Communication-Efficient Decentralized Optimization via Double-Communication Symmetric ADMM

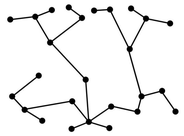
Jinrui Huang[1], Xueqin Wang[1], Dong Liu[1,3], Jingguo Lan[1], Runxiong Wu[2]
[1]University of Science and Technology of China, [2]University of Wisconsin–Madison, [3]Princeton University

MOTIVATION

- Decentralized optimization usually uses one round of communication round per iteration.
- Simply adding more consensus rounds often does not reduce total communication.
- **We ask:** Can fixed multi-round communication reduce overall cost if built into the algorithmic structure?

SETUP

- Decentralized composite optimization
- Each agent only communicates with neighbors over a connected graph



$$\min_{x\in\mathbb{R}^d} F(x) = \sum_{i=1}^{n} \big(f_i(x) + g_i(x)\big)$$

CONTRIBUTION

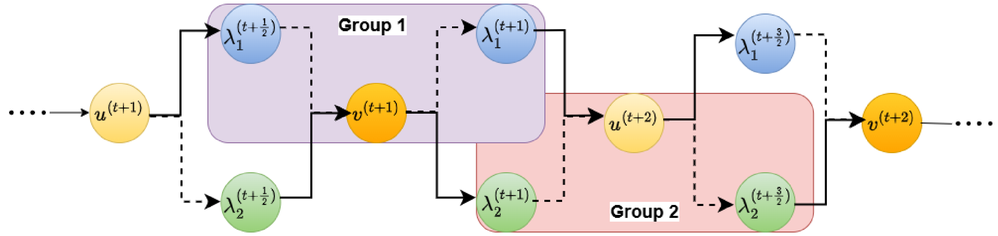
- New **DS-ADMM framework** for decentralized composite optimization
- **Two-round symmetric communication** built into ADMM constraints
- **Linear convergence** under metric subregularity
- Better communication efficiency **in experiments**

METHOD OVERVIEW

- Consensus is encoded through a symmetric pair of constraints: $u = \widetilde{W}v, \qquad v = \widetilde{W}u$
- **Symmetric constraint construction** enables a balanced S-ADMM update and naturally induces multi-round communication.

- **Graph-aware proximal linearization** removes non-separable quadratic coupling and makes updates decentralized.

- **Communication structure designed** to minimize per-iteration overhead:
- Dual variables are split into two half-step updates
- Variables are organized into two update groups
- Transmitted vectors are carefully chosen to reduce communication cost

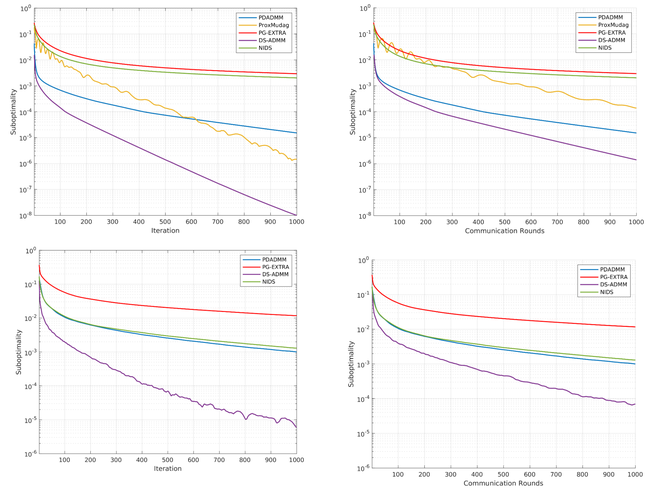


- Only two communication rounds per iteration, each with two transmitted d-vectors.

THEORY

- O(1/t) sublinear convergence in general convex settings
- Linear convergence under metric subregularity
- Applies to broad composite ML models

EXPERIMENTS



- **Robust** across different graph structures, connectivity levels, and network sizes
- **Faster convergence and fewer communication rounds** on LASSO and SVM tasks