Energy-aware Job Scheduling (N = 4) MDP loss, property ExpUtil, shallow 3-layered network 104 = One-hot encoding Ordinal encoding One-hot encoding (Double Q-learning) Ordinal encoding (Double Q-learning) 10³ 10² 10^1 Average loss 10^{0} 10^{-1} 10-2 100 200 300 400 500 0 Number of episodes