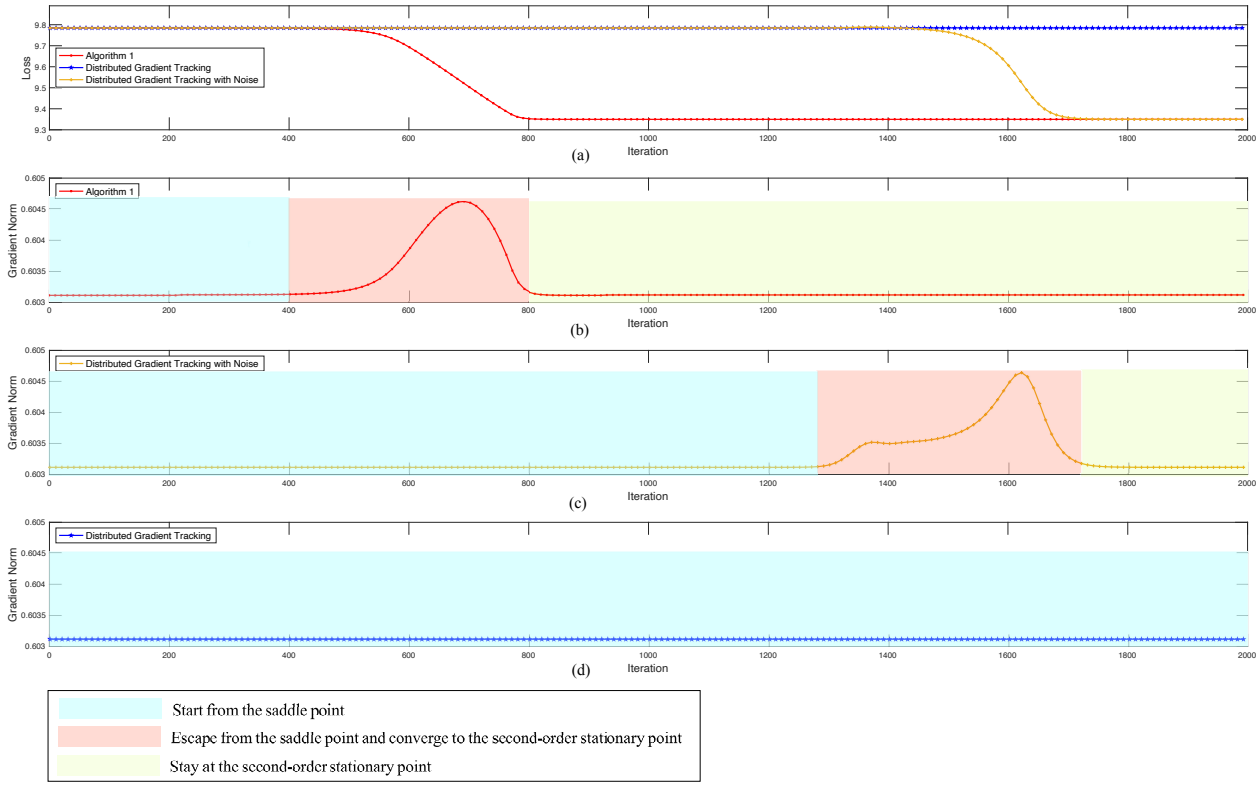

Fig. 4. Comparison using Regularized Logistic Regression under the Gisette dataset. (a) Comparison of loss function values. (b) Starting from a saddle point, our algorithm 1 enables saddle-point escaping from iteration around 400. (Note that the gradient norm at the saddle point is non-negotiable since the dimension of the state variable is as high as 5002.) After escaping from the saddle point, our algorithm enables convergence to a second-order stationary point starting from iteration around 685 (the peak of the curve). Starting from iteration around 800, the states reach the second-order stationary point. (c) Starting from a saddle point, the distributed gradient tracking algorithm with noise enables saddle-point escaping from iteration around 1300. After escaping from the saddle point, the algorithm enables convergence to a second-order stationary point starting from iteration around 1620 (the peak of the curve). Starting from iteration around 1700, the states reach the second-order stationary point. (d) The state variable is trapped at the saddle point all the time.