Algorithm 1 On line EM with Optimal Control Learning

Input: Observed data $\{y_i^{(t)}, x_i^{(t)}\}$, initial parameters $\{B^0, \Sigma^0, \theta_i^0, \gamma_{ij}^0, \beta_c^0, \sigma_0^2\}$, number of clusters $\mathrm{N}_c, initialclusterpriors \{\pi_c\}_c$.
for each time t **do**
  E-Step
  Compute the **cluster assignment probabilities:** $p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)}) = \frac{\pi_{c_i^{(t)}} p(y_i^{(t)}, x_i^{(t)}|c_i^{(t)})}{\sum_c \pi_c p(y_i^{(t)}, x_i^{(t)}|c)}$
  for each cluster c **do**:
    for each i **do**:
      Compute posterior mean: $\mu_{z|x}^{(i,t)} = B_c^\top (B_c B_c^\top + \Sigma_c)^{-1} x_i^{(t)}$
      Compute posterior covariance: $\Sigma_{z|x} = I - B_c^\top (B_c B_c^\top + \Sigma_c)^{-1} B_c$
    end for
  end for

  M-Step: Update model parameters
  for each cluster c **do**
    Update B: $B_c^{new} = \left(\sum_{i,t} p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)}) x_i^{(t)} \mu_{z|x}^{(i,t)\top}\right) \left(\sum_{i,t} p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)})(\Sigma_{z|x} + \mu_{z|x}^{(i,t)} \mu_{z|x}^{(i,t)\top})\right)^{-1}$

    Update Σ: $\Sigma_c^{new} = \frac{1}{\sum_{i,t} p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)})} \sum_{i,t} p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)}) \left(x_i^{(t)} x_i^{(t)\top} - B_c^{new} \mu_{z|x}^{(i,t)} x_i^{(t)\top} - x_i^{(t)} \mu_{z|x}^{(i,t)\top} B_c^{new\top} + B\right.$
    for each i **do**
      Update θ_c:
      $\theta_c^{new} = \left(\sum_t p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)}) x_i^{(t)} x_i^{(t)\top}\right)^{-1} \left(\sum_t p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)}) x_i^{(t)} y_i^{(t)}\right)$
    end for
    for each i, j **do**
      Update γ_{ij}:
      $\gamma_{ij}^{new} = \left(\sum_t p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)}) x_j^{(t)} x_j^{(t)\top}\right)^{-1} \left(\sum_t p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)}) x_j^{(t)} (y_i^{(t)} - \theta_i^\top x_i^{(t)})\right)$
    end for
    for each cluster c **do**
      Update β_c:
      $\beta_c^{new} = \frac{\sum_{i, c^*(i)=c} \sum_t p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)}) u_c^{(t)} (y_i^{(t)} - \theta_i^{new\top} x_i^{(t)} - \sum_{j \neq i, c^*(j)=c^*(i)} \gamma_{ij}^{new\top} x_j^{(t)})}{\sum_{i, c^*(i)=c} \sum_t p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)}) u_c^{(t)} u_c^{(t)}}$
    end for
    Update σ^2:
    $\sigma_c^{2,new} = \frac{1}{nT} \sum_{i,t} p(c_i^{(t)}|x_i^{(t)}, y_i^{(t)}) \left(y_i^{(t)} - \theta_i new^\top x_i^{(t)} - \sum_{j \neq i, c^*(j)=c^*(i)} \gamma_{ij}^{new} {}^\top x_j^{(t)} - \beta_{c^*(i)}^{new} u_{c^*(i)}^{(t)}\right)^2$

    Final-Step: Optimal Control Learning
    for each cluster c **do**
      $u_c^{(t)} = \frac{\sum_{i, c^*(i)=c} \left(y_i^{(t)} - \theta_i^{new\top} x_i^{(t)} - \sum_{j \neq i, c^*(j)=c} \gamma_{ij}^{new\top} x_j^{(t)}\right)}{\sum_{i, c^*(i)=c} \beta_c^{new\top} \beta_c^{new}}$
    end for