Robust Optimization over Networks Using Distributed Restarting of Accelerated Dynamics

Daniel E. Ochoa, Jorge I. Poveda, César A. Uribe, Nicanor Quijano

Abstract

Motivated by the several engineering applications it has in areas such as machine learning, power, transportation, and water distribution systems, and distributed network control, see [1] and references therein, we study the accelerated, efficient and robust solution of accelerated distributed optimization problems over network systems characterized by connected and undirected graphs \( G := (\mathcal{V}, \mathcal{E}) \), where \( \mathcal{V} = \{1, 2, \ldots, n\} \) is the set of nodes, and \( \mathcal{E} \subset \mathcal{V} \times \mathcal{V} \) is the set of edges. Specifically, we consider the setting where each node \( i \) has a local function \( f_i: \mathbb{R}^p \rightarrow \mathbb{R} \), and the network cooperates to find a common point \( z^* \in \mathbb{R}^p \) that minimizes a global function defined as the summation of the local costs. This distributed optimization problem can be written as

\[
\min_{z_1, z_2, \ldots, z_n \in \mathbb{R}^p} \sum_{i=1}^n f_i(z_i), \quad \text{s.t.} \quad z_i = z_j, \quad \forall \ i, j \in \mathcal{V},
\]

which is also known in the literature as the consensus-optimization problem [2].

Discrete-time and continuous-time approaches to solve problem (1) have been extensively studied using gradient descent and Newton-based dynamics [3], [4], primal-dual dynamics [5], and projected dynamics [6], to name just a few. However, a persistent challenge in the solution of problem (1) is to achieve fast rates of convergence without sacrificing essential robustness properties of the algorithms. As recently shown in [7], [8], this task is not trivial given that certain classes of accelerated continuous-time algorithms, such as Nesterov’s ODE [9], [10], [11], can be destabilized under arbitrarily small disturbances on the states or gradients. Since these disturbances are unavoidable in practice, there is an urgent need for the development of robust, accelerated and distributed algorithms for the solution of problem (1).

In the literature of accelerated centralized optimization, one of the approaches that has received significant attention during the last years is the incorporation of restarting techniques. As a matter of fact, as shown in [9], [12], [13], [14], and [15], accelerated algorithms with restarting techniques can achieve exponential convergence rates in strongly convex optimization problems without having perfect knowledge of the condition number of the cost function. Moreover, restarting can also be used to induce suitable robustness properties in the Nesterov’s ODE, provided the combination of the continuous-time dynamics and the discrete-time dynamics is carefully carried out [7]. While these ideas have been explored and validated in centralized optimization problems, as mentioned in [16], it remains an

---

* D. E. Ochoa and J. I. Poveda are with the Department of Electrical, Computer and Energy Eng. at the University of Colorado, Boulder. Email: daniel.ochoa, jorge.poveda@colorado.edu. C. Uribe is with the LIDS lab at the Massachusetts Institute of Technology. Email: cauribe@mit.edu. N. Quijano is with the Department of Electronics Eng. at the University of Los Andes, Colombia. Email: niquijano@uniandes.edu.co.

† Work supported in part by CU Boulder ASIRT Seed Grant, NSF grant CNS-1947613, and the the Yahoo! Faculty Engagement Program.

We design an algorithm to solve the consensus-optimization problem from a complete dual perspective, by constructing dynamics for the dual variable $\mathbf{x} \in \mathbb{R}^{np}$. The proposed accelerated dynamics make use of a momentum variable $\mathbf{y} \in \mathbb{R}^{np}$, and timers $\tau_i$ for each one of the agents. In order to obtain the value of the primal variable $\mathbf{z}$ we make use of the relation $\mathbf{z} = \arg \max_{\mathbf{z} \in \mathbb{R}^{np}} \{\langle \mathbf{Lx}, \mathbf{z} \rangle - F(\mathbf{z})\}$, where $\mathbf{L} = \mathbf{L} \otimes \mathbf{I}_p \in \mathbb{R}^{np \times np}$ and $\mathbf{L}$ is the Laplacian matrix of the communication graph $G$.

Figure 1: We design an algorithm to solve the consensus-optimization problem from a complete dual perspective, by constructing dynamics for the dual variable $\mathbf{x} \in \mathbb{R}^{np}$. The proposed accelerated dynamics make use of a momentum variable $\mathbf{y} \in \mathbb{R}^{np}$, and timers $\tau_i$ for each one of the agents. In order to obtain the value of the primal variable $\mathbf{z}$ we make use of the relation $\mathbf{z} = \arg \max_{\mathbf{z} \in \mathbb{R}^{np}} \{\langle \mathbf{Lx}, \mathbf{z} \rangle - F(\mathbf{z})\}$, where $\mathbf{L} = \mathbf{L} \otimes \mathbf{I}_p \in \mathbb{R}^{np \times np}$ and $\mathbf{L}$ is the Laplacian matrix of the communication graph $G$.
References


