Training loss: workers = 8, τ = 4, β = 3.96 2.0 1.5 Mean loss 1.0 0.5 0.0 250 1000 1250 500 750 Wall-clock time (s)