

(a) Learning diffusion discriminator

(b) Learning policy with diffusion reward