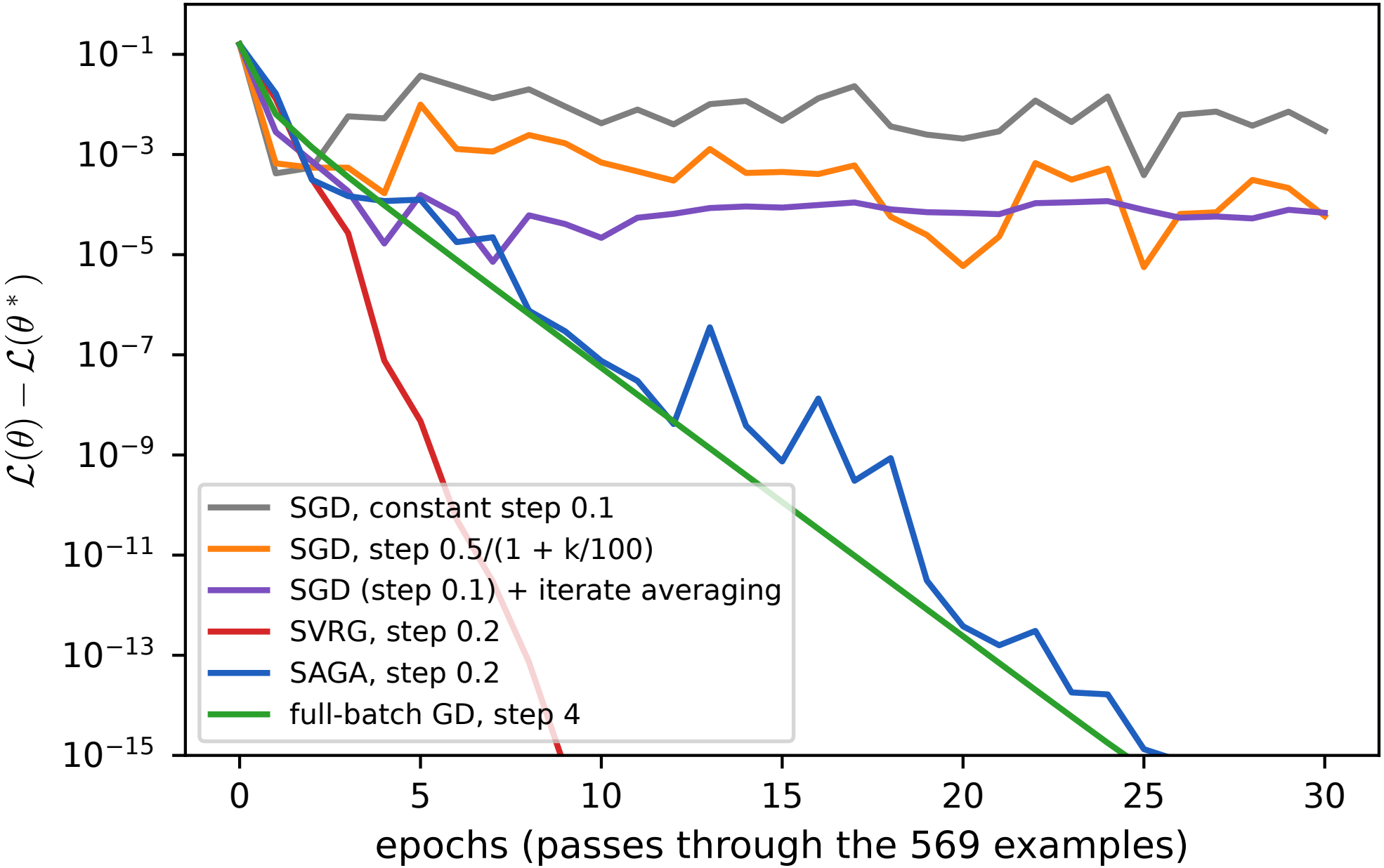

10^{-1}
10^{-3}
10^{-5}
10^{-7}
10^{-9}
10^{-11}
10^{-13}
10^{-15}
$\mathcal{L}(\theta) - \mathcal{L}(\theta^*)$
0
5
10
15
20
25
30
epochs (passes through the 569 examples)
SGD, constant step 0.1
SGD, step 0.5/(1 + k/100)
SGD (step 0.1) + iterate averaging
SVRG, step 0.2
SAGA, step 0.2
full-batch GD, step 4