



Figure 1: Training evolution of the policy for the stochastic cart-pole experiment. We estimate the EIG after each training epoch, and we report the mean and standard deviation over 25 seeds.