a9a; heterogeneous; batch = 20; workers = 20 Local SGD 10 local steps ★ Local SGD 20 local steps Local SGD 30 local steps Local SGD 50 local steps 10⁰ Local SGD 100 local steps $||\nabla f(x)||$ 25 50 75 100 125 200 0 150 175 communications