Solve CartPole: Success Probability 60%-Cls with MDP-Regularization CartPole: Mean/ Median (dashed/ solid) Marginal Trial Rewards AlphaZero 500 MuZero 1.0 -MuZero h_{θ}^{-1} MuZero $D_{KL}(h_{\theta} || g_{\theta})$ Success Rewards one one **o** 0.6 Probability ... Average 100 100 0.2 0.0 + 10^{-2} 10⁰ 10^{-1} 20 70 80 Regularization Strength ω Training Episode