

Error on  $E[||x|| + ||\nabla f(x)||]$

- SGD vs SGD SDE
- USAM vs SGD SDE
- USAM vs USAM SDE
- SAM vs SGD SDE
- SAM vs SAM SDE

$\rho = \eta$

$\rho = \sqrt{\eta}$

$10^{-3}$

$10^{-2}$

$10^{-1}$

regularization strength  $\rho$

$10^0$

$10^{-1}$

