Training loss: workers = 2, τ = 8, β = 1.98 1.4 1.2 1.0 Mean loss 90 00 00 0.4 0.2 0.0 100 200 300 400 500 Wall-clock time (s)