$$\mathcal{L}_{\text{TMD}}\left(\phi, \psi; \overline{\psi}, \mathcal{B}\right) = \mathcal{L}_{\text{NCE}}\left(\phi, \psi; \mathcal{B}\right) + \zeta\Big(\mathcal{L}_{\mathcal{I}}\left(\phi, \psi; \mathcal{B}\right) + \mathcal{L}_{\mathcal{T}}\left(\phi, \overline{\psi}; \mathcal{B}\right)\Big)$$

$$\mathcal{L}_{\pi}\left(\pi; \phi, \psi, \{s_i, a_i, s'_i, g_i\}_{i=1}^{N}\right) = \sum_{i,j=1}^{N} (1-\lambda) d_{\text{MRN}}\left(\phi(s_i, \hat{a}_{ij}), \psi(g_j)\right), g_j\Big)$$

$$+ \lambda d_{\text{MRN}}\left(\phi(s_i, \hat{a}_{ii}), \psi(g_i)\right) + \alpha \left\| \hat{a}_{ii} - a_i \right\|_2^2$$

where $\hat{a}_{ij} = \pi(s_i, g_j)$, batch $\mathcal{B} \sim p^{\pi_\beta} = \{s_i, a_i, s'_i, g_i\}_{i=1}^{N}$.