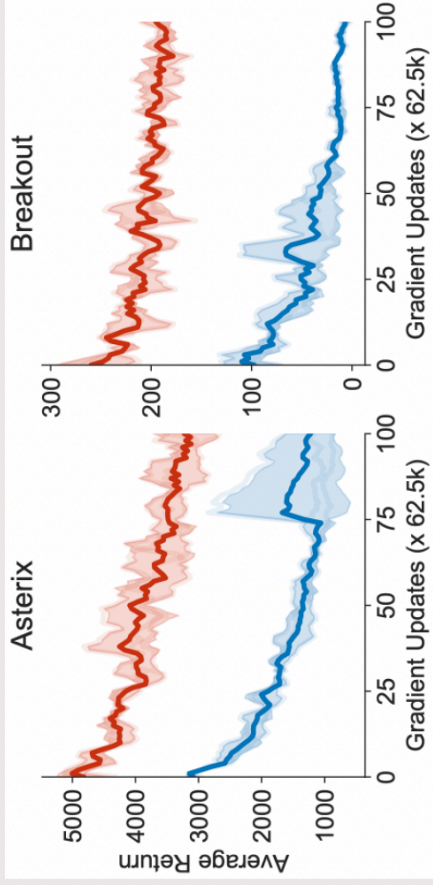
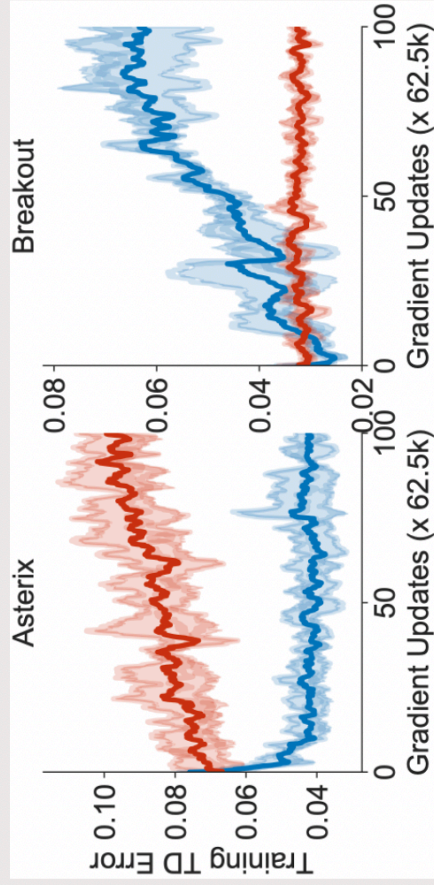


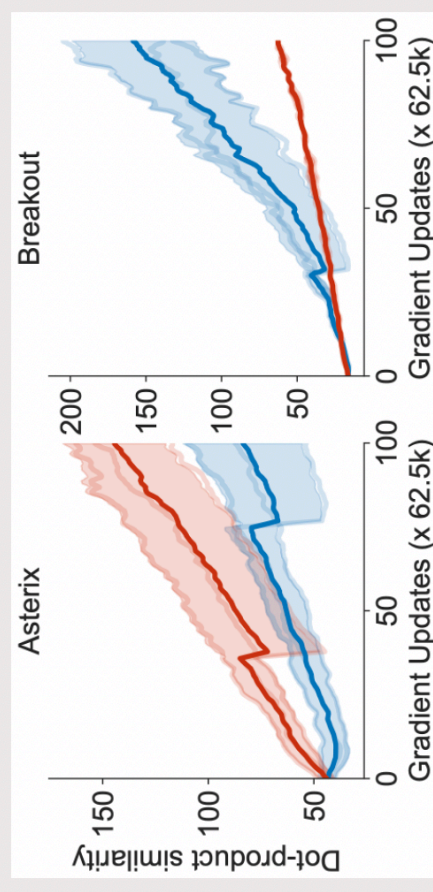
Return degrades even starting from good solutions



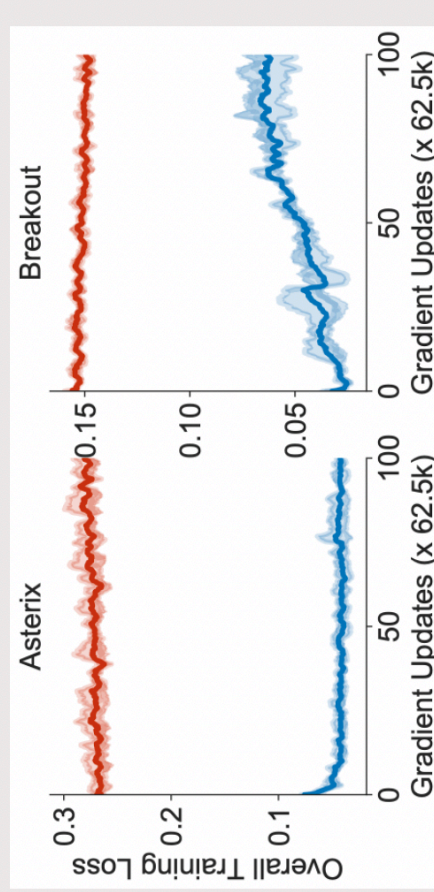
In particular, training TD error is low and decreasing



Feature dot products increase as returns decrease



Overall training error for CQL/DQN is still low!



DQN

CQL