

Here we present additional numerical experiments showcasing the performance of the proposed method in the case of regression problem (17) with respect to reviewers comments.

- Figure 10 contains plots for the regression problem (17) with $i = 1, 2$ and correlated data (setting similar to Fig 1 in the paper). All algorithms are initialized at the same randomly chosen point, which includes the stochastic version of Algorithm 1 as suggested by reviewer YmeS.

- Figure 11 contains plots for the regression problem (17) with $i = 1, 2$ and correlated data (setting similar to Fig 1 in the paper). All algorithms are initialized at the same randomly chosen point. This includes comparisons of different cases of Algorithm 1 with various values of ε , as well as the case where the constant term ε/L_f^2 is dropped, as suggested by reviewer YmeS.

- Figure 12 contains plots for the regression problem (17) with $i = 1, 2$ and correlated data (setting similar to Fig 1 in the paper). All algorithms are initialized at the same randomly chosen point, which includes Algorithm 1 with different update rules for μ_k , as suggested by reviewer gzzq4.

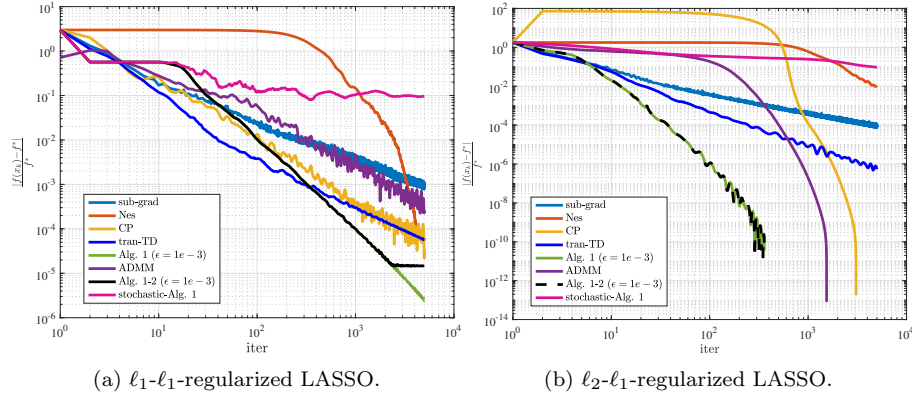


Figure 10. Numerical results for regression problems with stochastic version of Algorithm 1.

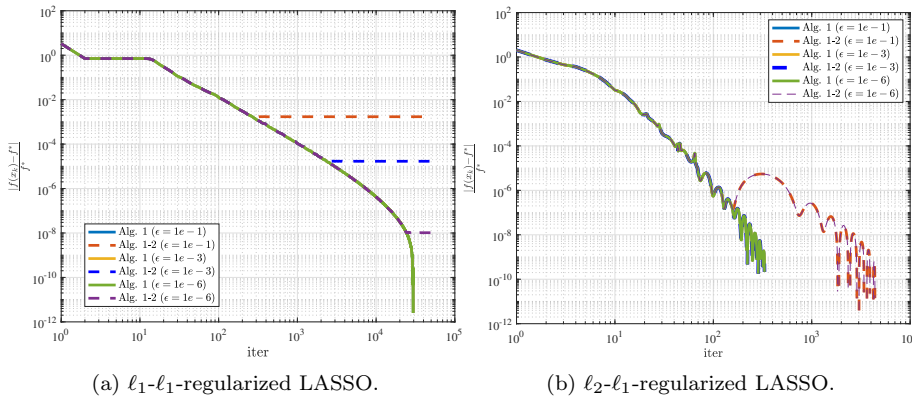


Figure 11. Numerical results for regression problems with different values of ε of Algorithm 1.

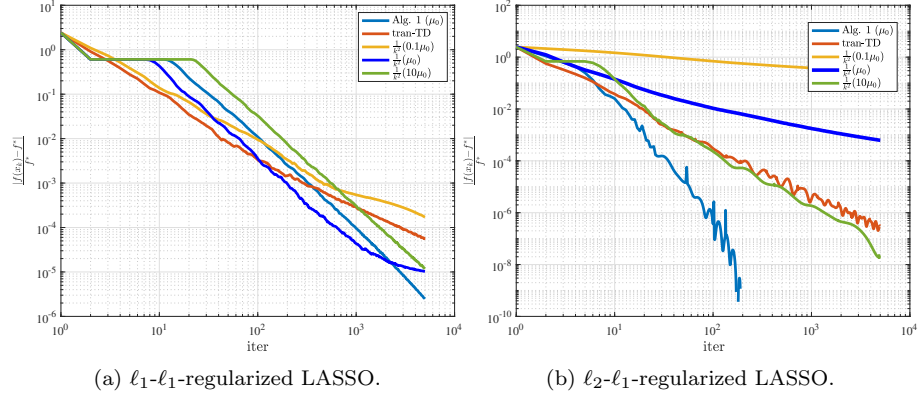


Figure 12. Numerical results for regression problems with different update rule of Algorithm 1.