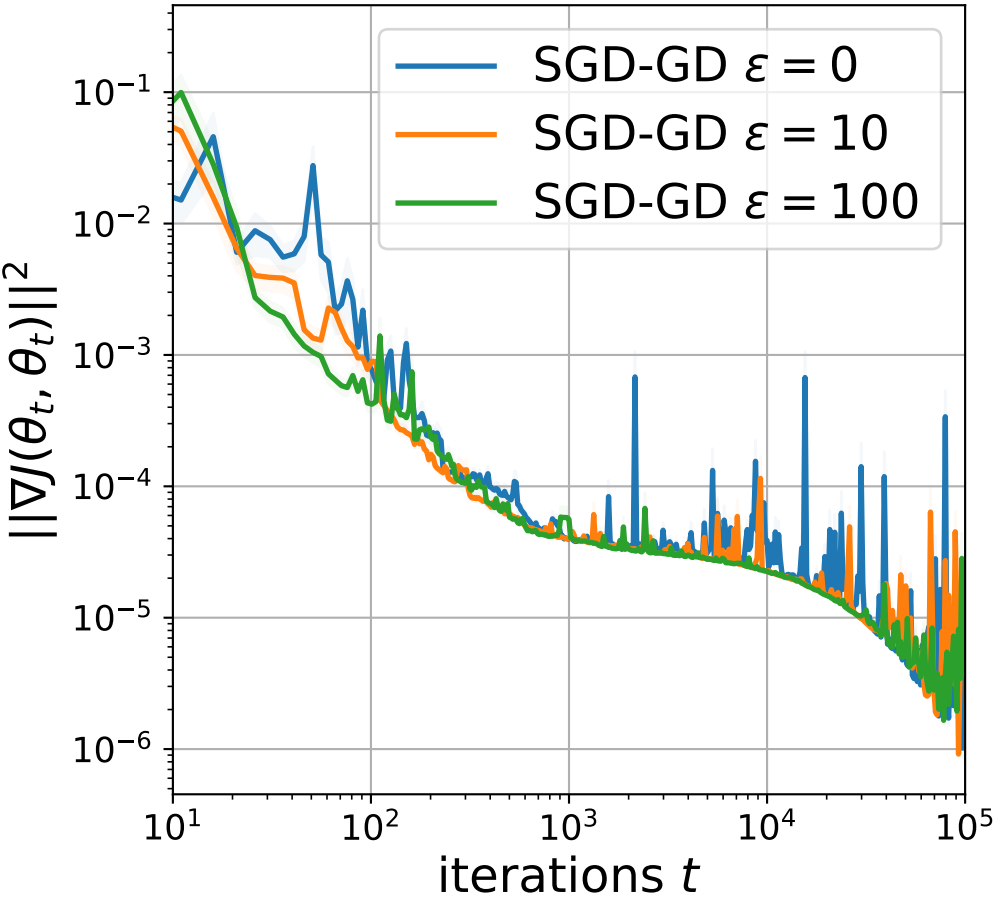
SGD-GD $\varepsilon = 0$
SGD-GD $\varepsilon = 10$
SGD-GD $\varepsilon = 100$
$||\nabla J(\theta_t, \theta_t)||^2$
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{1}
10^{2}
10^{3}
10^{4}
10^{5}
iterations t