a9a; heterogeneous; batch = 20; workers = 100; local steps = 10Local SGD Local SPIDER Local Adam 10<sup>0</sup> FedAdam  $||\nabla f(x)||$  $10^{-1}$ 50 100 150 200 250 300 0 communications