The main motivation in this dissertation has been to explore how coding will enhance efficiency and scalability in solving distributed optimization problems over networks. Specifically, we focused on developing first-order methods for solving convex optimization problems, where the total loss is defined as a sum of smooth component functions. To this end, coding-based distributed gradient descent algorithms were studied, allowing an efficient, adaptive, scalable, and robust to stragglers implementation. We studied the application of such coded algorithms both on centralized and decentralized (distributed) networks. Although the algorithms' brute application permits robustness to stragglers it also augments the convergence rate by a scalable factor. Moreover, the convergence of the proposed algorithms is released from the dependency on the network topology and molded according to the applied coding method.

First, we have studied a centralized algorithm on parameter server networks that is robust to an allowed number of stragglers. We analyzed convergence rates of the algorithm under an assumption of bounded information delays for synchronous implementations. We also proved the convergence of this algorithm and provided an investigation on its convergence rates under such conditions through simulation. Additionally, we have analyzed two different variants of coded gradient descent methods on distributed networks. We investigated two synchronous algorithms over static distributed networks to solve stochastic optimization problems with smooth loss functions. We further proposed a variant synchronous algorithm on time-varying distributed networks. Our approach is based on the proposition that the use of coding will not only allow robustness against stragglers but also can improve the convergence rate. Subsequently, the constructed methods were analyzed from an analytical perspective to prove their convergence and to investigate their convergence rates. Our evaluations are also assessed from a numerical perspective which provides nominal adherence to the described theory.