IPD, 3 players 1 SHAPER 2 SHAPERs Reward Per Timestep 4 2 Shaper 1 Shaper 1 NL 1 Shaper 2 NL 2 NL 1 0 0 250 500 1000 750 1000 250 500 750 1.0 -1.0 State distribution 8.0 8.0 CD, C 0.6 0.6 C, CC DC, C D, CC 0.4 0.4 DD, C 0.2 0.2 0.0 0.0 250 750 500 750 1000 250 500 1000 Eval episode Eval episode