train_time 10 adam / tanh / (20, 20) adam / tanh / (50, 50) adam / tanh / (100, 100) adam / relu / (20, 20) 8 adam / relu / (50, 50) adam / relu / (100, 100) lbfgs / tanh / (20, 20) lbfgs / tanh / (50, 50) lbfgs / tanh / (100, 100) 6 lbfgs / relu / (20, 20) lbfgs / relu / (50, 50) lbfgs / relu / (100, 100) 4 2 0 10⁻³ 10^{-2} 10^{-1} 10^{0} 10^{1} alpha