Algorithm 1 Online EM algorithm with Control Adjustment

Input: Observed data $\{y_i^{(t)}\}$ and number of clusters N_c
Output: Optimal control $\{u_c^*\}_c$ of each cluster c and the final estimated model parameters $\Theta^* = \{A^*, B_c^*, C_c^*, Q^*, \Sigma_c^*, \gamma^*, \theta_c^*, \bar{\theta}^*, \beta_c^*\}$

Begin
Initialize Model parameters Θ^0, initial (fixed) control $\{u_c^0\}_c$
for $t = 0, ..., T$ **do**
  if $t \in T_1$ **then**
    Phase 1: Parameter Estimation (with fixed control)
    E-step
    Compute the posterior distribution of $x_i^{(t)}$ and $z_i^{(t)}$ for all i
    M-step
    Update model parameters $\{A^{new}, B_c^{new}, C_c^{new}, Q^{new}, \Sigma_c^{new}, \gamma^{new}, \theta_c^{new}, \bar{\theta}^{new}, \beta_c^{new}\}$
  else
    $t \in T_2$
    Phase 2: Control adjustment (with fixed parameters)
    Simulate future states $\hat{x}_i^{(t')}$ and $\hat{z}_i^{(t')}$ on a horizon of N timesteps using the estimated parameters Θ^{new}
    Update control (solving the MSE of future states):
    for each cluster c **do**
      $u_c^{new} = \arg\min_u \sum_{t'=1}^{N} \sum_{i,\ c_i=c} \left\| y_{target} - \hat{y}_i^{(t')} \right\|_2^2$
    end for
  end if
end for
End

N.B. : $\hat{x}_i^{(t')}$, $\hat{z}_i^{(t')}$ and $\hat{y}_i^{(t')}$ refer respectively to the simulated future states of x and z and the simulated future observations y, where $\forall t' \in \{1, ..., N\}$:

$$\hat{y}_i^{(t')} = \theta_{c_i}^{new} \hat{x}_i^{(t')} + \bar{\theta}^{new} \hat{z}_i^{(t')} + \sum_{j \neq i,\ c_j = c_i} \gamma_{ij}^{new} \hat{x}_j^{(t')} + \beta_c^{new} u$$