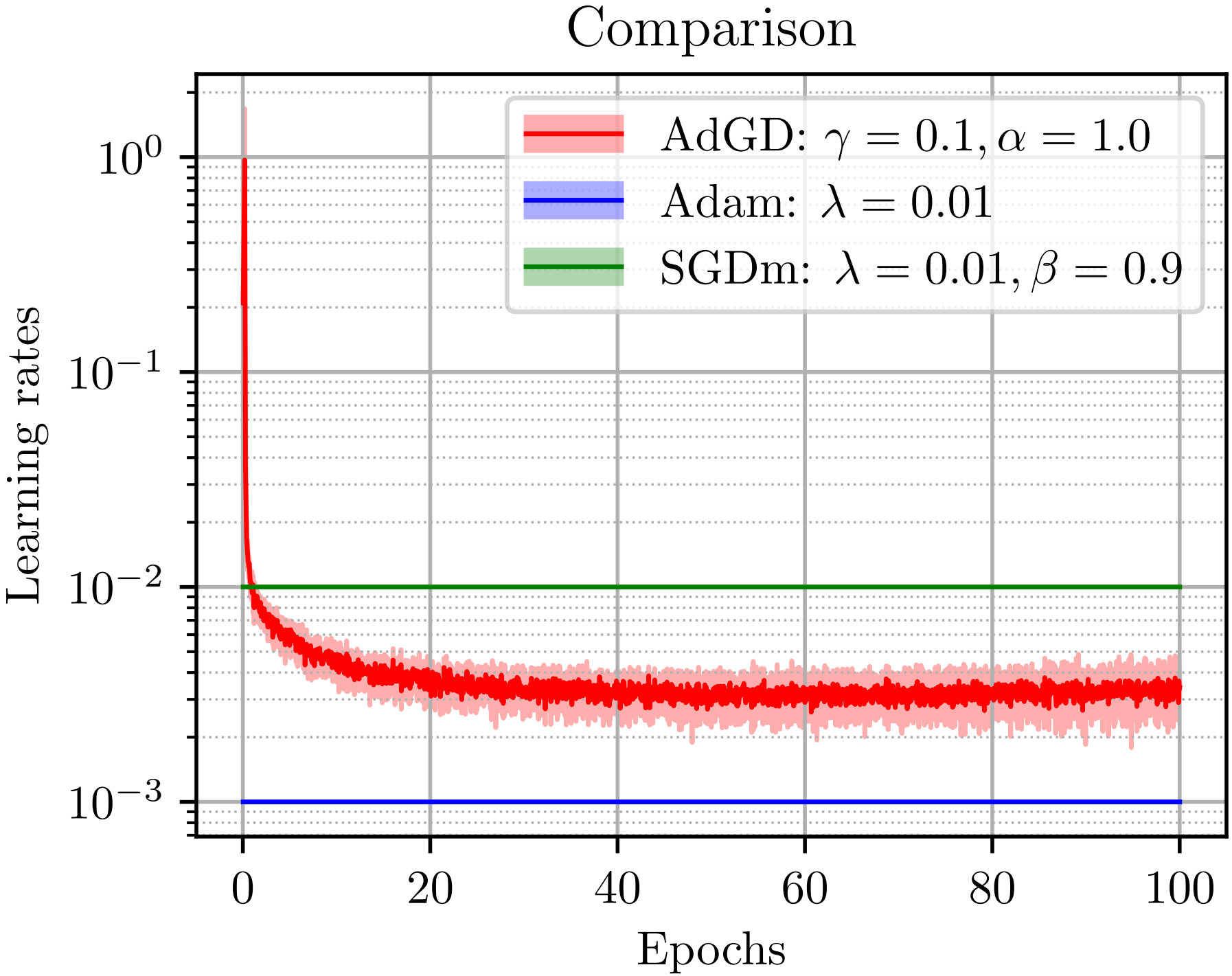
Comparison
AdGD: $\gamma = 0.1, \alpha = 1.0$
Adam: $\lambda = 0.01$
SGDm: $\lambda = 0.01, \beta = 0.9$
Learning rates
10^0
10^{-1}
10^{-2}
10^{-3}
0
20
40
60
80
100
Epochs