Training loss: workers = 2, τ = 8, β = 1.98 2.0 Mean loss 1.0 0.5 0.0 100 200 400 500 300 Wall-clock time (s)