Algorithm 1: Temporal Metric Distillation (TMD)

input: dataset $\mathcal{D}$, learning rate η
initialize representations ϕ, ψ, policy π
while training **do**
    sample $\mathcal{B} = \{s_i, a_i, s'_i, g_i\}_{i=1}^{N} \sim \mathcal{D}$
    $\overline{\psi} \leftarrow \psi$
    $(\phi, \psi) \leftarrow (\phi, \psi) - \eta \nabla_{\phi,\psi} \mathcal{L}_{\text{TMD}}(\phi, \psi; \overline{\psi}, \mathcal{B})$ ▷ *Eq.* (26)
    $\pi \leftarrow \pi - \eta \nabla_{\pi} \mathcal{L}_{\pi}(\phi, \psi, \pi; \mathcal{B})$ ▷ *Eq.* (27)
return π