

$$\mathcal{L}_{\text{NCE}}(\{x_i, y_i\}_{i=1}^K; f, h) = \sum_{i=1}^K \log \left(\frac{e^{f(y_i)^T h(x_i)}}{\sum_{j=1}^K e^{f(y_i)^T h(x_j)}} \right) + \sum_{j=1}^K \log \left(\frac{e^{f(y_i)^T h(x_i)}}{\sum_{i=1}^K e^{f(y_i)^T h(x_j)}} \right) \quad (8)$$

$$\mathcal{L}_{\text{BC}}(\{s_i, a_i, s_i^+, \ell_i\}_{i=1}^K; \pi) = \sum_{i=1}^K \log \pi(a_i | s_i, \xi(\ell_i)) + \log \pi(a_i | s_i, \psi(s_i^+)) \quad (9)$$

$$\begin{aligned} \mathcal{L}_{\text{TRA}}(\{s_i, a_i, s_i^+, g_i, \ell_i\}_{i=1}^K; \pi, \phi, \psi, \xi) &= \underbrace{\mathcal{L}_{\text{BC}}(\{s_i, a_i, s_i^+, \ell_i\}_{i=1}^K; \pi, \psi, \xi)}_{\text{behavioral cloning}} \\ &+ \underbrace{\mathcal{L}_{\text{NCE}}(\{s_i, s_i^+\}_{i=1}^K; \phi, \psi)}_{\text{temporal alignment}} + \underbrace{\mathcal{L}_{\text{NCE}}(\{g_i, \ell_i\}_{i=1}^K; \psi, \xi)}_{\text{task alignment}} \end{aligned} \quad (10)$$

Algorithm 1: Temporal Representation Alignment

- 1: **input:** dataset $\mathcal{D} = (\{s_{t,i}, a_{t,i}\}_{t=1}^H, \ell_i)_{i=1}^N$
 - 2: initialize networks $\Theta \triangleq (\pi, \phi, \psi, \xi)$
 - 3: **while** training **do**
 - 4: sample batch $\{(s_{t,i}, a_{t,i}, s_{t+k,i}, \ell_i)\}_{i=1}^K \sim \mathcal{D}$
 for $k \sim \text{Geom}(1 - \gamma)$
 - 5: $\Theta \leftarrow \Theta - \alpha \nabla_{\Theta} \mathcal{L}_{\text{TRA}}(\{s_{t,i}, a_{t,i}, s_{t+k,i}, \ell_i\}_{i=1}^K; \Theta)$
 - 6: **output:** language-conditioned policy $\pi(a_t | s_t, \xi(\ell))$
 goal-conditioned policy $\pi(a_t | s_t, \psi(g))$
-