ANALYSIS OF DISTRIBUTED ADMM ALGORITHM FOR CONSENSUS OPTIMIZATION IN PRESENCE OF ERROR

Layla Majzoobi+, Farshad Lahouti*

+School of Electrical Engineering, University of Tehran

*Electrical Engineering Department, California Institute of Technology

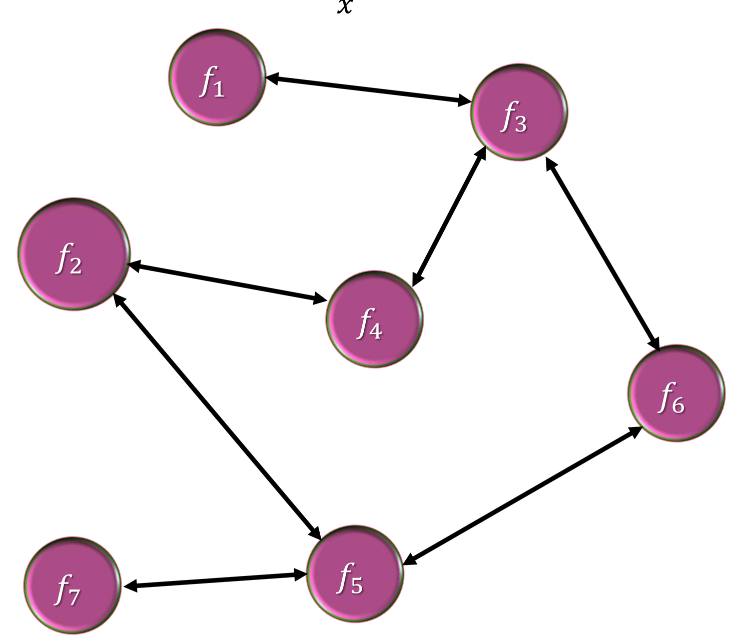
ABSTRACT

ADMM is a popular algorithm for solving convex optimization problems. Applying this algorithm to distributed consensus optimization problem results in a fully distributed iterative solution which relies on processing at the nodes and communication between neighbors. Local computations usually suffer from different types of errors, due to e.g., observation or quantization noise, which can degrade the performance of the algorithm. In this work, we focus on analyzing the convergence behavior of distributed ADMM for consensus optimization in presence of additive node error. We specifically show that (a noisy) ADMM converges linearly under certain conditions and also examine the associated convergence point. Numerical results are provided which demonstrate the effectiveness of the presented analysis.

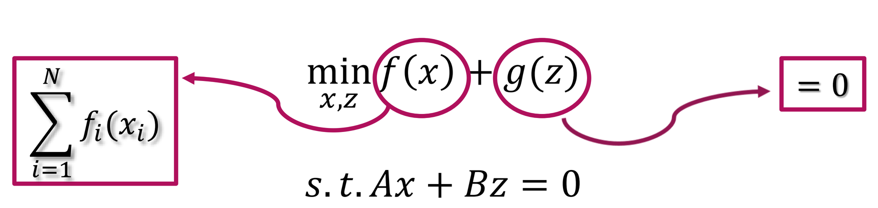
Introduction

- Consensus optimization is a popular distributed optimization problem which arises in various domains and is formulated by

$$\min_{\tilde{x}} . \sum_{i=1}^{N} f_i(\tilde{x})$$



- Considering local variable x_i at each node, variable z_{ij} over each arc (i, j) ,to enforce equality of local variables x_i's, and concatenating x_i's and z_{ij}'s in x *and* z, respectively, the problem can be rewritten in ADMM form



- ADMM based distributed algorithm is proposed to solve consensus problem over a connected graph

Algorithm1
Input functions f_i; Initialization: for all $i \in \mathcal{V}$, set $x_i^0 = \alpha_i^0 = 0$, Set parameter $C > 0$
For all $k = 1, 2, \ldots$ every node i do 1. Update x_i^{k+1} by solving $\nabla f_i(x_i^{k+1}) + \alpha_i^k + 2C\lvert N_i \rvert x_i^{k+1} - C(\lvert N_i \rvert x_i^k + \sum_{j \in N_i} x_j^k) = 0$ 2.Update $\alpha_i^{k+1} = \alpha_i^k + C(\lvert N_i \rvert x_i^k - \sum_{j \in N_i} x_j^k)$

- If f_i's are strongly convex and their gradients are Lipschitz continuous, x^k is R-linearly convergent to its optimal point x^*.

CONVERGENCE ANALYSIS OF NOISY ADMM

The local objective functions in a consensus are usually computed based on some observations and measurements, which are noisy in general. The computed results at a given node are always quantized prior to communication to another node in an iteration. In addition, certain local computations are too complex to carry out exactly and are usually replaced by approximations.

$$\hat{x}_i^k = x_i^k + e_i^k \;\Rightarrow\; \hat{x}^k = x^k + e_x^k \;\Rightarrow\; \hat{z} = z^k + e_z^k$$

Theorem: Consider optimization problem

$$\min_{x,z} f(x) + g(z)$$
$$s.t. Ax + Bz = 0$$

and its optimal points x^*andz^*. Assume thatf_i's (f) are strongly convex functions with moduli m_{f_i} (m_f) and have Lipschitz continuous gradients $\nabla f_i(\nabla f)$ with constant M_{f_i} $(M_f$). if $\|e_z^k\|_2 \le \|x^{k+1} - x^*\|_2$ and $m_f - \frac{c}{2}\sigma_{max}^2(M_+) \ge 0$, then $\forall\, \mu > 1$, $\exists \delta > 0$

$$\|u^{k+1} - u^*\|_G^2 \le \frac{1}{1+\delta}\|u^k - u^*\|_G^2$$

$$\|x^{k+1} - x^*\|_2^2 \le \frac{1}{m_f - \frac{c}{2}\sigma_{max}^2(M_+)}\|u^k - u^*\|_G^2$$

where $u^k \triangleq \binom{z^k}{\beta^k}$, $G = \begin{pmatrix} cI & 0 \\ 0 & \frac{1}{c}I \end{pmatrix}$, $M_- \beta^k = \alpha^k$, and M_+ and M_- are the extended unoriented and oriented network incidence matrices, respectively.

Corollary: If $\|e_x\|_2 = \sigma_e$ and constant over all iterations, an upper bound on the error $\|x^k - x^*\|_2$ for $k \longrightarrow \infty$ is $\sqrt{\max_i \lvert N_i \rvert}\, \sigma_e$.

NUMERICAL RESULTS

We consider Algorithm 1 to solve the optimization problem

$$\min_{\tilde{x}} \sum_{i=1}^{N} \frac{1}{2}\|y_i - M_i\tilde{x}\|_2^2$$

$$\tilde{x} \in \mathbb{R}^3, M_i \in \mathbb{R}^{3\times3} , N = 200$$

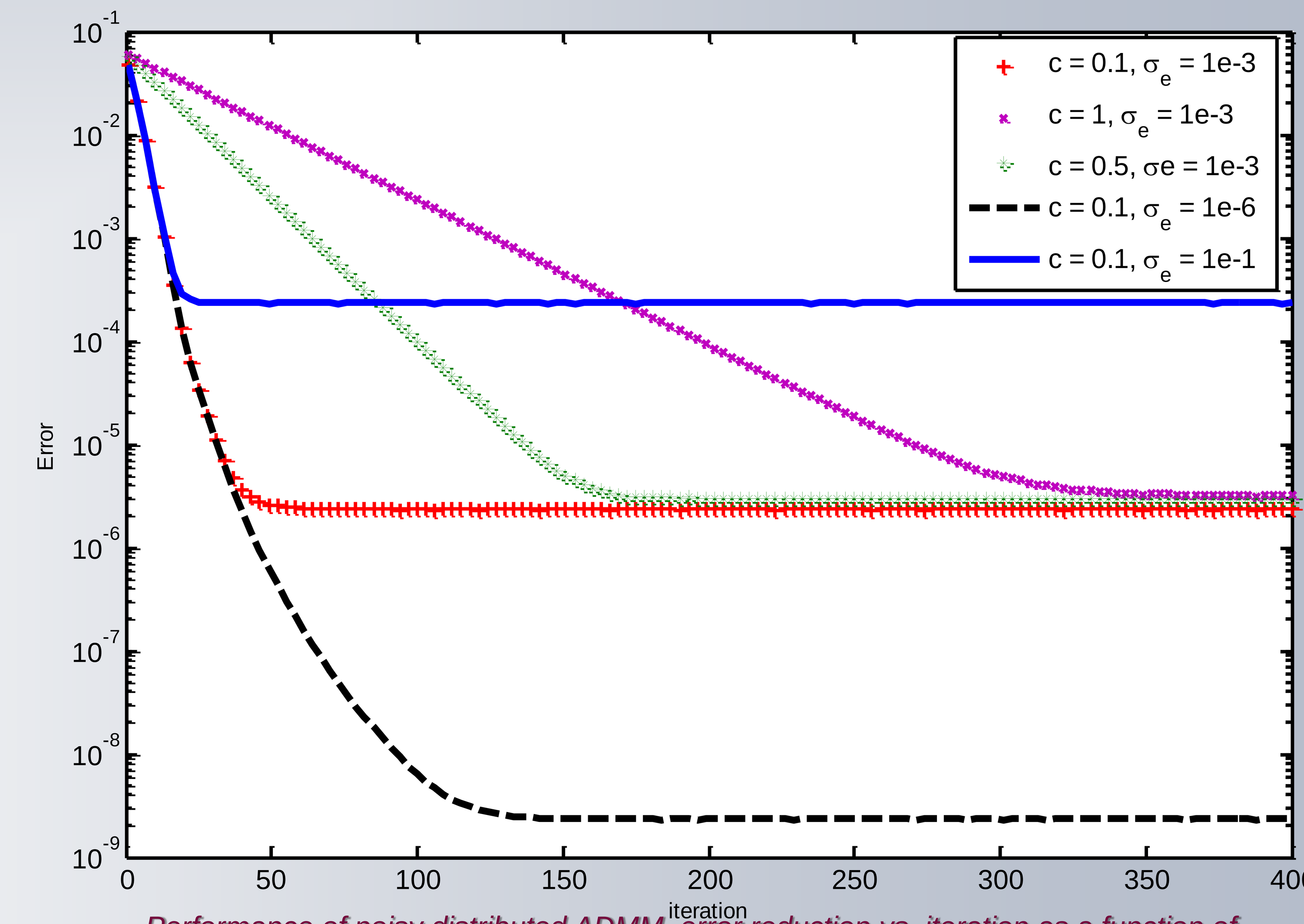
$$y_i = M_i\tilde{x} + n_i \;, n_i \sim \mathcal{N}(0, 10^{-3}I_3)$$

$$e_x = \mathcal{N}(0, \sigma_e)$$

Define connectivity ratio as $\rho = \frac{E}{E_c}$ where E is the number of edges in the network and E_c is the number of edges in a corresponding complete graph.

Consider $\tilde{x}_{i,k}^D$and $\tilde{x}^C$ as respectively the distributed and centralized estimates of $\tilde{x}$ (at node i and iteration$k$). Our performance metrics is error $\mathcal{E}_{i,k}^{D,C} = \frac{\|\tilde{x}_{i,k}^D - \tilde{x}^C\|_2^2}{\|\tilde{x}^C\|_2^2}$.

NUMERICAL RESULTS



Performance of noisy distributed ADMM, error reduction vs. iteration as a function of σ_e and c. ρ=0.04.

CONCLUSIONS

We analyzed the convergence behavior of distributed ADMM for consensus optimization in presence of additive computation error, which is for example due to observation or quantization noise at the nodes. Specifically, we showed that (a noisy) ADMM converges linearly under certain conditions and also examined the associated convergence point. Numerical results in the case of collaborative mean squared error estimation was presented. Next steps of research include analytical assessment of optimized ADMM parameters in this setting and with different system and network parameters.

REFERENCES

[1] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[2] W. Shi, Q. Ling, K. Yuan, G. Wu, and W. Yin, "On the linear convergence of the ADMM in decentralized consensus optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 7, pp. 1750–1761, 2014.

[3] W. Deng and W. Yin, "On the global and linear convergence of the generalized alternating direction method of multipliers," *Journal of Scientific Computing*, pp. 1–28, 2012.

[4] D. Cvetkovi´c, P. Rowlinson, and S. K. Simi´c, "Signless laplacians of finite graphs," *Linear Algebra and its applications*, vol. 423, no. 1, pp. 155–171, 2007.