol{d}_1^{(i)}\right) \le \mathbf{2}\right\}} \cdot \boldsymbol{\rho}\left(\dfrac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_2^{(i)} - \boldsymbol{d}_1^{(i)}\right)}\right) \cdot (\mathbf{1} - \boldsymbol{\alpha}_i - \boldsymbol{\beta}_i), & \boldsymbol{s} = \mathbf{1} \end{cases}$$

1.1.2. For each $n \in \{3,4,\cdots,\mathrm{N}_i\}$ and $s \in \{1,2,\cdots,\mathrm{T}\}$,

$$\mathbf{F}_{\boldsymbol{i}}(\boldsymbol{n},\mathbf{1},\boldsymbol{s}) \triangleq p\left(v_{1:n}^{(i)}, \mathrm{C}_{n-1}^{(i)}=\mathrm{A}_1, \mathrm{C}_n^{(i)}=\mathrm{A}_s, 1{:}\,(n-2) \text{ were not aligned} \,\middle|\, d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)$$

$$= p\left(v_n^{(i)} \middle| \mathrm{C}_n^{(i)}=\mathrm{A}_s\right) p\left(\mathrm{C}_n^{(i)}=\mathrm{A}_s \middle| \mathrm{C}_{n-1}^{(i)}=\mathrm{A}_1, d_{n-1}^{(i)}, d_n^{(i)}, a_n^{(i)}, b_n^{(i)}\right) p\left(v_{n-1}^{(i)} \middle| \mathrm{C}_{n-1}^{(i)}=\mathrm{A}_1\right) p\left(\mathrm{C}_{n-1}^{(i)}=\mathrm{A}_1 \middle| a_{n-1}^{(i)}, b_{n-1}^{(i)}\right) p(1{:}\,(n-2) \text{ were not aligned})$$

$$= \begin{cases} \mathbf{N}\left(\boldsymbol{v}_n^{(i)} \middle| \boldsymbol{\tau}^{(i)}+\boldsymbol{\mu}_s, \boldsymbol{\sigma}_s\right)\cdot \mathbf{N}\left(\boldsymbol{v}_{n-1}^{(i)} \middle| \boldsymbol{\tau}^{(i)}+\boldsymbol{\mu}_1, \boldsymbol{\sigma}_1\right)\cdot \mathbf{1}_{\left\{\mathbf{A}_s \in \left(\boldsymbol{a}_n^{(i)}, \boldsymbol{b}_n^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathbf{A}_1 \in \left(\boldsymbol{a}_{n-1}^{(i)}, \boldsymbol{b}_{n-1}^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathbf{0.25} \le \frac{\mathbf{A}_s - \mathbf{A}_1}{\mathbf{R}_i\left(\boldsymbol{d}_n^{(i)} - \boldsymbol{d}_{n-1}^{(i)}\right)} \le \mathbf{4}\right\}} \cdot \boldsymbol{\rho}\left(\dfrac{\mathbf{A}_s - \mathbf{A}_1}{\mathbf{R}_i\left(\boldsymbol{d}_n^{(i)} - \boldsymbol{d}_{n-1}^{(i)}\right)}\right) \cdot \boldsymbol{\beta}_i \boldsymbol{\psi}_i^{n-2} \boldsymbol{\delta}_i^{n-2} \cdot (\mathbf{1} - \boldsymbol{\psi}_i), & \boldsymbol{s} > \mathbf{1} \\ \mathbf{N}\left(\boldsymbol{v}_n^{(i)} \middle| \boldsymbol{\tau}^{(i)}+\boldsymbol{\mu}_1, \boldsymbol{\sigma}_1\right)\cdot \mathbf{N}\left(\boldsymbol{v}_{n-1}^{(i)} \middle| \boldsymbol{\tau}^{(i)}+\boldsymbol{\mu}_1, \boldsymbol{\sigma}_1\right)\cdot \mathbf{1}_{\left\{\mathbf{A}_1 \in \left(\boldsymbol{a}_n^{(i)}, \boldsymbol{b}_n^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathbf{A}_1 \in \left(\boldsymbol{a}_{n-1}^{(i)}, \boldsymbol{b}_{n-1}^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathbf{0.125} \le \mathbf{R}_i\left(\boldsymbol{d}_n^{(i)} - \boldsymbol{d}_{n-1}^{(i)}\right) \le \mathbf{2}\right\}} \cdot \boldsymbol{\rho}\left(\dfrac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_n^{(i)} - \boldsymbol{d}_{n-1}^{(i)}\right)}\right) \cdot \boldsymbol{\beta}_i \boldsymbol{\psi}_i^{n-2} \boldsymbol{\delta}_i^{n-2} \cdot (\mathbf{1} - \boldsymbol{\psi}_i), & \boldsymbol{s} = \mathbf{1} \end{cases}$$

1.1.3. For each $t, s \in \{2,3,\cdots,\mathrm{T}\}$,

$$\mathbf{F}_{\boldsymbol{i}}(\mathbf{2},\boldsymbol{t},\boldsymbol{s}) \triangleq p\left(v_{1:2}^{(i)}, \mathrm{C}_1^{(i)}=\mathrm{A}_t, \mathrm{C}_2^{(i)}=\mathrm{A}_s, 1{:}\,(t-1) \text{ were skipped}, t \text{ was not skipped} \,\middle|\, d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)$$

$$= p\left(v_2^{(i)} \middle| \mathrm{C}_2^{(i)}=\mathrm{A}_s\right) p\left(\mathrm{C}_2^{(i)}=\mathrm{A}_s \middle| \mathrm{C}_1^{(i)}=\mathrm{A}_t, d_1^{(i)}, d_2^{(i)}, a_2^{(i)}, b_2^{(i)}\right) p\left(v_1^{(i)} \middle| \mathrm{C}_1^{(i)}=\mathrm{A}_t\right) p\left(\mathrm{C}_1^{(i)}=\mathrm{A}_t \middle| a_1^{(i)}, b_1^{(i)}\right) p(1{:}\,(t-1) \text{ were skipped}, t \text{ was not skipped})$$

$$
= \begin{cases} \mathbf{N}\left(\boldsymbol{v}_2^{(i)} \middle| \boldsymbol{\tau}^{(i)} + \boldsymbol{\mu}_s, \boldsymbol{\sigma}_s\right) \cdot \mathbf{N}\left(\boldsymbol{v}_1^{(i)} \middle| \boldsymbol{\tau}^{(i)} + \boldsymbol{\mu}_t, \boldsymbol{\sigma}_t\right) \cdot \mathbf{1}_{\left\{\mathrm{A}_s \in \left(a_2^{(i)}, b_2^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathrm{A}_t \in \left(a_1^{(i)}, b_1^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{0.25 \le \frac{\mathrm{A}_s - \mathrm{A}_t}{\mathrm{R}_i\left(d_2^{(i)} - d_1^{(i)}\right)} \le 4\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{A}_s - \mathbf{A}_t}{\mathbf{R}_i\left(\boldsymbol{d}_2^{(i)} - \boldsymbol{d}_1^{(i)}\right)}\right) \cdot \boldsymbol{\alpha}_i \boldsymbol{\phi}_i^{t-1} \cdot (\mathbf{1} - \boldsymbol{\phi}_i - \boldsymbol{\gamma}_i) \cdot (\mathbf{1} - \boldsymbol{\phi}_i), & \boldsymbol{s} > \boldsymbol{t} \\ \mathbf{N}\left(\boldsymbol{v}_2^{(i)} \middle| \boldsymbol{\tau}^{(i)} + \boldsymbol{\mu}_s, \boldsymbol{\sigma}_s\right) \cdot \mathbf{N}\left(\boldsymbol{v}_1^{(i)} \middle| \boldsymbol{\tau}^{(i)} + \boldsymbol{\mu}_s, \boldsymbol{\sigma}_s\right) \cdot \mathbf{1}_{\left\{\mathrm{A}_s \in \left(a_2^{(i)}, b_2^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathrm{A}_s \in \left(a_1^{(i)}, b_1^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{0.125 \le \mathrm{R}_i\left(d_2^{(i)} - d_1^{(i)}\right) \le 2\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_2^{(i)} - \boldsymbol{d}_1^{(i)}\right)}\right) \cdot \boldsymbol{\alpha}_i \boldsymbol{\phi}_i^{s-1} \cdot (\mathbf{1} - \boldsymbol{\phi}_i - \boldsymbol{\gamma}_i) \cdot (\mathbf{1} - \boldsymbol{\phi}_i), & \boldsymbol{s} = \boldsymbol{t} \end{cases}
$$

1.2. Iterations: for each $n \in \{3,4,\cdots,\mathrm{N}_i\}$ and $t, s \in \{2,3,\cdots,\mathrm{T}\}$,

$$
\mathbf{F}_{\boldsymbol{i}}(\boldsymbol{n}, \boldsymbol{t}, \boldsymbol{s}) \triangleq p\left(v_{1:n}^{(i)}, \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t, \mathrm{C}_n^{(i)} = \mathrm{A}_s \middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)
$$

$$
= p\left(v_{1:n}^{(i)}, \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t, \mathrm{C}_n^{(i)} = \mathrm{A}_s, \text{staying in the match state} \middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right) + p\left(v_{1:n}^{(i)}, \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t, \mathrm{C}_n^{(i)} = \mathrm{A}_s, 1{:}\,(n-2) \text{ were not aligned} \middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)
$$

$$
p\left(v_{1:n}^{(i)}, \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t, \mathrm{C}_n^{(i)} = \mathrm{A}_s, \text{staying in the match state} \middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)
$$

$$
= \sum_{u=1}^{\mathrm{T}} p\left(v_n^{(i)} \middle| \mathrm{C}_n^{(i)} = \mathrm{A}_s\right) p\left(\mathrm{C}_n^{(i)} = \mathrm{A}_s \middle| \mathrm{C}_{n-2}^{(i)} = \mathrm{A}_u, \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t, d_{n-2}^{(i)}, d_{n-1}^{(i)}, d_n^{(i)}, a_n^{(i)}, b_n^{(i)}\right) p\left(v_{1:n-1}^{(i)}, \mathrm{C}_{n-2}^{(i)} = \mathrm{A}_u, \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t \middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right) p(\text{staying in the match state})
$$

$$
= \begin{cases} \boldsymbol{\eta}_i \cdot \mathbf{N}\left(\boldsymbol{v}_n^{(i)} \middle| \boldsymbol{\tau}^{(i)} + \boldsymbol{\mu}_s, \boldsymbol{\sigma}_s\right) \cdot \mathbf{1}_{\left\{\mathrm{A}_s \in \left(a_n^{(i)}, b_n^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{0.25 \le \frac{\mathrm{A}_s - \mathrm{A}_t}{\mathrm{R}_i\left(d_n^{(i)} - d_{n-1}^{(i)}\right)} \le 4\right\}} \left[\mathbf{1}_{\left\{0.125 \le \mathrm{R}_i\left(d_{n-1}^{(i)} - d_{n-2}^{(i)}\right) \le 2\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_{n-1}^{(i)} - \boldsymbol{d}_{n-2}^{(i)}\right)}, \frac{\mathbf{A}_s - \mathbf{A}_t}{\mathbf{R}_i\left(\boldsymbol{d}_n^{(i)} - \boldsymbol{d}_{n-1}^{(i)}\right)}\right) \cdot \mathbf{F}_i(\boldsymbol{n}-\mathbf{1}, \boldsymbol{t}, \boldsymbol{t}) + \sum_{u \neq t} \left(\mathbf{1}_{\left\{0.25 \le \frac{\mathrm{A}_t - \mathrm{A}_u}{\mathrm{R}_i\left(d_{n-1}^{(i)} - d_{n-2}^{(i)}\right)} \le 4\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{A}_t - \mathbf{A}_u}{\mathbf{R}_i\left(\boldsymbol{d}_{n-1}^{(i)} - \boldsymbol{d}_{n-2}^{(i)}\right)}, \frac{\mathbf{A}_s - \mathbf{A}_t}{\mathbf{R}_i\left(\boldsymbol{d}_n^{(i)} - \boldsymbol{d}_{n-1}^{(i)}\right)}\right) \cdot \mathbf{F}_i(\boldsymbol{n}-\mathbf{1}, \boldsymbol{u}, \boldsymbol{t}) \right) \right], & \mathbf{s} > \mathbf{t} \\ \boldsymbol{\eta}_i \cdot \mathbf{N}\left(\boldsymbol{v}_n^{(i)} \middle| \boldsymbol{\tau}^{(i)} + \boldsymbol{\mu}_s, \boldsymbol{\sigma}_s\right) \cdot \mathbf{1}_{\left\{\mathrm{A}_s \in \left(a_n^{(i)}, b_n^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{0.125 \le \mathrm{R}_i\left(d_n^{(i)} - d_{n-1}^{(i)}\right) \le 2\right\}} \left[\mathbf{1}_{\left\{0.125 \le \mathrm{R}_i\left(d_{n-1}^{(i)} - d_{n-2}^{(i)}\right) \le 2\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_{n-1}^{(i)} - \boldsymbol{d}_{n-2}^{(i)}\right)}, \frac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_n^{(i)} - \boldsymbol{d}_{n-1}^{(i)}\right)}\right) \cdot \mathbf{F}_i(\boldsymbol{n}-\mathbf{1}, \boldsymbol{s}, \boldsymbol{s}) + \sum_{u \neq s} \left(\mathbf{1}_{\left\{0.25 \le \frac{\mathrm{A}_s - \mathrm{A}_u}{\mathrm{R}_i\left(d_{n-1}^{(i)} - d_{n-2}^{(i)}\right)} \le 4\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{A}_s - \mathbf{A}_u}{\mathbf{R}_i\left(\boldsymbol{d}_{n-1}^{(i)} - \boldsymbol{d}_{n-2}^{(i)}\right)}, \frac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_n^{(i)} - \boldsymbol{d}_{n-1}^{(i)}\right)}\right) \cdot \mathbf{F}_i(\boldsymbol{n}-\mathbf{1}, \boldsymbol{u}, \boldsymbol{s}) \right) \right], & \mathbf{s} = \mathbf{t} \end{cases}
$$

$$
p\left(v_{1:n}^{(i)}, \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t, \mathrm{C}_n^{(i)} = \mathrm{A}_s, 1{:}\,(n-2) \text{ were not aligned} \middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)
$$

$$
= p\left(v_n^{(i)} \middle| \mathrm{C}_n^{(i)} = \mathrm{A}_s\right) p\left(\mathrm{C}_n^{(i)} = \mathrm{A}_s \middle| \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t, d_{n-1}^{(i)}, d_n^{(i)}, a_n^{(i)}, b_n^{(i)}\right) p\left(v_{n-1}^{(i)} \middle| \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t\right) p\left(\mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t \middle| a_{n-1}^{(i)}, b_{n-1}^{(i)}\right) p(1{:}\,(n-2) \text{ were not aligned})
$$

$$
= \begin{cases} \mathbf{N}\left(\boldsymbol{v}_n^{(i)} \middle| \boldsymbol{\tau}^{(i)} + \boldsymbol{\mu}_s, \boldsymbol{\sigma}_s\right) \cdot \mathbf{N}\left(\boldsymbol{v}_{n-1}^{(i)} \middle| \boldsymbol{\tau}^{(i)} + \boldsymbol{\mu}_t, \boldsymbol{\sigma}_t\right) \cdot \mathbf{1}_{\left\{\mathrm{A}_s \in \left(a_n^{(i)}, b_n^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathrm{A}_t \in \left(a_{n-1}^{(i)}, b_{n-1}^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{0.25 \le \frac{\mathrm{A}_s - \mathrm{A}_t}{\mathrm{R}_i\left(d_n^{(i)} - d_{n-1}^{(i)}\right)} \le 4\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{A}_s - \mathbf{A}_t}{\mathbf{R}_i\left(\boldsymbol{d}_n^{(i)} - \boldsymbol{d}_{n-1}^{(i)}\right)}\right) \cdot \boldsymbol{\alpha}_i \boldsymbol{\phi}_i^{t-1} \cdot \boldsymbol{\gamma}_i \boldsymbol{\psi}_i^{n-2} \boldsymbol{\delta}_i^{n-2} \cdot (\mathbf{1} - \boldsymbol{\psi}_i), & \boldsymbol{s} > \boldsymbol{t} \\ \mathbf{N}\left(\boldsymbol{v}_n^{(i)} \middle| \boldsymbol{\tau}^{(i)} + \boldsymbol{\mu}_s, \boldsymbol{\sigma}_s\right) \cdot \mathbf{N}\left(\boldsymbol{v}_{n-1}^{(i)} \middle| \boldsymbol{\tau}^{(i)} + \boldsymbol{\mu}_s, \boldsymbol{\sigma}_s\right) \cdot \mathbf{1}_{\left\{\mathrm{A}_s \in \left(a_n^{(i)}, b_n^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{\mathrm{A}_s \in \left(a_{n-1}^{(i)}, b_{n-1}^{(i)}\right)\right\}} \cdot \mathbf{1}_{\left\{0.125 \le \mathrm{R}_i\left(d_n^{(i)} - d_{n-1}^{(i)}\right) \le 2\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_n^{(i)} - \boldsymbol{d}_{n-1}^{(i)}\right)}\right) \cdot \boldsymbol{\alpha}_i \boldsymbol{\phi}_i^{t-1} \cdot \boldsymbol{\gamma}_i \boldsymbol{\psi}_i^{n-2} \boldsymbol{\delta}_i^{n-2} \cdot (\mathbf{1} - \boldsymbol{\psi}_i), & \boldsymbol{s} = \boldsymbol{t} \end{cases}
$$

1.3. Termination:

1.3.1. for each $n \in \{2,3,\cdots,\mathrm{N}_i - 1\}$ and $t,s \in \{1,2,\cdots,\mathrm{T}\}$,

$$\mathbf{E}_{\boldsymbol{i}}(\boldsymbol{n},\boldsymbol{t},\boldsymbol{s}) \triangleq p\left(v_{1:\mathrm{N}_i}^{(i)}, \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t, \mathrm{C}_n^{(i)} = \mathrm{A}_s, (n+1){:}\,\mathrm{N}_i \text{ were not aligned}\middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)$$

$$= p\left(v_{n+1:\mathrm{N}_i}^{(i)}, (n+1){:}\,\mathrm{N}_i \text{ were not aligned}\right) p\left(v_{1:n}^{(i)}, \mathrm{C}_{n-1}^{(i)} = \mathrm{A}_t, \mathrm{C}_n^{(i)} = \mathrm{A}_s\middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right) = (\mathbf{1} - \boldsymbol{\eta}_{\boldsymbol{i}}) \cdot \boldsymbol{\delta}_{\boldsymbol{i}}^{\mathrm{N}_i - \boldsymbol{n}} \cdot \mathbf{F}_{\boldsymbol{i}}(\boldsymbol{n},\boldsymbol{t},\boldsymbol{s})$$

1.3.2. for each $t,s \in \{1,2,\cdots,\mathrm{T}\}$,

$$\mathbf{E}_{\boldsymbol{i}}(\mathbf{N}_{\boldsymbol{i}},\boldsymbol{t},\boldsymbol{s}) \triangleq p\left(v_{1:\mathrm{N}_i}^{(i)}, \mathrm{C}_{\mathrm{N}_i-1}^{(i)} = \mathrm{A}_t, \mathrm{C}_{\mathrm{N}_i}^{(i)} = \mathrm{A}_s\middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right) = \mathbf{F}_{\boldsymbol{i}}(\mathbf{N}_{\boldsymbol{i}},\boldsymbol{t},\boldsymbol{s})$$

2. Backward Sampling:

2.1. Initialization: for each $m \in \{2,3,\cdots,\mathrm{N}_i\}$ and $t_{m-1}^{(i),(l)}, t_m^{(i),(l)} \in \{1,2,3,\cdots,\mathrm{T}\}$,

$$\mathbb{P}\left(\mathbf{C}_{\boldsymbol{m-1}}^{(\boldsymbol{i}),(\boldsymbol{l})} = \mathbf{A}_{t_{m-1}^{(i),(l)}}, \mathbf{C}_{\boldsymbol{m}}^{(\boldsymbol{i}),(\boldsymbol{l})} = \mathbf{A}_{t_m^{(i),(l)}}\middle| \boldsymbol{v}_{\mathbf{1}:\mathbf{N}_{\boldsymbol{i}}}^{(\boldsymbol{i})}, \boldsymbol{a}_{\mathbf{1}:\mathbf{N}_{\boldsymbol{i}}}^{(\boldsymbol{i})}, \boldsymbol{b}_{\mathbf{1}:\mathbf{N}_{\boldsymbol{i}}}^{(\boldsymbol{i})}, \boldsymbol{d}_{\mathbf{1}:\mathbf{N}_{\boldsymbol{i}}}^{(\boldsymbol{i})}\right)$$

$$\propto p\left(v_{1:\mathrm{N}_i}^{(i)}, \mathrm{C}_{m-1}^{(i),(l)} = \mathrm{A}_{t_{m-1}^{(i),(l)}}, \mathrm{C}_m^{(i),(l)} = \mathrm{A}_{t_m^{(i),(l)}}, (m+1){:}\,\mathrm{N}_i \text{ were not aligned}\middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)$$

$$= \mathbf{E}_{\boldsymbol{i}}\left(\boldsymbol{m}, \boldsymbol{t}_{\boldsymbol{m-1}}^{(\boldsymbol{i}),(\boldsymbol{l})}, \boldsymbol{t}_{\boldsymbol{m}}^{(\boldsymbol{i}),(\boldsymbol{l})}\right)$$

2.2. Iterations: for each $n \in \{m-2, m-3,\cdots,1\}$ and $t_n^{(i),(l)} \in \{1,2,3,\cdots,\mathrm{T}\}$,

$$\mathbb{P}\left(\mathbf{C}_{\boldsymbol{n}}^{(\boldsymbol{i}),(\boldsymbol{l})} = \mathbf{A}_{\boldsymbol{t}_{\boldsymbol{n}}^{(i),(l)}}\middle| \mathbf{C}_{\boldsymbol{n+1}}^{(\boldsymbol{i}),(\boldsymbol{l})} = \mathbf{A}_{\boldsymbol{t}_{\boldsymbol{n+1}}^{(i),(l)}}, \mathbf{C}_{\boldsymbol{n+2}}^{(\boldsymbol{i}),(\boldsymbol{l})} = \mathbf{A}_{\boldsymbol{t}_{\boldsymbol{n+2}}^{(i),(l)}}, \boldsymbol{v}_{\mathbf{1}:\boldsymbol{n+2}}^{(\boldsymbol{i})}, \boldsymbol{a}_{\mathbf{1}:\mathbf{N}_{\boldsymbol{i}}}^{(\boldsymbol{i})}, \boldsymbol{b}_{\mathbf{1}:\mathbf{N}_{\boldsymbol{i}}}^{(\boldsymbol{i})}, \boldsymbol{d}_{\mathbf{1}:\mathbf{N}_{\boldsymbol{i}}}^{(\boldsymbol{i})}\right)$$

$$\propto \mathbb{P}\left(\mathrm{C}_n^{(i),(l)} = \mathrm{A}_{t_n^{(i),(l)}}, \mathrm{C}_{n+1}^{(i),(l)} = \mathrm{A}_{t_{n+1}^{(i),(l)}}, \mathrm{C}_{n+2}^{(i),(l)} = \mathrm{A}_{t_{n+2}^{(i),(l)}}, v_{1:n+2}^{(i)}\middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)$$

$$= p\left(v_{n+2}^{(i)}\middle| \mathrm{C}_{n+2}^{(i),(l)} = \mathrm{A}_{t_{n+2}^{(i),(l)}}\right) p\left(\mathrm{C}_{n+2}^{(i),(l)} = \mathrm{A}_{t_{n+2}^{(i),(l)}}\middle| \mathrm{C}_n^{(i),(l)} = \mathrm{A}_{t_n^{(i),(l)}}, \mathrm{C}_{n+1}^{(i),(l)} = \mathrm{A}_{t_{n+1}^{(i),(l)}}, d_n^{(i)}, d_{n+1}^{(i)}, d_{n+2}^{(i)}, a_{n+2}^{(i)}, b_{n+2}^{(i)}\right) p\left(\mathrm{C}_n^{(i),(l)} = \mathrm{A}_{t_n^{(i),(l)}}, \mathrm{C}_{n+1}^{(i),(l)} = \mathrm{A}_{t_{n+1}^{(i),(l)}}, v_{1:n+1}^{(i)}\middle| d_{1:\mathrm{N}_i}^{(i)}, a_{1:\mathrm{N}_i}^{(i)}, b_{1:\mathrm{N}_i}^{(i)}\right)$$

$$= \mathrm{N}\left(v_{n+2}^{(i)} \middle| \tau^i + \mu_{t_{n+2}^{(i),(l)}}, \sigma_{t_{n+2}^{(i),(l)}}\right) 1_{\left\{\mathrm{A}_{t_{n+2}^{(i),(l)}} \in \left(a_{n+2}^{(i)}, b_{n+2}^{(i)}\right)\right\}} \rho\left(\frac{\mathrm{A}_{t_{n+1}^{(i),(l)}} - \mathrm{A}_{t_n^{(i),(l)}}}{\mathrm{R}_i\left(d_{n+1}^{(i)} - d_n^{(i)}\right)}, \frac{\mathrm{A}_{t_{n+2}^{(i),(l)}} - \mathrm{A}_{t_{n+1}^{(i),(l)}}}{\mathrm{R}_i\left(d_{n+2}^{(i)} - d_{n+1}^{(i)}\right)}\right) \mathrm{F}_i\left(n+1, t_n^{(i),(l)}, t_{n+1}^{(i),(l)}\right)$$

2.2.1. $t_{n+1}^{(i),(l)} < t_{n+2}^{(i),(l)}$:

$$= \begin{cases} \mathrm{N}\left(v_{n+2}^{(i)} \middle| \tau^i + \mu_{t_{n+2}^{(i),(l)}}, \sigma_{t_{n+2}^{(i),(l)}}\right) \cdot 1_{\left\{\mathrm{A}_{t_{n+2}^{(i),(l)}} \in \left(a_{n+2}^{(i)}, b_{n+2}^{(i)}\right)\right\}} \cdot 1_{\left\{0.25 \le \frac{\mathrm{A}_{t_{n+1}^{(i),(l)}} - \mathrm{A}_{t_n^{(i),(l)}}}{\mathrm{R}_i\left(d_{n+1}^{(i)} - d_n^{(i)}\right)} \le 4\right\}} \cdot \rho\left(\frac{\mathrm{A}_{t_{n+1}^{(i),(l)}} - \mathrm{A}_{t_n^{(i),(l)}}}{\mathrm{R}_i\left(d_{n+1}^{(i)} - d_n^{(i)}\right)}, \frac{\mathrm{A}_{t_{n+2}^{(i),(l)}} - \mathrm{A}_{t_{n+1}^{(i),(l)}}}{\mathrm{R}_i\left(d_{n+2}^{(i)} - d_{n+1}^{(i)}\right)}\right) \cdot \mathrm{F}_i\left(n+1, t_n^{(i),(l)}, t_{n+1}^{(i),(l)}\right), & t_n^{(i),(l)} < t_{n+1}^{(i),(l)} \\ \mathrm{N}\left(v_{n+2}^{(i)} \middle| \tau^i + \mu_{t_{n+2}^{(i),(l)}}, \sigma_{t_{n+2}^{(i),(l)}}\right) \cdot 1_{\left\{\mathrm{A}_{t_{n+2}^{(i),(l)}} \in \left(a_{n+2}^{(i)}, b_{n+2}^{(i)}\right)\right\}} \cdot 1_{\left\{0.125 \le \mathrm{R}_i\left(d_{n+1}^{(i)} - d_n^{(i)}\right) \le 2\right\}} \cdot \rho\left(\frac{0.5}{\mathrm{R}_i\left(d_{n+1}^{(i)} - d_n^{(i)}\right)}, \frac{\mathrm{A}_{t_{n+2}^{(i),(l)}} - \mathrm{A}_{t_{n+1}^{(i),(l)}}}{\mathrm{R}_i\left(d_{n+2}^{(i)} - d_{n+1}^{(i)}\right)}\right) \mathrm{F}_i\left(n+1, t_{n+1}^{(i),(l)}, t_{n+1}^{(i),(l)}\right), & t_n^{(i),(l)} = t_{n+1}^{(i),(l)} \end{cases}$$

$$\propto \begin{cases} \boldsymbol{1}_{\left\{\mathbf{0.25} \le \frac{\mathbf{A}_{\boldsymbol{t}_{n+1}^{(i),(l)}} - \mathbf{A}_{\boldsymbol{t}_n^{(i),(l)}}}{\mathbf{R}_i\left(\boldsymbol{d}_{n+1}^{(i)} - \boldsymbol{d}_n^{(i)}\right)} \le \mathbf{4}\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{A}_{\boldsymbol{t}_{n+1}^{(i),(l)}} - \mathbf{A}_{\boldsymbol{t}_n^{(i),(l)}}}{\mathbf{R}_i\left(\boldsymbol{d}_{n+1}^{(i)} - \boldsymbol{d}_n^{(i)}\right)}, \frac{\mathbf{A}_{\boldsymbol{t}_{n+2}^{(i),(l)}} - \mathbf{A}_{\boldsymbol{t}_{n+1}^{(i),(l)}}}{\mathbf{R}_i\left(\boldsymbol{d}_{n+2}^{(i)} - \boldsymbol{d}_{n+1}^{(i)}\right)}\right) \cdot \mathbf{F}_i\left(\boldsymbol{n}+\mathbf{1}, \boldsymbol{t}_n^{(i),(l)}, \boldsymbol{t}_{n+1}^{(i),(l)}\right), & \boldsymbol{t}_n^{(i),(l)} < \boldsymbol{t}_{n+1}^{(i),(l)} \\ \boldsymbol{1}_{\left\{\mathbf{0.125} \le \mathbf{R}_i\left(\boldsymbol{d}_{n+1}^{(i)} - \boldsymbol{d}_n^{(i)}\right) \le \mathbf{2}\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_{n+1}^{(i)} - \boldsymbol{d}_n^{(i)}\right)}, \frac{\mathbf{A}_{\boldsymbol{t}_{n+2}^{(i),(l)}} - \mathbf{A}_{\boldsymbol{t}_{n+1}^{(i),(l)}}}{\mathbf{R}_i\left(\boldsymbol{d}_{n+2}^{(i)} - \boldsymbol{d}_{n+1}^{(i)}\right)}\right) \mathbf{F}_i\left(\boldsymbol{n}+\mathbf{1}, \boldsymbol{t}_{n+1}^{(i),(l)}, \boldsymbol{t}_{n+1}^{(i),(l)}\right), & \boldsymbol{t}_n^{(i),(l)} = \boldsymbol{t}_{n+1}^{(i),(l)} \end{cases}$$

2.2.2. $t_{n+1}^{(i),(l)} = t_{n+2}^{(i),(l)}$:

$$= \begin{cases} \mathrm{N}\left(v_{n+2}^{(i)} \middle| \tau^i + \mu_{t_{n+2}^{(i),(l)}}, \sigma_{t_{n+2}^{(i),(l)}}\right) \cdot 1_{\left\{\mathrm{A}_{t_{n+2}^{(i),(l)}} \in \left(a_{n+2}^{(i)}, b_{n+2}^{(i)}\right)\right\}} \cdot 1_{\left\{0.25 \le \frac{\mathrm{A}_{t_{n+1}^{(i),(l)}} - \mathrm{A}_{t_n^{(i),(l)}}}{\mathrm{R}_i\left(d_{n+1}^{(i)} - d_n^{(i)}\right)} \le 4\right\}} \cdot \rho\left(\frac{\mathrm{A}_{t_{n+1}^{(i),(l)}} - \mathrm{A}_{t_n^{(i),(l)}}}{\mathrm{R}_i\left(d_{n+1}^{(i)} - d_n^{(i)}\right)}, \frac{0.5}{\mathrm{R}_i\left(d_{n+2}^{(i)} - d_{n+1}^{(i)}\right)}\right) \cdot \mathrm{F}_i\left(n+1, t_n^{(i),(l)}, t_{n+1}^{(i),(l)}\right), & t_n^{(i),(l)} < t_{n+1}^{(i),(l)} \\ \mathrm{N}\left(v_{n+2}^{(i)} \middle| \tau^i + \mu_{t_{n+2}^{(i),(l)}}, \sigma_{t_{n+2}^{(i),(l)}}\right) \cdot 1_{\left\{\mathrm{A}_{t_{n+2}^{(i),(l)}} \in \left(a_{n+2}^{(i)}, b_{n+2}^{(i)}\right)\right\}} \cdot 1_{\left\{0.125 \le \mathrm{R}_i\left(d_{n+1}^{(i)} - d_n^{(i)}\right) \le 2\right\}} \cdot \rho\left(\frac{0.5}{\mathrm{R}_i\left(d_{n+1}^{(i)} - d_n^{(i)}\right)}, \frac{0.5}{\mathrm{R}_i\left(d_{n+2}^{(i)} - d_{n+1}^{(i)}\right)}\right) \mathrm{F}_i\left(n+1, t_{n+1}^{(i),(l)}, t_{n+1}^{(i),(l)}\right), & t_n^{(i),(l)} = t_{n+1}^{(i),(l)} \end{cases}$$

$$\propto \begin{cases} \boldsymbol{1}_{\left\{\mathbf{0.25} \le \frac{\mathbf{A}_{\boldsymbol{t}_{n+1}^{(i),(l)}} - \mathbf{A}_{\boldsymbol{t}_n^{(i),(l)}}}{\mathbf{R}_i\left(\boldsymbol{d}_{n+1}^{(i)} - \boldsymbol{d}_n^{(i)}\right)} \le \mathbf{4}\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{A}_{\boldsymbol{t}_{n+1}^{(i),(l)}} - \mathbf{A}_{\boldsymbol{t}_n^{(i),(l)}}}{\mathbf{R}_i\left(\boldsymbol{d}_{n+1}^{(i)} - \boldsymbol{d}_n^{(i)}\right)}, \frac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_{n+2}^{(i)} - \boldsymbol{d}_{n+1}^{(i)}\right)}\right) \cdot \mathbf{F}_i\left(\boldsymbol{n}+\mathbf{1}, \boldsymbol{t}_n^{(i),(l)}, \boldsymbol{t}_{n+1}^{(i),(l)}\right), & \boldsymbol{t}_n^{(i),(l)} < \boldsymbol{t}_{n+1}^{(i),(l)} \\ \boldsymbol{1}_{\left\{\mathbf{0.125} \le \mathbf{R}_i\left(\boldsymbol{d}_{n+1}^{(i)} - \boldsymbol{d}_n^{(i)}\right) \le \mathbf{2}\right\}} \cdot \boldsymbol{\rho}\left(\frac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_{n+1}^{(i)} - \boldsymbol{d}_n^{(i)}\right)}, \frac{\mathbf{0.5}}{\mathbf{R}_i\left(\boldsymbol{d}_{n+2}^{(i)} - \boldsymbol{d}_{n+1}^{(i)}\right)}\right) \mathbf{F}_i\left(\boldsymbol{n}+\mathbf{1}, \boldsymbol{t}_{n+1}^{(i),(l)}, \boldsymbol{t}_{n+1}^{(i),(l)}\right), & \boldsymbol{t}_n^{(i),(l)} = \boldsymbol{t}_{n+1}^{(i),(l)} \end{cases}$$

2.3. Termination: iterate the above sampling steps to get $t_m^{(i),(l)}$, $t_{m-1}^{(i),(l)}$, ⋯, $t_k^{(i),(l)}$ recursively until possible.

a. If $t_{k-1}^{(i),(l)}$ was not sampled because $\mathrm{A}_{t_k^{(i),(l)}} - 0.25\mathrm{R}_i\left(d_k^{(i)} - d_{k-1}^{(i)}\right) < \mathrm{A}_1$ and $1_{\left\{0.125 \le \mathrm{R}_i\left(d_k^{(i)} - d_{k-1}^{(i)}\right) \le 2\right\}} = 0$, keep the samples in the form of:

$$\left\{\left(\tilde{t}_1^{(i),(l)}, \tilde{n}_1^{(i),(l)}\right), \left(\tilde{t}_2^{(i),(l)}, \tilde{n}_2^{(i),(l)}\right), \cdots, \left(\tilde{t}_{m-k+1}^{(i),(l)}, \tilde{n}_{m-k+1}^{(i),(l)}\right)\right\} \triangleq \left\{\left(t_k^{(i),(l)}, k\right), \left(t_{k+1}^{(i),(l)}, k+1\right), \cdots, \left(t_m^{(i),(l)}, m\right)\right\}$$

b. Otherwise, discard them and go back to 2.1 for starting the sampling procedure again.

3. Updating transition parameters by ML estimators:

Suppose that we get $\mathrm{L}^{(i)}$ samples for the core i by running the above backward sampling algorithm.

Let:

$$\lambda_{i,l} \triangleq \text{length of the sample } l \text{ for the core } i$$

$$\mathrm{A} \triangleq \#\text{ of samples where } \tilde{t}_1^{(i),(l)} > 1 \text{ and } \tilde{n}_1^{(i),(l)} > 1$$

$$\mathrm{B} \triangleq \#\text{ of samples where } \tilde{t}_1^{(i),(l)} > 1 \text{ and } \tilde{n}_1^{(i),(l)} = 1$$

$$\mathrm{C} \triangleq \#\text{ of samples where } \tilde{t}_1^{(i),(l)} = 1 \text{ and } \tilde{n}_1^{(i),(l)} > 1$$

$$\mathrm{D} \triangleq \#\text{ of samples where } \tilde{n}_{\lambda_{i,l}}^{(i),(l)} < \mathrm{N}_i$$

Then,

$$\widehat{\boldsymbol{\alpha}}_i \triangleq \frac{\mathbf{A}+\mathbf{B}}{\mathbf{L}^{(i)}}, \qquad \widehat{\boldsymbol{\beta}}_i \triangleq \frac{\mathbf{C}}{\mathbf{L}^{(i)}}, \qquad \widehat{\boldsymbol{\gamma}}_i \triangleq \frac{\mathbf{A}}{\mathbf{A}+\mathbf{B}+\sum_{l=1}^{\mathbf{L}^{(i)}} \tilde{\boldsymbol{t}}_1^{(i),(l)} - \mathbf{L}^{(i)}}$$

$$\widehat{\boldsymbol{\phi}}_i \triangleq \frac{\sum_{l=1}^{\mathbf{L}^{(i)}} \tilde{\boldsymbol{t}}_1^{(i),(l)} - \mathbf{L}^{(i)}}{\mathbf{A}+\mathbf{B}+\sum_{l=1}^{\mathbf{L}^{(i)}} \tilde{\boldsymbol{t}}_1^{(i),(l)} - \mathbf{L}^{(i)}}, \qquad \widehat{\boldsymbol{\psi}}_i \triangleq \frac{\sum_{l=1}^{\mathbf{L}^{(i)}} \tilde{\boldsymbol{n}}_1^{(i),(l)} - \sum_{l=1}^{\mathbf{L}^{(i)}} \tilde{\boldsymbol{n}}_{\lambda_{i,l}}^{(i),(l)} + \mathbf{L}^{(i)}(\mathbf{N}_i - \mathbf{1})}{(\mathbf{A}+\mathbf{C}) + \sum_{l=1}^{\mathbf{L}^{(i)}} \tilde{\boldsymbol{n}}_1^{(i),(l)} - \sum_{l=1}^{\mathbf{L}^{(i)}} \tilde{\boldsymbol{n}}_{\lambda_{i,l}}^{(i),(l)} + \mathbf{L}^{(i)}(\mathbf{N}_i - \mathbf{1})}, \qquad \widehat{\boldsymbol{\eta}}_i \triangleq \frac{\sum_{l=1}^{\mathbf{L}^{(i)}} \boldsymbol{\lambda}_{i,l} - \mathbf{L}^{(i)}}{\mathbf{D} + \sum_{l=1}^{\mathbf{L}^{(i)}} \boldsymbol{\lambda}_{i,l} - \mathbf{L}^{(i)}}$$

$$\widehat{\mathbf{R}}_i \triangleq \frac{\mathbf{1}}{\mathbf{L}^{(i)}} \sum_{l=1}^{\mathbf{L}^{(i)}} \frac{\mathbf{A}_{\tilde{t}_{\lambda_{i,l}}^{(i),(l)}} - \mathbf{A}_{\tilde{t}_1^{(i),(l)}}}{\boldsymbol{d}_{\tilde{n}_{\lambda_{i,l}}^{(i),(l)}} - \boldsymbol{d}_{\tilde{n}_1^{(i),(l)}}}$$

Of course, it is possible to do above in the Bayesian style with appropriate priors. For example, if we use beta priors on α_i, the hyperparameters act like pseudo-counts, like the estimates in the next part. Priors usually contribute the stability of inferred parameters, which prevent some extreme estimations (e.g. $\pm\infty$) caused by the lack of data and reflect our strong/weak beliefs on the parameters from our past experiences.

3.1. About δ_i: its ML estimator is always 1, which is not proper to our purpose of maximizing the length of aligned data, so…

Suppose that we give a prior on $\delta_i \sim \mathrm{Beta}(z, w)$ for some hyperparameters z and w, then

$$\hat{\boldsymbol{\delta}}_i \triangleq \frac{\sum_{l=1}^{\mathrm{L}^{(i)}} \tilde{\boldsymbol{n}}_1^{(i),(l)} - \sum_{l=1}^{\mathrm{L}^{(i)}} \tilde{\boldsymbol{n}}_{\lambda_{i,l}}^{(i),(l)} + \mathbf{NL}^{(i)} - \mathbf{L}^{(i)} + \boldsymbol{z} - \mathbf{1}}{\sum_{l=1}^{\mathrm{L}^{(i)}} \tilde{\boldsymbol{n}}_1^{(i),(l)} - \sum_{l=1}^{\mathrm{L}^{(i)}} \tilde{\boldsymbol{n}}_{\lambda_{i,l}}^{(i),(l)} + \mathbf{NL}^{(i)} - \mathbf{L}^{(i)} + \boldsymbol{z} + \boldsymbol{w} - \mathbf{2}}$$

To make this algorithm “meaningful”, it is recommended to set $1 \approx z \ll w$, which makes $\hat{\delta}_i$ stay in the range of low values near to 0.

4. Updating stack parameters by ML estimators:

If we are only interested in aligning each core to the stack, it is just needed to learn shift parameters τ_i’s for the cores. If we are interested in constructing the new stack from the data as well, it is also required to learn means and standard deviations $\{\mu_t\}_{t=1}^{\mathrm{T}}$ and $\{\sigma_t\}_{t=1}^{\mathrm{T}}$. An EM algorithm can learn both the alignments and stack parameters simultaneously: to be specific, it first aligns each core to the stack by generating samples, and then updates the stack parameters by using samples.

$$\hat{\boldsymbol{\tau}}_i \triangleq \frac{\sum_{l=1}^{\mathrm{L}^{(i)}} \sum_{s=1}^{\lambda_{i,l}} \left(\boldsymbol{v}_{\tilde{\boldsymbol{n}}_s^{(i),(l)}}^{(i)} - \tilde{\boldsymbol{\mu}}_{\tilde{\boldsymbol{t}}_s^{(i),(l)}} \right)}{\sum_{l=1}^{\mathrm{L}^{(i)}} \boldsymbol{\lambda}_{i,l}}$$

$$\hat{\boldsymbol{\mu}}_t \triangleq \frac{\sum_{i=1}^{\mathrm{K}} \sum_{l=1}^{\mathrm{L}^{(i)}} \sum_{s=1}^{\lambda_{i,l}} \left(\boldsymbol{v}_{\tilde{\boldsymbol{n}}_s^{(i),(l)}}^{(i)} - \hat{\boldsymbol{\tau}}_i \right) \mathbf{1}_{\left\{\tilde{\boldsymbol{t}}_s^{(i),(l)} = \boldsymbol{t}\right\}}}{\sum_{i=1}^{\mathrm{K}} \sum_{l=1}^{\mathrm{L}^{(i)}} \sum_{s=1}^{\lambda_{i,l}} \mathbf{1}_{\left\{\tilde{\boldsymbol{t}}_s^{(i),(l)} = \boldsymbol{t}\right\}}}$$

$$\hat{\boldsymbol{\sigma}}_t \triangleq \frac{\sum_{i=1}^{\mathrm{K}} \sum_{l=1}^{\mathrm{L}^{(i)}} \sum_{s=1}^{\lambda_{i,l}} \left(\boldsymbol{v}_{\tilde{\boldsymbol{n}}_s^{(i),(l)}}^{(i)} - \hat{\boldsymbol{\tau}}_i - \hat{\boldsymbol{\mu}}_t \right)^{\mathbf{2}} \mathbf{1}_{\left\{\tilde{\boldsymbol{t}}_s^{(i),(l)} = \boldsymbol{t}\right\}}}{\sum_{i=1}^{\mathrm{K}} \sum_{l=1}^{\mathrm{L}^{(i)}} \sum_{s=1}^{\lambda_{i,l}} \mathbf{1}_{\left\{\tilde{\boldsymbol{t}}_s^{(i),(l)} = \boldsymbol{t}\right\}}}$$

, where $\{\tilde{\mu}_t\}_{t=1}^{\mathrm{T}}$ are the updated means in the previous iteration.

Note that EM algorithms are quite sensitive to the initial conditions, which means that it is, in general, not possible to see good inference if we do not use “proper” initializations. One suggestion to deal with this defect is initializing stack parameters by using any existing stacks (for example, LR04). Then, EM algorithms will “refine” our initial guesses on the parameters iteratively by using data.