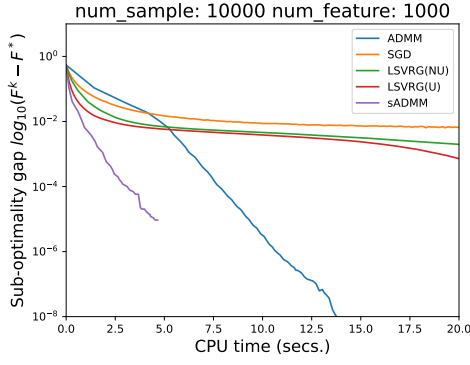
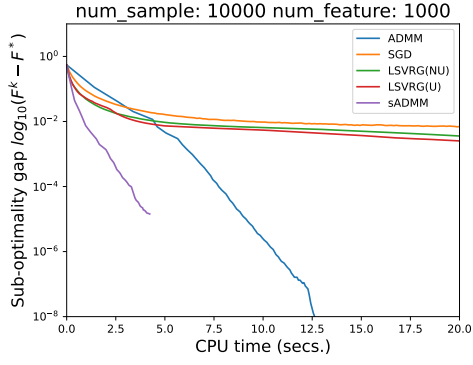
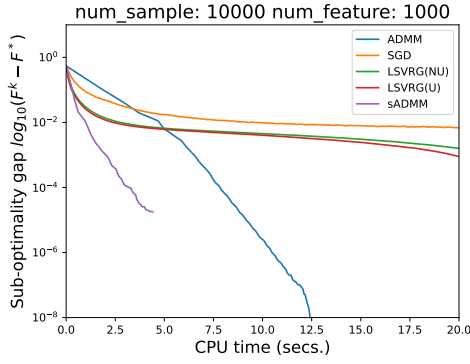




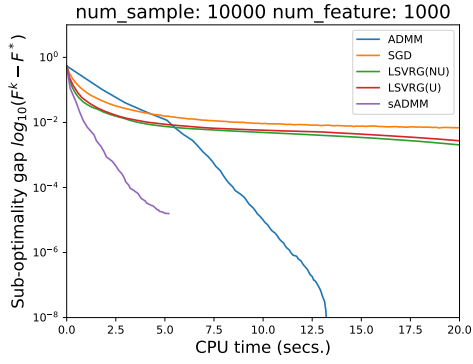
(a)



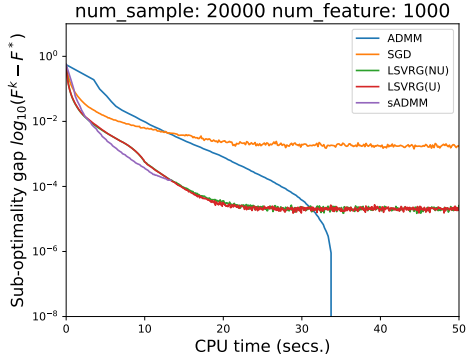
(b)



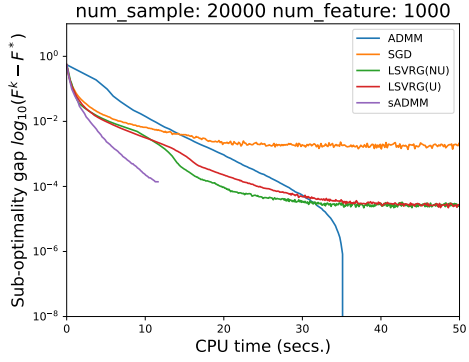
(c)



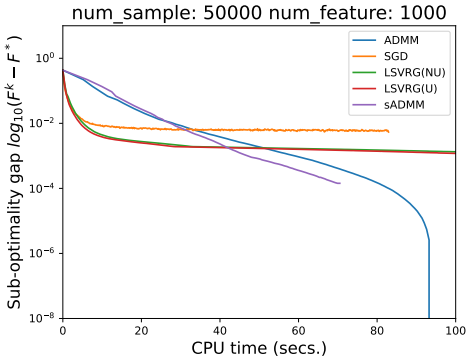
(d)



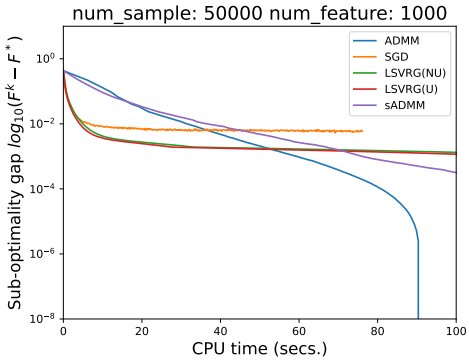
(e)



(f)



(g)



(h)

Figure 1: Time vs. Sub-optimality gap in synthetic dataset with ERM framework and ℓ_1 regularization. The datasets with the same number of samples are generated by different random number seeds. Sub-optimality is defined as $F^k - F^*$, where F^k represents the objective function value at the k -th iteration or epoch and F^* denotes the minimum value obtained by all algorithms. Plots are truncated when $F^k - F^* < 10^{-8}$.