Homework #2

Mahan Fathi

February 18, 2022

1 Problem 1

The model error plots:

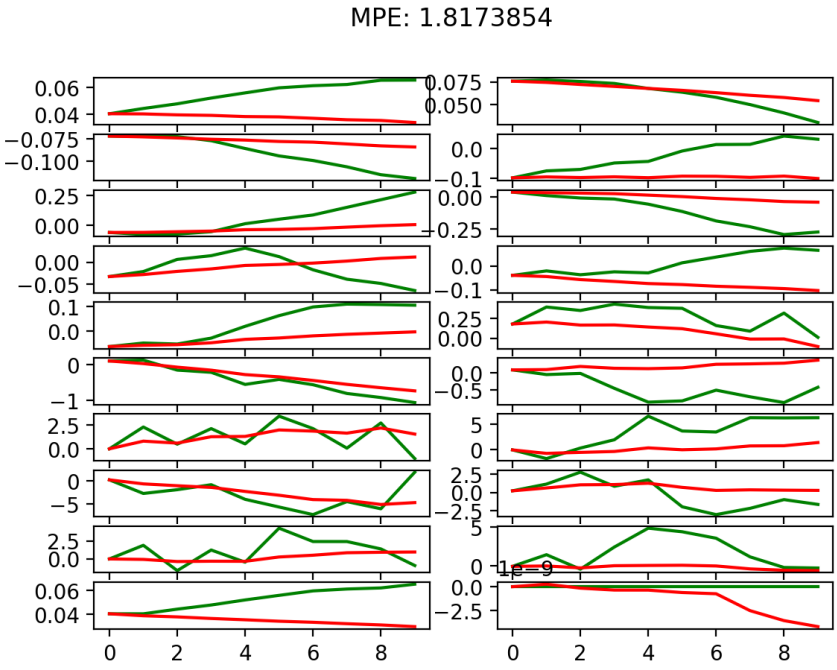

Figure 1: Experiment: `cheetah_n5_arch2x250`

The models corresponding to the last plot seem to be performing best. The first is trained only for a few number of iterations, i.e. 5. The second one has low network capacity, which apparently is not able to capture all the nuances of the underlying dynamics of the environment.

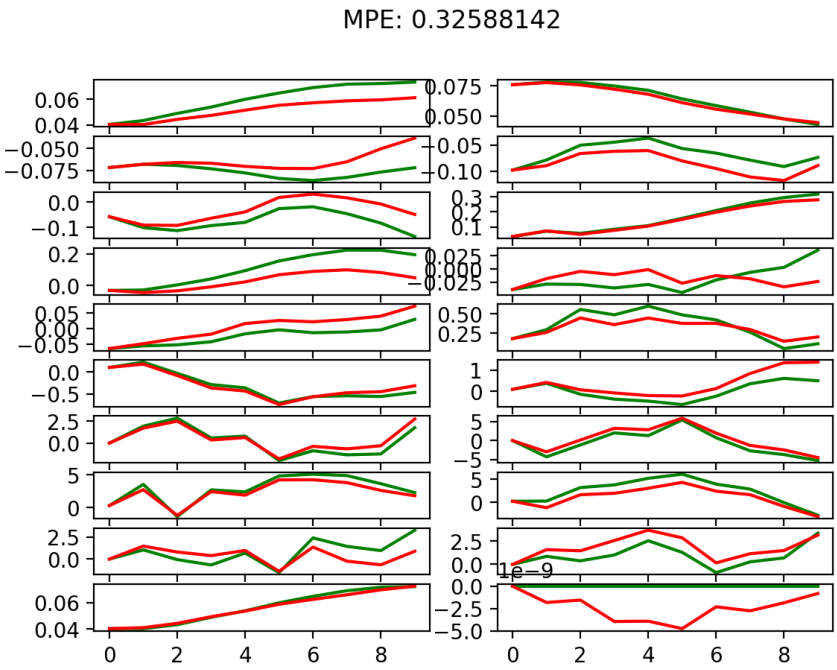

Figure 2: Experiment: `cheetah_n500_arch1x32`

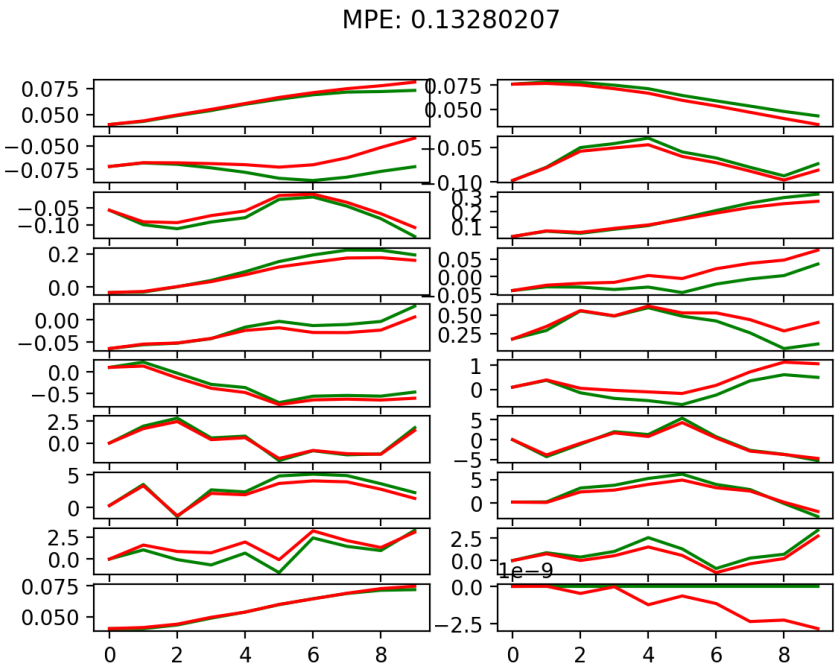

Figure 3: Experiment: `cheetah_n500_arch2x250`

2 Problem 2

Table 1: Problem 2 results

cheetah	Eval_AverageReturn	Train_AverageReturn
AverageReturn	-32.15	-167.19

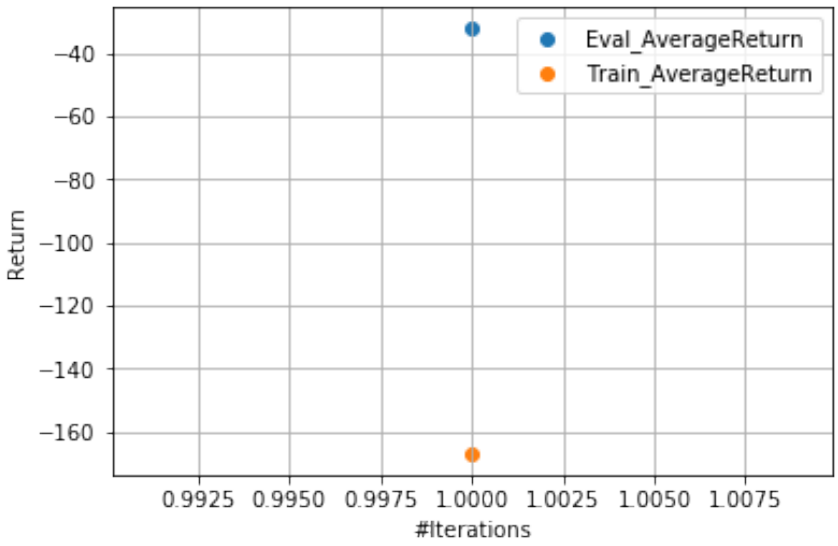

Figure 4: Problem 2 plot

3 Problem 3

Performance plots:

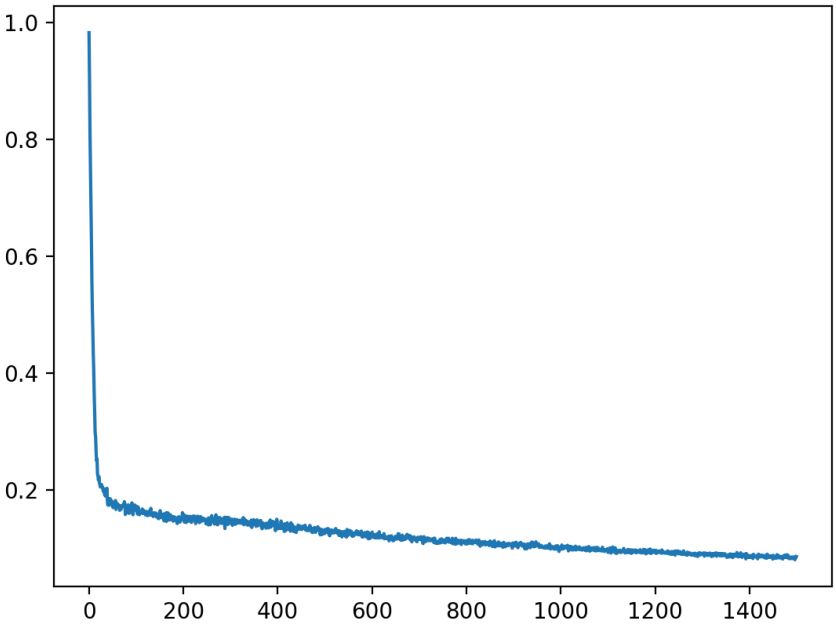

Figure 5: Training loss for `q3_cheetah`

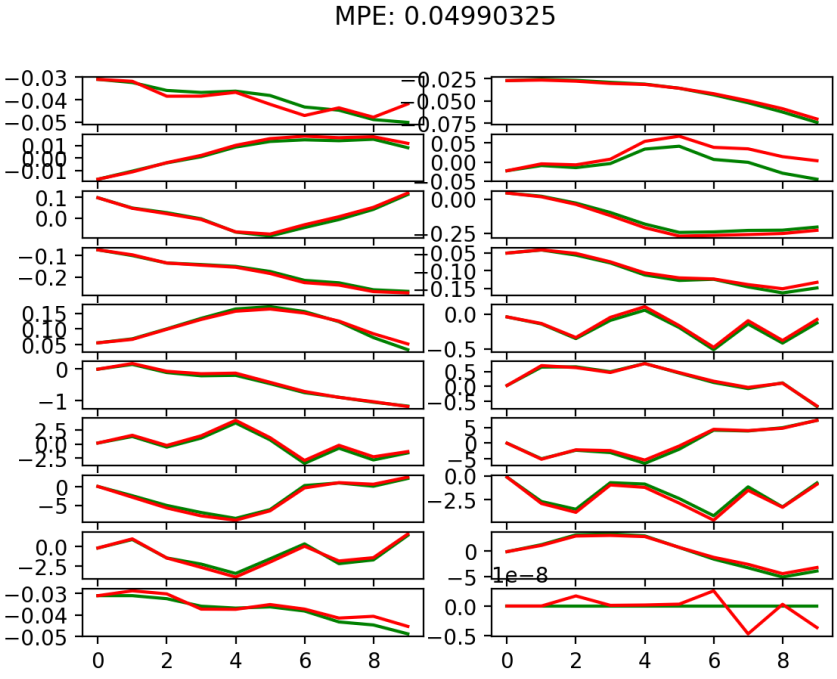

Figure 6: Model errors for `q3_cheetah`

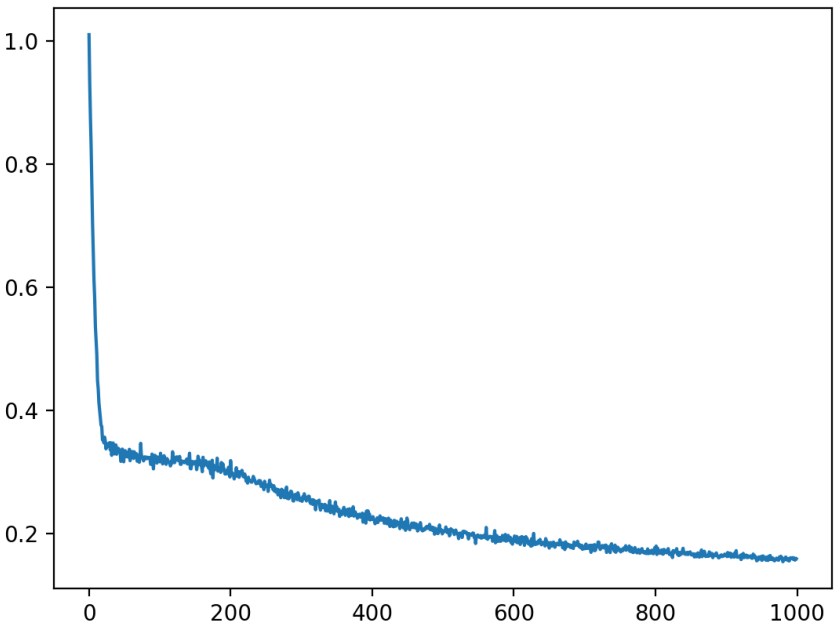

Figure 7: Training loss for `q3_reacher`

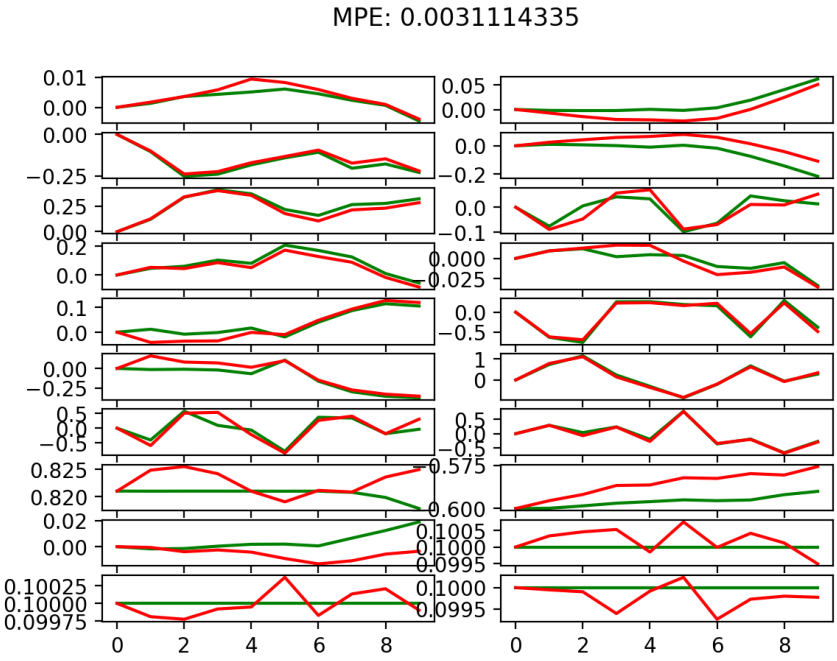

Figure 8: Model errors for `q3_reacher`

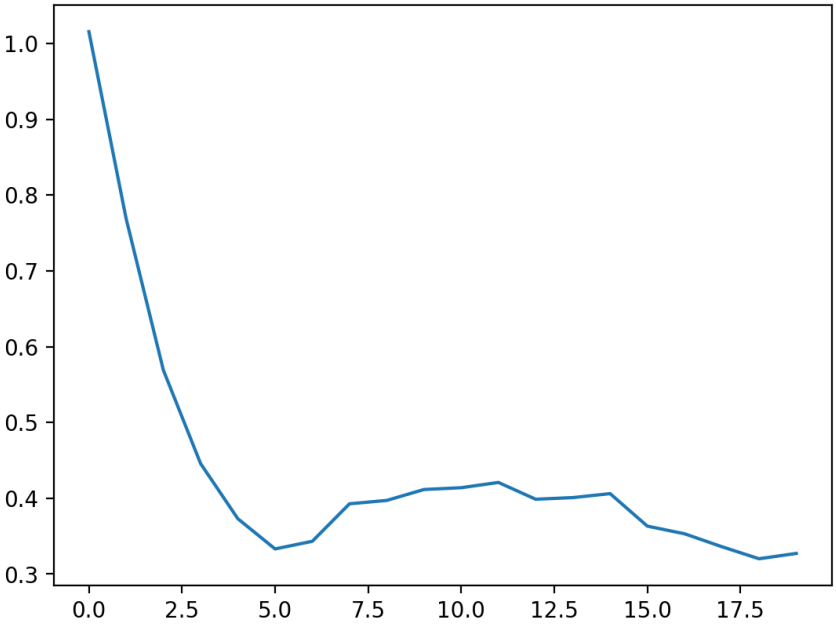

Figure 9: Training loss for `q3_obstacles`

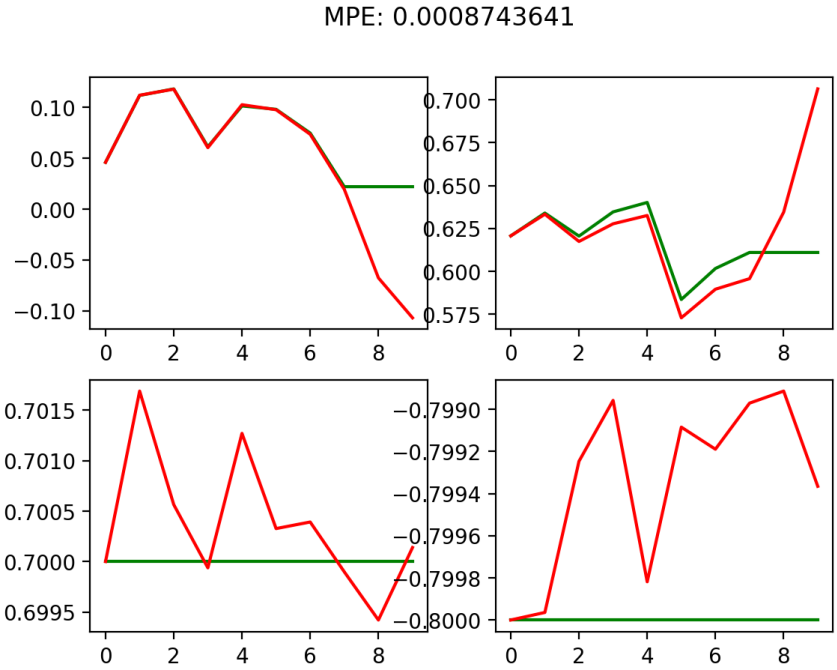

Figure 10: Model errors for `q3_obstacles`

4 Problem 4

4.1 Ensembles

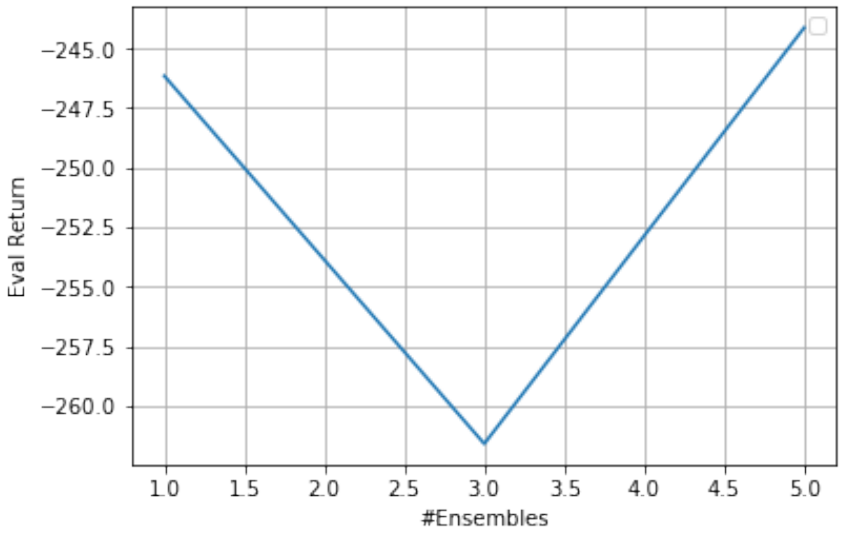

Figure 11: Ablation with regards to the number of ensembles

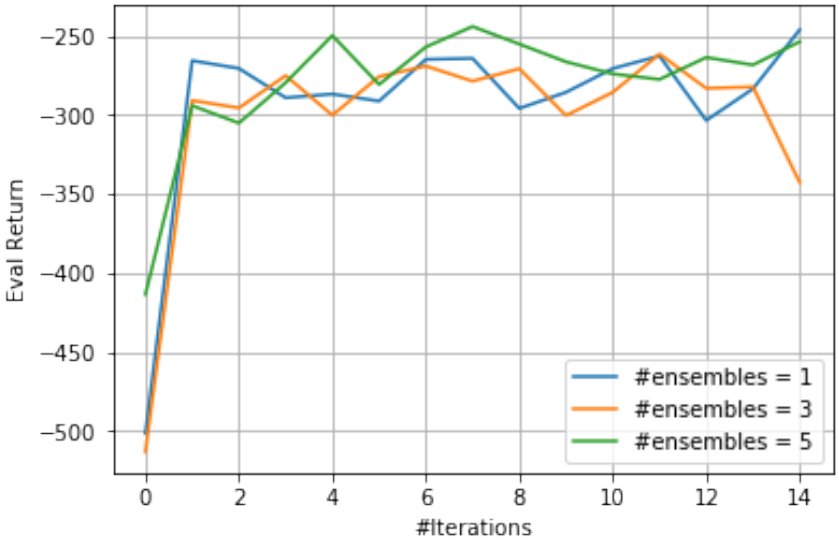

Figure 12: `Eval_AverageReturn` with different number of ensembles

4.2 MPC # Action Sequences

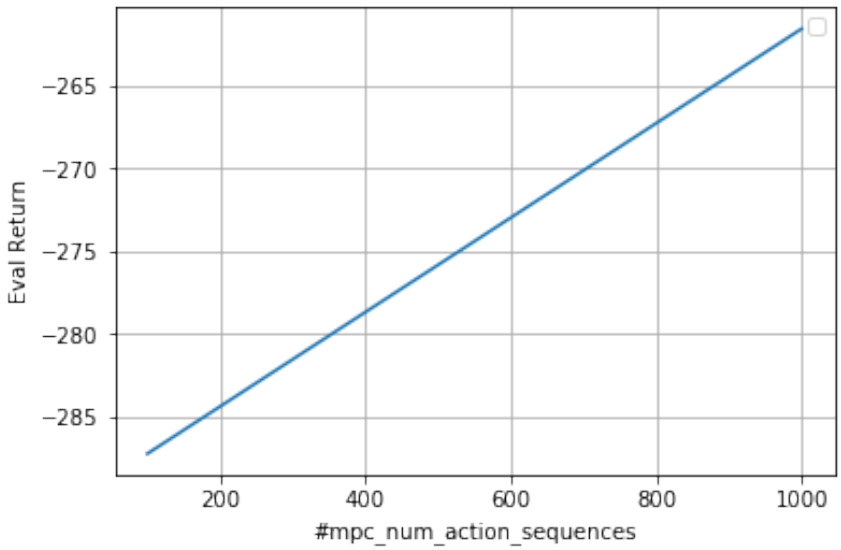

Figure 13: Ablation with regards to the number sequence candidates

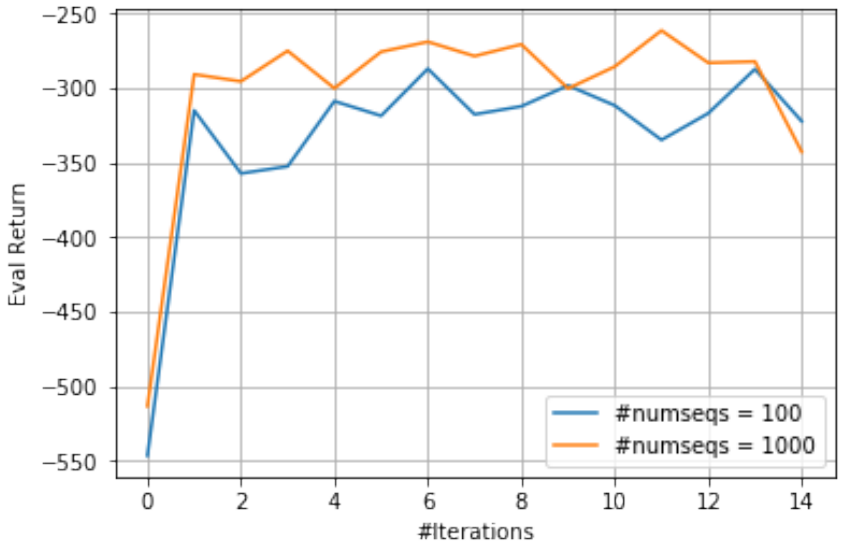

Figure 14: `Eval_AverageReturn` with different number of sequence candidates

4.3 Horizon

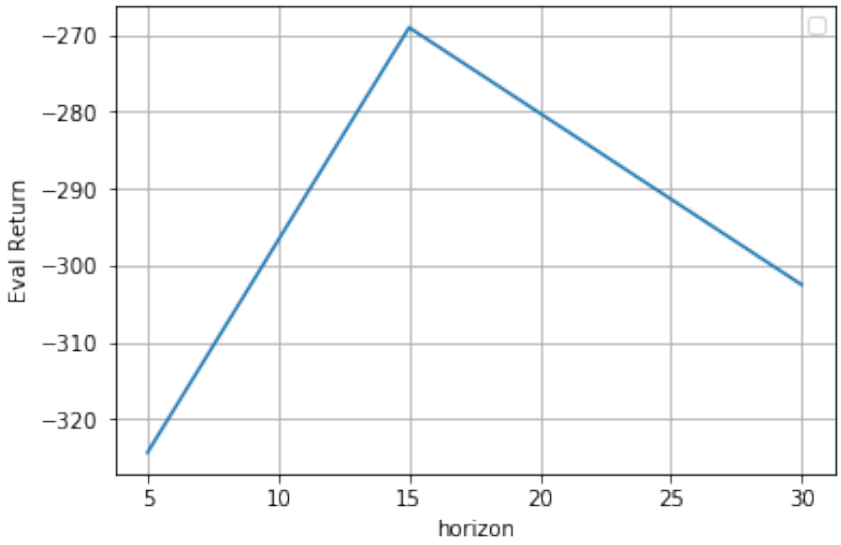

Figure 15: Ablation with regards to the planning horizon

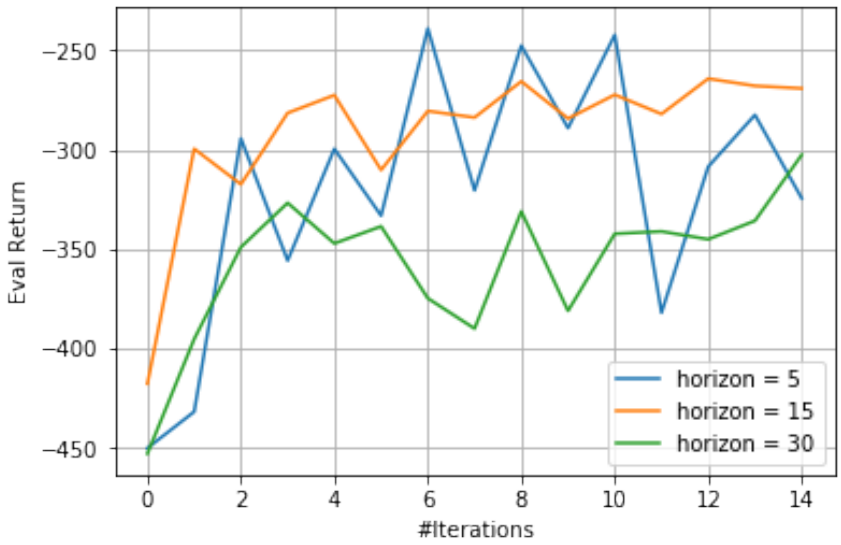

Figure 16: `Eval_AverageReturn` with different horizons.

5 Problem 5

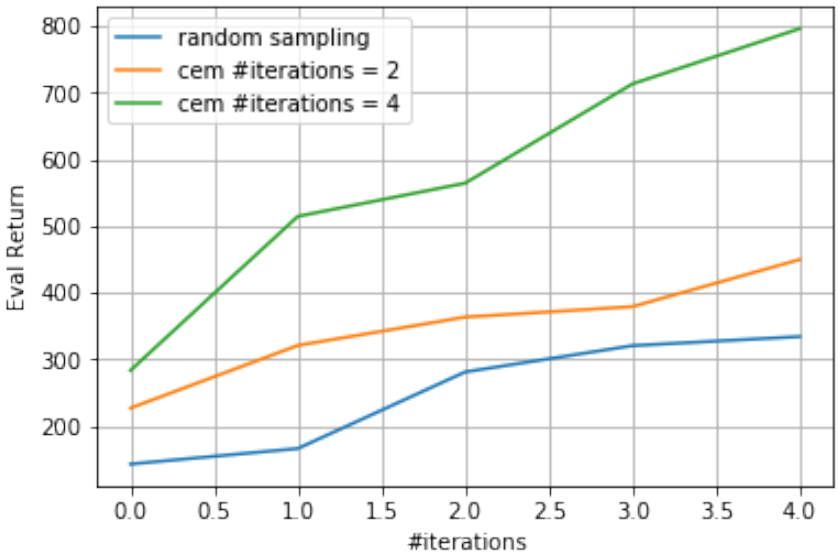

Figure 17: CEM compared to random actions. CEM clearly outperforms random sampling method, since directs the search using some sort of a heuristic. Moreover, more iterations in CEM leads to more accurate results and a thorough search over the planning space.