

Supplement to the submission titled “Learning Distinguishable Representations in Deep Q-Networks for Linear Transfer”

1 Hyperparameters used in the experiments

Table 1: λ and k values used in each environment

Environment	λ_{max}	k
LunarLander	0.05	25
Mountain Car	0.05	10
Breakout	0.01	25
SpaceInvaders	0.01	25
Asterix	0.01	25

Table 1 presents the values of k (clusters) and λ_{max} (maximum correlation coefficient) for all the tasks.

Table 2: Exit conditions for different environments

Environment	Last n	Average Return
LunarLander	10	200
Mountain Car	10	-200
Breakout	10	8
SpaceInvaders	10	25
Asterix	10	5

Table 2 outlines the exit conditions for each environment. Here, ‘Last n’ column specifies the number of episodes over which returns are averaged to determine if the condition is satisfied. For example, in the Lunar Lander environment, LFA will stop training once the average return over the last 10 episodes exceeds 200.