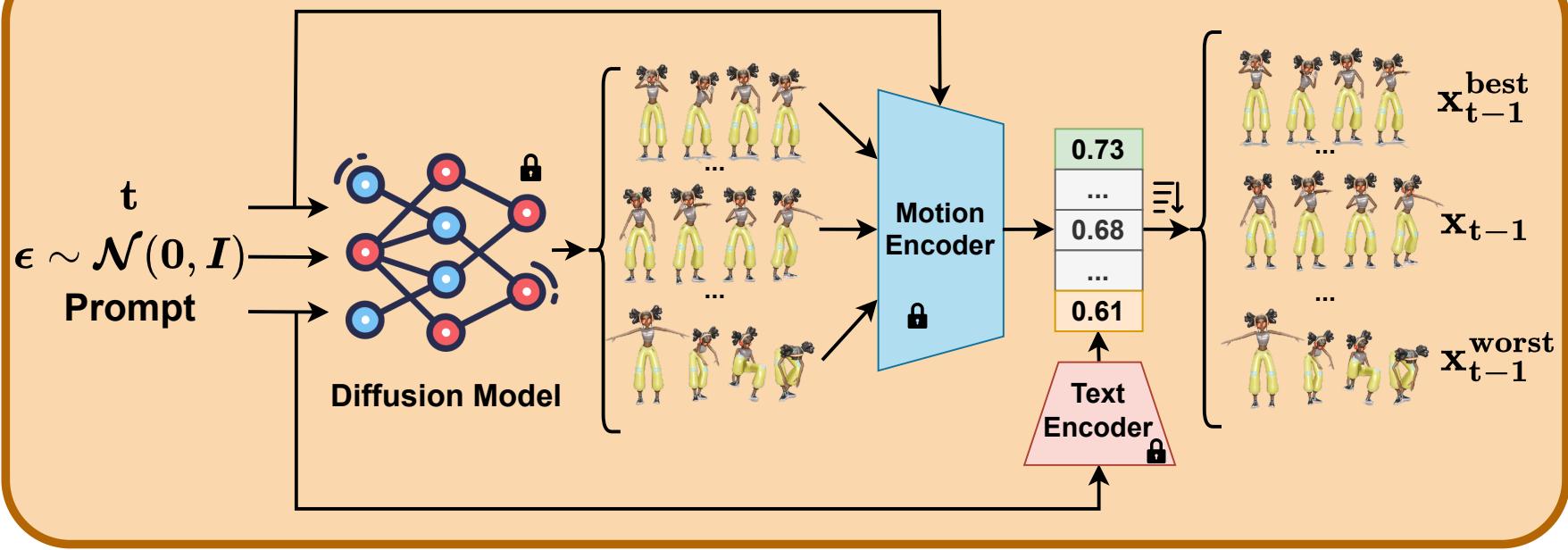
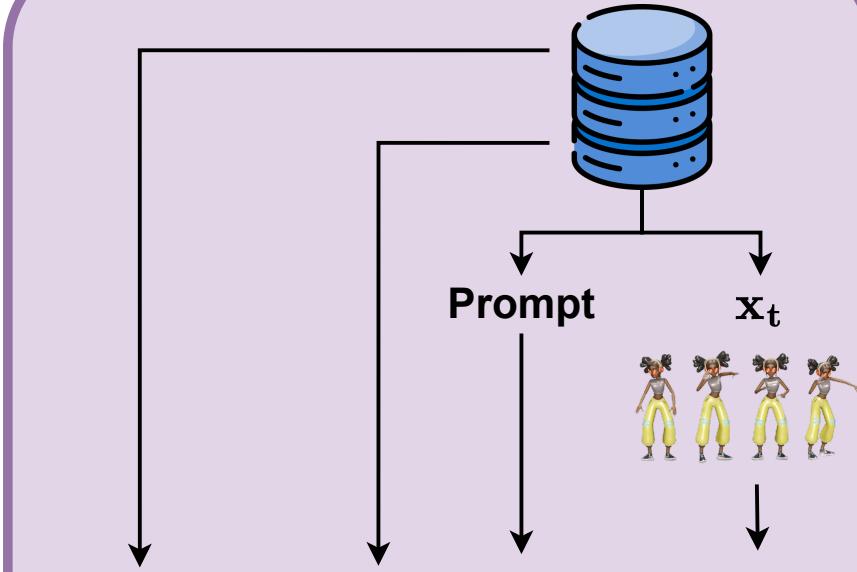


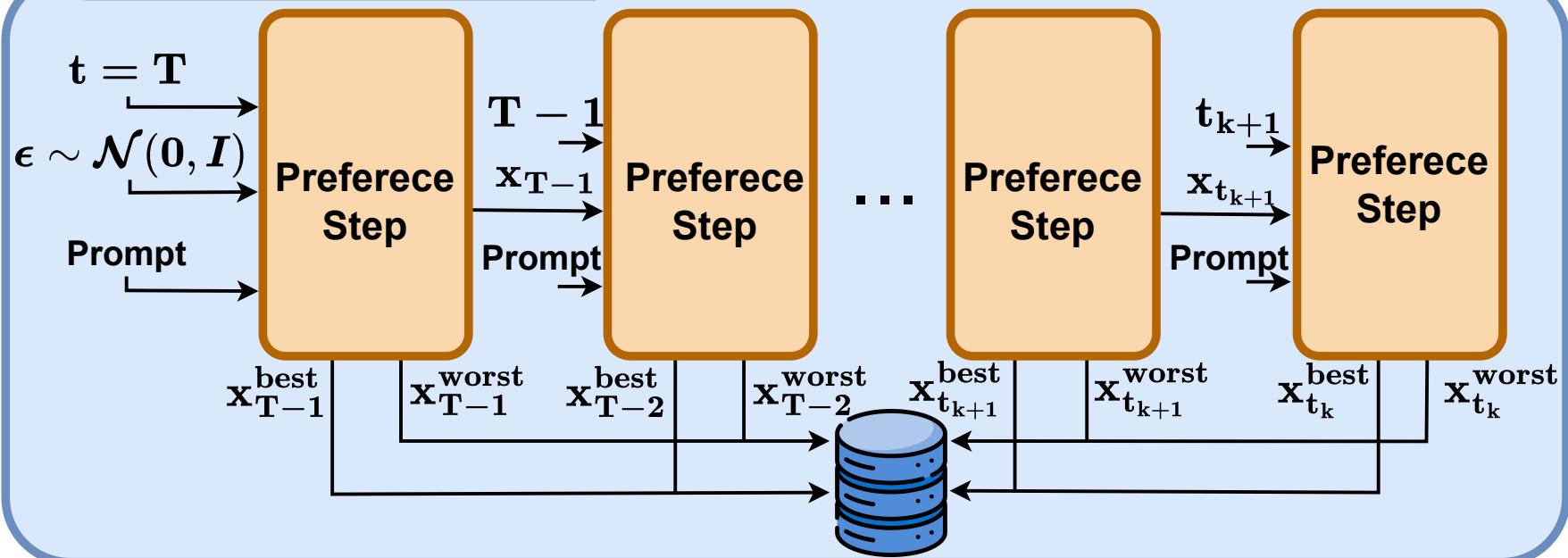
## Preference Step



## Preference Optimization



## Preference Collection



$$\left. \begin{array}{l} \uparrow p_{\theta}(x_{t-1}^{\text{best}} | x_t, c) \\ \downarrow p_{\theta}(x_{t-1}^{\text{worse}} | x_t, c) \end{array} \right\} \mathcal{L}_{\text{SPO}}(\theta)$$

