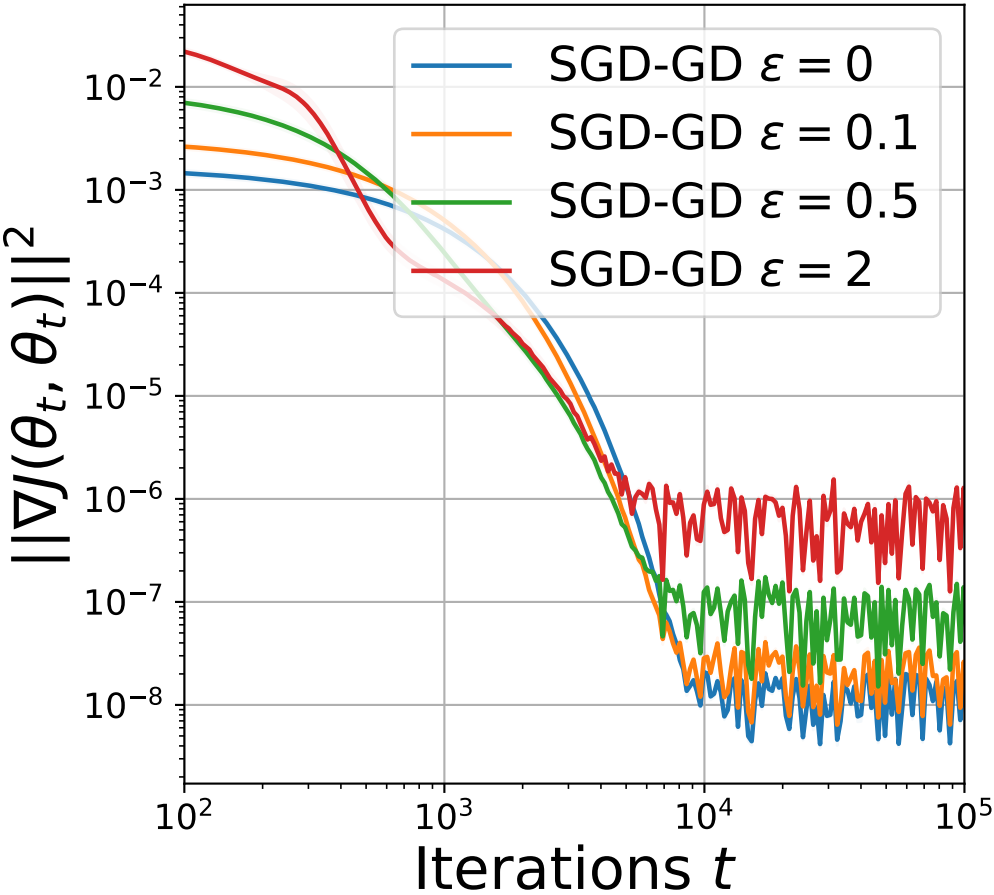
SGD-GD $\varepsilon = 0$
SGD-GD $\varepsilon = 0.1$
SGD-GD $\varepsilon = 0.5$
SGD-GD $\varepsilon = 2$
$||\nabla J(\theta_t, \theta_t)||^2$
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{2}
10^{3}
10^{4}
10^{5}
Iterations t