

Figure 1: Response to **Question 5 by Reviewer vhg**d: Evaluation in the humanoid locomotion task. The shaded region corresponds to standard deviation using 5 different random seeds.

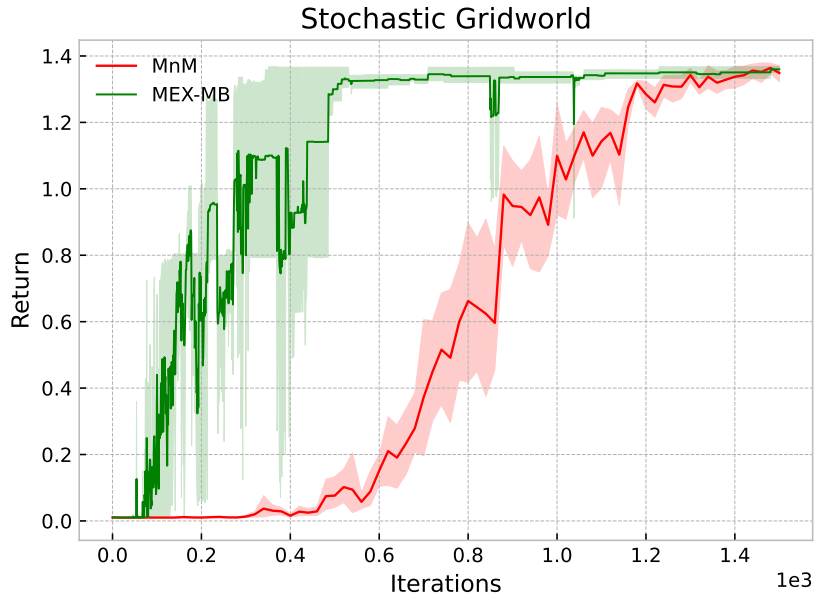


Figure 2: Response to **Originality by Reviewer wpG**c: Evaluation in the stochastic gridworld environment. The shaded region corresponds to standard deviation using 5 different random seeds.