Proceedings of the 39th Chinese Control Conference
July 27-29, 2020, Shenyang, China

Primal-dual algorithm for distributed optimization with local domains on signed networks

Xiaoxing Ren[1], Dewei Li[1], Yugeng Xi[1], Lulu Pan [1], Haibin Shao [1]

1. Department of Automation, Shanghai Jiao Tong University, Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai

E-mail: {renxiaoxing, dwli, ygxi, llpan, shore}@sjtu.edu.cn.

Abstract: We consider the distributed optimization problem on signed networks. Each agent has a local function which depends on a subset of the components of the variable and is subject to a local constraint set. A primal-dual algorithm with fixed step size is proposed. The algorithm ensures that the agents' estimates converge to a subset of the components of an optimal solution or its opposite. Note that each component of the variable is allowed to be associated with more than one agents, our algorithm guarantees that those coupled agents achieve bipartite consensus on estimates for the intersection components. Numerical results are provided to demonstrate the theoretical analysis.

Key Words: distributed optimization, multi-agent systems, signed network, convex optimization, primal-dual method.

1 Introduction

Distributed multi-agent optimization has received a variety of attentions due to its ubiquity in scenarios such as power systems [1, 2], estimation over sensor networks [3, 4], distributed learning and regression problems in machine learning [5, 6], etc. There exist many studies of distributed optimization methods, the first-order algorithms have received much attention since they are simple, easy for implementation and generalize, see, e.g., [7–12]. Considering that there may exist various constraints in physical applications, authors of [13–15] developed algorithms to solve distributed constrained optimization problems.

In the above algorithms, each agent has an estimate of the entire variable, which may not be reliable or necessary in some important scenarios such as resource allocation. In this paper, we consider distributed optimization problems where each local function can depend on some components of the variable. There are some works [16–20] consider distributed partitioned problems that are similar to the problem studied in this work. The authors of [16–18] proposed distributed ADMM-based methods. The work [19] developed a parallel coordinate algorithm and the work [20] used an asynchronous dual decomposition method. In this paper, we solve the distributed optimization problems with local domains using primal-dual algorithm.

In the existing studies, a main assumption is that the connections among agents are non-negative. However, in most network settings, there are also negative effects at work. Networks with antagonistic interactions between agents can be represented as signed networks and describe a broad range of systems in the real world [21–25]. Recently the dynamics of signed networks have attracted great attention of the researchers, and it is worth extending distributed optimization problems to signed networks. Only collaboration may not be appropriate in some applications of distributed optimization since some agents would not cooperate. For instance, in the investment problem, the personal relationships between investors can be described as being either positive or negative in effect, evaluation, or trust.

The main contributions of our work are as follows. First, we extend the distributed optimization problem to signed networks. We propose a primal-dual algorithm to solve distributed constrained optimization problems with local domains defined on a signed network. The idea is to manipulate the optimization problem by a gauge transformation so that the primal-dual algorithm can lead agents to find the optimal solution, and agents achieve bipartite consensus with negative interactions among them. Second, we generalize the typical distributed problems to the case where local functions depend on a subset of the components of the variable. Our approach reduces the computation and communication burden in comparison with the algorithms where each agent is interested in all components of the solution regardless of the actual components on which local function depends, the dimension of the variable is also scalable with the network.

The remainder of this paper is organized as follows. First, we provide the preliminaries and problem formulation in Section 2. The distributed primal-dual algorithm is proposed in Section 3. The convergence analysis of the proposed algorithm is provided in Section 4 which is followed by the numerical results in Section 5. Finally, the conclusions are summarized in Section 6.

Notations: Denote x_i as the ith component of a column vector $x \in \mathbb{R}^n$. We use $x = \operatorname{col}\left\{x^i\right\}_{i\in\{1,\dots,n\}}$ to denote a column vector $(x^1,\dots,x^n)^T$ stacked with $x^1,\dots,x^n$. The inner product of two vectors $x, y \in \mathbb{R}^n$ is denoted by $\langle x, y\rangle$ $\|x\| = \sqrt{x^T x}$. The set of all $n\times n$ gauge transformations is given by $\mathsf{D} = \{D = \operatorname{diag}(\sigma), \sigma = [\sigma_1,\dots,\sigma_n], \sigma_i \in \{1,-1\}\}$. For a smooth (differentiable) convex function $f : \mathbb{R}^n \to \mathbb{R}$ and a vector $x \in \mathbb{R}^n$, the gradient of function f at x denoted by $\triangledown f(x)$ satisfies $\langle \triangledown f(x), (\tilde{x}-x)\rangle \le f(\tilde{x}) - f(x)$ for any $\tilde{x} \in \mathbb{R}^n$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous if there exists a constant C such that $\|f(x)-f(y)\| \le C\|x-y\|$, $\forall x, y \in \mathbb{R}^n$. The cardinality of a set Ω with finite elements is denoted by $|\Omega|$. For a nonempty convex set $\Omega \subset \mathbb{R}^n$, we call the point in Ω that is the closest to x the projection of x on Ω and denoted it by $\mathbf{P}_\Omega(x)$. A pair of vectors $x^* \in \Omega$ and $\lambda^* \in \psi$ is called the saddle point of the function $\phi(x,\lambda)$ in

This work is supported by the National Science Foundation of China (Grants No. 61973214, 61590924 and 61963030), the Natural Science Foundation of Shanghai (Grant No. 19ZR1476200).

$\Omega \times \psi$ if

$$\phi(x^*,\lambda) \le \phi(x^*,\lambda^*) \le \phi(x,\lambda^*) \qquad \forall x \in \Omega,\ \lambda \in \psi.$$

2 Preliminaries and Problem Formulation

2.1 Preliminaries

Consider a network of N agents. We assume that the network has a static topology and is represented by an undirected signed graph $G(V,E)$, where $V = \{1,...,N\}$ is the node set, E is the set of undirected edges (i,j). There are no self-loops in the graph, $(i,j) \in E$ if and only if nodes i and j can communicate with each other. A path of G is a concatenation of edges in E: $\{(i_1,i_2),(i_2,i_3),...,(i_{p-1},i_p)\} \subset E$, in which all nodes $i_1,...,i_p$ are distinct. A cycle is a path beginning and ending with the same node $i_1 = i_p$. A cycle is positive if it contains an even number of negative edge weights, it is negative otherwise. A graph is connected if there exists an undirected path between any $i,j \in V$. $\mathcal{N}_i = \{j \in V | (i,j) \in E\}$. The adjacency matrix of $G(V,E)$ is $A = [a_{ij}] \in \mathbb{R}^{N\times N}$, where $a_{ij} = a_{ji} = 0$ if $(i,j) \notin E$, $a_{ij} = a_{ji} = 1$ if edge (i,j) is positive and $a_{ij} = a_{ji} = -1$ if edge (i,j) is negative. $C \in \mathbb{R}^{N\times N}$ is a diagonal matrix of node degrees with entries $C_{i,i} = \sum_{j\in\mathcal{N}_i} |a_{ij}|$, the Laplacian matrix of the graph is $L = C - A = [l_{ij}]$ and

$$l_{i,k} = \begin{cases} \sum_{j\in\mathcal{N}_i} |a_{ij}|, & i = k; \\ -a_{ik}, & i \ne k. \end{cases}$$

Denote $M = [m_{ij}] \in \mathbb{R}^{n\times m}$ as the incidence matrix of undirected signed graph $G(V,E)$, where n and m are the numbers of nodes and undirected edges respectively. The column of M with a positive edge has a 1 in the row related to the node of the edge with smaller index and a -1 in the row related to the node of the edge with larger index, the column of a negative edge has either a 1 or -1 in both rows related to the two nodes of the edge, and zeros elsewhere. Denote the Laplacian matrix of the undirected signed graph by L, we have the following equation

$$L = MM^T. \tag{1}$$

We introduce the following lemmas.

Lemma 1. *([21], Lemma 1) A connected signed graph G is structurally balanced if and only if all cycles of G are positive.*

Lemma 2. *([21], Definition 2) A structurally balanced signed graph admits a bipartition of the agents V_1, V_2, $V_1 \cup V_2 = V$, $V_1 \cap V_2 = \emptyset$, such that $a_{ij} \ge 0\ \forall i,j \in V_q$ ($q \in \{1,2\}$), $a_{ij} \le 0\ \forall i \in V_r, j \in V_q$ ($r,q \in \{1,2\}, r \ne q$).*

Lemma 3. *([26], Proposition 1.1.9) Let C be a nonempty closed convex subset of $\mathbb{R}^n$, and let z be a vector in $\mathbb{R}^n$. There exists a unique vector that minimizes $\|z - x\|$ over x in C, called the projection of z on C. Furthermore, a vector x^* is the projection of z on C if and only if*

$$(z-x^*)^T(x-x^*) \le 0, \qquad x \in C.$$

Lemma 4. *([26], Proposition 1.1.8) Let C be a nonempty convex subset of $\mathbb{R}^n$ and let f be convex and differentiable over an open set that contains C. Then a vector $x^* \in C$ minimizes f over C if and only if*

$$\nabla f(x^*)(z-x^*) \ge 0, \qquad \forall z \in C.$$

2.2 Problem Formulation

We formulate the optimization problem defined on a signed network $G(V,E)$ of N agents. The goal of agents is to solve the following problem.

$$\min_{x\in\mathbb{R}^n} \quad f(x) = \sum_{i=1}^{N} f_i(x_{S_i}) \quad \text{st}: \quad x_{S_i} \in \Omega_i,\ i \in V \tag{2}$$

where Ω_i is a closed constraint convex set of agent i, only agent i knows $f_i : \mathbb{R}^n \to \mathbb{R}$ and Ω_i, $i \in V$. We denote the optimal value of this problem by f^*, which is assumed to be finite. We also denote the nonempty optimal solution set by Ω^*. We use x_{S_i}, $S_i \subseteq \{1,...,n\}$ to denote the components of $x \in \mathbb{R}^n$ that f_i depends on for $i \in V$. For example, if $S_i = \{2,4\}$, then $x_{S_i} = (x_2,x_4)^T$. Moreover, the neighbors of agent i know the set of components x_{S_i} that f_i depends on.

In problem (2), more than one agents may depend on the component x_p, $p \in \{1,...,n\}$. The subgraph induced by x_p is $G_p(V_p,E_p)$, where V_p is the set of agents whose local functions depend on x_p and an edge $(i,j) \in E$ belongs to E_p if both i and j are in V_p.

We introduce our assumptions on the network and problem.

Assumption 1. *(a) The undirected signed graph $G(V,E)$ is structurally balanced. (b) All induced subgraphs $G_p(V_p,E_p)$, $p \in \{1,...,n\}$ are connected.*

Assumption 2. *(a) For all $i \in V$, f_i is convex and smooth (differentiable). (b) For all $i \in V$, ∇f_i is Lipschitz continuous on constraint set Ω_i.*

Assumption 3. *The intersection of local constraints $\bigcap_{i=1}^N \Omega_i$ has at least one relative interior point.*

3 Algorithm Design

We first present problem (5) and demonstrate the relation between its optimal solutions and that of problem (2). We use $x_{S_i}^{i,k} \in \mathbb{R}^n$ to denote the estimates for x_{S_i} by agent i at time k, the estimates for x_p by agent $i \in G_p$ is denoted by x_p^i, $i \in V_p$. Denote $\bar{x}_p = \mathrm{col}\{x_p^i\}_{i\in V_p}$ as the estimates for x_p by all agents in G_p.

Let Assumption 1(a) holds, by Lemma 2 the structurally balanced graph G admits a bipartition of the agents v_1, v_2. All cycles of G are positive by Lemma 1, then it is easy to derive that the subgraphs G_p, $p \in \{1,...,n\}$ are structurally balanced. Thus, each subgraph G_p admits a bipartition of the agents v_{1p}, v_{2p}, where $v_{1p} \subset v_1$. Introduce $\sigma = (\sigma_1,...,\sigma_N)^T \in \mathbb{R}^N$, $\sigma_i \in \{1,-1\}$ for $\sigma_i = 1$ when $i \in v_1$, $\sigma_i = -1$ when $i \in v_2$. We define the gauge transformation matrices $D = \mathrm{diag}(D_p)$, where $D_p = \mathrm{diag}(\sigma_p)$, $\sigma_p = \mathrm{col}\{\sigma_i\}_{i\in V_p}$.

Denote matrix $B = \mathrm{diag}(B_p)$, where B_p, $p \in \{1,...,n\}$ is the transpose of the incidence matrix of subgraph G_p. From (1) we have $L_p = B_p^T B_p$. Thus

$$L = \mathrm{diag}(L_p) = B^T B. \tag{3}$$

We introduce the following optimization problem

$$\min F(X) = \sum_{i=1}^{N} f_i(\sigma_i x_{S_i}^i)$$

$$\text{st}: \quad BX = 0, \quad x^i_{S_i} \in \sigma_i \Omega_i,\ i \in V, \tag{4}$$

where $X = (\bar{x}_1^T, ..., \bar{x}_n^T)^T \in \mathbb{R}^m$, $m = \sum_{i=1}^N |S_i|$.

By the definition of the incidence matrix introduced in Section 2.1 and Assumption 1(b), $BX = 0$ is equivalent with $x^i_p = a_{ij} x^j_p$, where $(i,j) \in E_p$ and $p = \{1, ..., n\}$. Denote $\tilde{\Omega} = \Omega_1 \times ... \times \Omega_N$ the Cartesian product of local constraint sets, $\tilde{X} = (x^T_{S_1}, ..., x^T_{S_N})^T \in \mathbb{R}^m$. The constraints $x^i_{S_i} \in \Omega_i$ for $i \in V$ can be written in a compact form $\tilde{X} \in \tilde{\Omega}$, there is a convex set Ω such that $X \in \Omega$ is equivalent with $\tilde{X} \in \tilde{\Omega}$. Constraints $x^i_{S_i} \in \sigma_i \Omega_i$ for $i \in V$ can be written as $X \in D\Omega$. Then problem (4) is equivalently with

$$\min F(X) = \sum_{i=1}^N f_i(\sigma_i x^i_{S_i})$$

$$\text{st:} \quad X \in D\Omega, \quad BX = 0. \tag{5}$$

$F(X)$ is also a convex smooth function ([26], Proposition 1.1.4).

Lemma 5. *Let Assumption 1 holds. Let $X^* = ((x_1^*)^T, ..., (x_n^*)^T)^T$ be an optimal solution of problem (5), then we have that $\sigma_i x^*_{S_i} = \omega^*_{S_i}$ where $\omega^* \in \Omega^*$ is an optimal solution of problem (2).*

Proof. Introduce the variable Z obtained by the gauge transformation matrix D

$$Z = DX = (\bar{z}_1^T, ..., \bar{z}_n^T)^T.$$

Then $z^i_{S_i} = \sigma_i x^i_{S_i}$, $X = DZ$, rewrite problem (5) as

$$\min \tilde{F}(Z) = \sum_{i=1}^N f_i(z^i_{S_i})$$

$$\text{st:} \quad Z \in \Omega, \quad BDZ = 0. \tag{6}$$

Denote an optimal solution of (6) as Z^*, the related optimal solution of (5) as $X^* = DZ^*$. All subgraphs are structurally balanced by Assumption 1(a). From the former analysis, by the definition of B and D, $BDZ = 0$ indicates that $z^i_p = z^j_p =$ for $(i,j) \in E_p$, $p = \{1, ..., n\}$. So problem (6) is a usual distributed optimization problem on an undirected graph, since all subgraphs are connected by Assumption 1(b), it has the same solutions with problem (2). Thus, $Z^* \in \Omega^*$ and $DX^* \in \Omega^*$, i.e., $\sigma_i x^*_{S_i} = \omega^*_{S_i}$, where $i \in V$ and $\omega^* \in \Omega^*$. □

Define the Lagrange function $\Phi(X, \Lambda)$ and augmented Lagrange function $\tilde{\Phi}(X, \Lambda)$ on constraint set $D\Omega$, $\Lambda \in \mathbb{R}^P$ is the Lagrange multiplier vector where $P = \sum_{p=1}^n |E_p|$,

$$\Phi(X, \Lambda) = F(X) + \langle \Lambda, BX \rangle,$$

$$\tilde{\Phi}(X, \Lambda) = F(X) + \langle \Lambda, BX \rangle + \frac{1}{2} \langle BX, BX \rangle.$$

Then problem (5) can be rewritten as $\inf_{X \in D\Omega} \sup_{\Lambda \in \mathbb{R}^P} \Phi(X, \Lambda)$, while the dual problem is defined as $\sup_{\Lambda \in \mathbb{R}^P} \inf_{X \in D\Omega} \Phi(X, \Lambda)$.

Lemma 6. *([14], Lemma 3.3) Let Assumption 1(b), Assumptions 2 and 3 hold, then $\Phi(X, \Lambda)$ has at least one saddle point in $\Omega \times \mathbb{R}^P$. A pair (X^*, Λ^*) is the primal-dual solution to problem (5) if and only if (X^*, Λ^*) is a saddle point of $\Phi(X, \Lambda)$ in $\Omega \times \mathbb{R}^P$.*

Now we present our distributed primal-dual algorithm.

$$X_{k+1} = \mathrm{P}_{D\Omega}\left\{X_k - \alpha \nabla F(X_k) - \alpha(B^T \Lambda_k + B^T B X_k)\right\}$$

$$= \mathrm{P}_{D\Omega}\left\{X_k - \alpha \nabla F(X_k) - \alpha(B^T \Lambda_k + L X_k)\right\}, \tag{7}$$

$$\Lambda_{k+1} = \Lambda_k + \alpha B X_k, \tag{8}$$

where L is defined in (3), α is a positive fixed step size.

Associate a dual variable λ^{ij}_p $(i < j)$ with the constraint $x^i_p - a_{ij} x^j_p = 0$ in $BX = 0$. Note that $A = [a_{ij}] \in \mathbb{R}^{N \times N}$ is the adjacency matrix of G. Translating (7) into component form, for all $i \in V$ and $p \in S_i$,

$$y^{i,k+1}_p = x^{i,k}_p - \alpha \sigma_i \frac{\partial f_i}{\partial x_p}\Big|_{x_{S_i} = \sigma_i x^{i,k}_{S_i}} - \alpha \sum_{j \in \mathcal{N}_i \cap V_P} (x^{i,k}_p - a_{ij} x^{j,k}_p),$$

$$-\alpha\Big(\sum_{j \in \mathcal{N}_i \cap V_P, j > i} \lambda^{ij,k}_p - \sum_{j \in \mathcal{N}_i \cap V_P, j < i} a_{ij} \lambda^{ji,k}_p\Big)$$

where $\frac{\partial f_i}{\partial x_p}$ is the partial derivative of $f_i(x_{S_i})$ with respect to x_p at $\sigma_i x^{i,k}_{S_i}$. Then each agent i projects its primal estimate onto the local constraint set: $x^{i,k+1}_{S_i} = P_{\sigma_i \Omega_i}\left\{y^{i,k+1}_{S_p}\right\}$.

Translating (8) into component form, for all $p \in \{1, ..., n\}$ and $(i,j) \in E_p$,

$$\lambda^{ij,k+1}_p = \lambda^{ij,k}_p + \alpha(x^{i,k}_p - a_{ij} x^{j,k}_p),$$

this task can be assigned to agent i or j.

Algorithm 1 summarizes the above steps.

Algorithm 1

1: **Initialization:** Depart agents $i \in V$ into two parts v_1 and v_2 such that all edges in each subset are positive while all edges joining agents of different subsets are negative, determine the value of σ_i. Set $x^{i,0}_{S_i} = 0$ for $i \in V$; set $\lambda^{ij,0}_p = 0$ for $(i,j) \in E_p$, $p \in \{1, ..., n\}$; set $k = 1$.

2: **For all** $i \in V$**:** for all $p \in S_i$

$$y^{i,k+1}_p = x^{i,k}_p - \alpha \sigma_i \frac{\partial f_i}{\partial x_p}\Big|_{x_{S_i} = \sigma_i x^{i,k}_{S_i}} - \alpha\Big(\sum_{j \in \mathcal{N}_i \cap V_P, j > i} \lambda^{ij,k}_p$$

$$- \sum_{j \in \mathcal{N}_i \cap V_P, j < i} a_{ij} \lambda^{ji,k}_p\Big) - \alpha \sum_{j \in \mathcal{N}_i \cap V_P} (x^{i,k}_p - a_{ij} x^{j,k}_p),$$

3: **Projection on local constraint set:** for $i \in V$

$$x^{i,k+1}_{S_p} = P_{\sigma_i \Omega_i}\left\{y^{i,k+1}_{S_i}\right\}$$

4: **For all** $(i,j) \in E_p$**,** $p \in \{1, ..., n\}$**:**

$$\lambda^{ij,k+1}_p = \lambda^{ij,k}_p + \alpha(x^{i,k}_p - a_{ij} x^{j,k}_p),$$

5: **Set** $k = k + 1$ and repeat step 2–4 until a predefined stopping criterion (e.g., a maximum iteration number) is satisfied.

---

4 Convergence Analysis

Let (X^*, Λ^*) be a saddle point of $\Phi(X, \Lambda)$ in Ω, by Lemma 6, X^* is an optimal solution to the problem (5), then

$$BX^* = 0. \tag{9}$$

From (8), we derive

$$\Lambda_{k+2} = \Lambda_k + \alpha B X_k + \alpha B X_{k+1}, \tag{10}$$

and thus by (9)

$$\Lambda_k - \Lambda^* = \Lambda_{k+2} - \Lambda^* - \alpha B(X_k - X^*) - \alpha B(X_{k+1} - X^*). \tag{11}$$

Then, by multiplying both sides of (11) with B^T, we obtain

$$\begin{aligned} B^T(\Lambda_k - \Lambda^*) = B^T(\Lambda_{k+2} - \Lambda^*) - \alpha L(X_k - X^*) \\ -\alpha L(X_{k+1} - X^*). \end{aligned} \tag{12}$$

Set

$$T_{k+1} = X_k - \alpha \nabla F(X) - \alpha B^T \Lambda_k - \alpha L X_k - X_{k+1} \tag{13}$$

By (9) and (12), we derive

$$\begin{aligned} X_{k+1} - X^* = (I_m - \alpha L + \alpha^2 L)(X_k - X^*) - T_{k+1} \\ -\alpha B^T(\Lambda_{k+2} - \Lambda^*) - \alpha\left(\nabla F(X^*) + B^T\Lambda^*\right) + \alpha^2 L(X_{k+1} - X^*) \\ -\alpha\left(\nabla F(X_k) - \nabla F(X^*)\right). \end{aligned}$$

Set

$$W = I_m - \alpha L + \alpha^2 L, \tag{14}$$

we obtain the following recursion

$$\begin{aligned} W(X_{k+1} - X_k) + \alpha B^T(\Lambda_{k+2} - \Lambda^*) = -(\alpha L - 2\alpha^2 L)(X_{k+1} - X^*) \\ -T_{k+1} - \alpha\left(\nabla F(X_k) - \nabla F(X^*)\right) - \alpha\left(\nabla F(X^*) + B^T\Lambda^*\right). \end{aligned} \tag{15}$$

Define the Lyapunov function as follows:

$$V(X,\Lambda) = \langle X - X^*, W(X - X^*)\rangle + \langle \Lambda - \Lambda^*, \Lambda - \Lambda^*\rangle.$$

Denote l_r as the Lipschitz constant of $\nabla F(X)$ on convex set $D\Omega$, κ_i as the ith largest eigenvalue of $L = \mathrm{diag}(L_p)$ defined in (3). Since the Laplacian matrix of an undirected structurally balanced signed graph is semi-definite[21], $\kappa_1 \geq \kappa_2 \geq \ldots \geq \kappa_m = 0$.

Theorem 1. *Let Assumptions 1, 2 and 3 hold, $\left\{x_{S_i}^{i,k}\right\}$, $i \in V$ generated by Algorithm 1. If the positive step size satisfies*

$$\alpha \leq \frac{1}{2} \quad and \quad \alpha < \frac{2}{2\kappa_1 + l_r}, \tag{16}$$

then $\left\{\sigma_i x_{S_i}^{i,k}\right\}$ converges to the components of an optimal solution of problem (2) indexed by S_i.

Proof. For $k \geq 0$ we have

$$\begin{aligned} V(X_{k+1},\Lambda_{k+2}) - V(X_k,\Lambda_{k+1}) \\ = -\|\Lambda_{k+2} - \Lambda_{k+1}\|^2 - \langle X_{k+1} - X_k, W(X_{k+1} - X_k)\rangle \\ +2\langle \Lambda_{k+2} - \Lambda_{k+1}, \Lambda_{k+2} - \Lambda^*\rangle + 2\langle X_{k+1} - X^*, W(X_{k+1} - X_k)\rangle. \end{aligned} \tag{17}$$

By (8) and (9)

$$\begin{aligned} \langle \Lambda_{k+2} - \Lambda_{k+1}, \Lambda_{k+2} - \Lambda^*\rangle = \langle \alpha B(X_{k+1} - X^*), \Lambda_{k+2} - \Lambda^*\rangle \\ = \left\langle \alpha B^T(\Lambda_{k+2} - \Lambda^*), X_{k+1} - X^*\right\rangle. \end{aligned} \tag{18}$$

Thus,

$$\begin{aligned} \langle \Lambda_{k+2} - \Lambda_{k+1}, \Lambda_{k+2} - \Lambda^*\rangle + \langle X_{k+1} - X^*, W(X_{k+1} - X_k)\rangle \\ = \left\langle X_{k+1} - X_k, \alpha B^T(\Lambda_{k+2} - \Lambda^*) + W(X_{k+1} - X_k)\right\rangle. \end{aligned} \tag{19}$$

Then by (15), we obtain

$$\begin{aligned} \langle \Lambda_{k+2} - \Lambda_{k+1}, \Lambda_{k+2} - \Lambda^*\rangle + \langle X_{k+1} - X^*, W(X_{k+1} - X_k\rangle \\ = -\left\langle X_{k+1} - X^*, \left(\alpha L - 2\alpha^2 L\right)(X_{k+1} - X^*)\right\rangle \\ -\alpha\langle X_{k+1} - X^*, \nabla F(X_k) - \nabla F(X^*)\rangle \\ -\left\langle X_{k+1} - X^*, T_{k+1} + \alpha(\nabla F(X^*) + B^T\Lambda^*)\right\rangle. \end{aligned} \tag{20}$$

By the definition of the saddle point, we have $\forall X \in D\Omega, \Lambda \in \mathbb{R}^P$

$$\Phi(X^*,\Lambda) \leq \Phi(X^*,\Lambda^*) \leq \Phi(X,\Lambda^*).$$

Therefore X^* minimizes the Lagrange function $F(X) + \left\langle X, B^T\Lambda^*\right\rangle$ over $D\Omega$. Since the Lagrange function is convex on $X \in D\Omega$ for each Λ, from Lemma 4, we have

$$\left\langle \nabla F(X^*) + B^T\Lambda^*, X_{k+1} - X^*\right\rangle \geq 0, \qquad \forall k \geq 0. \tag{21}$$

By (13) T_{k+1} is the projection distance of $X_k - \alpha B^T\Lambda_k - \alpha L X_k$ onto the convex set $D\Omega$, using Lemma 3, we have

$$\langle X_{k+1} - X^*, T_{k+1}\rangle \geq 0.$$

Then we obtain

$$\begin{aligned} \langle \Lambda_{k+2} - \Lambda_{k+1}, \Lambda_{k+2} - \Lambda^*\rangle + \langle X_{k+1} - X^*, W(X_{k+1} - X_k\rangle \\ \leq -\left\langle X_{k+1} - X^*, \left(\alpha L - 2\alpha^2 L\right)(X_{k+1} - X^*)\right\rangle \\ -\alpha\langle X_{k+1} - X^*, \nabla F(X_k) - \nabla F(X^*)\rangle. \end{aligned} \tag{22}$$

Substitute (22) into (17) we have

$$\begin{aligned} V(X_{k+1},\Lambda_{k+2}) - V(X_k,\Lambda_{k+1}) \leq -\|\Lambda_{k+2} - \Lambda_{k+1}\|^2 \\ -\langle X_{k+1} - X_k, W(X_{k+1} - X_k)\rangle \\ -2\left\langle X_{k+1} - X^*, \left(\alpha L - 2\alpha^2 L\right)(X_{k+1} - X^*)\right\rangle \\ -2\alpha\langle X_{k+1} - X^*, \nabla F(X_k) - \nabla F(X^*)\rangle. \end{aligned} \tag{23}$$

All possible eigenvalues of $\alpha L - 2\alpha^2 L$ are $\alpha\kappa_i - 2\alpha^2\kappa_i$, $i \in \{1,\ldots,m\}$. By (16) $\alpha \leq \frac{1}{2}$, then $\alpha L - 2\alpha^2 L$ is a positive semi-definite matrix, thus we have

$$\begin{aligned} V(X_{k+1},\Lambda_{k+2}) - V(X_k,\Lambda_{k+1}) \leq -\|\Lambda_{k+2} - \Lambda_{k+1}\|^2 \\ -\langle X_{k+1} - X_k, W(X_{k+1} - X_k)\rangle \\ -2\alpha\langle X_{k+1} - X^*, \nabla F(X_k) - \nabla F(X^*)\rangle. \end{aligned} \tag{24}$$

Note that l_r is the Lipschitz constant of $\nabla F(X)$ and $xy \leq \frac{x^2}{4} + y^2$, we can derive that $\forall k \geq 0$,

$$-\langle X_{k+1} - X^*, \nabla F(X_k) - \nabla F(X^*)\rangle \leq \frac{l_r}{4}\|X_{k+1} - X_k\|^2.$$

Then by (24) we obtain

$$\begin{aligned} V(X_{k+1},\Lambda_{k+2}) - V(X_k,\Lambda_{k+1}) \leq \|\Lambda_{k+2} - \Lambda_{k+1}\|^2 \\ -\left\langle X_{k+1} - X_k, (W - \frac{\alpha l_r}{2} I_m)(X_{k+1} - X_k)\right\rangle. \end{aligned} \tag{25}$$

The eigenvalues of $W - \frac{\alpha l_r}{2} I_m$ are $1 - \alpha\kappa_i + \alpha^2\kappa_i - \frac{\alpha}{2} l_r$, $i \in \{1,...,m\}$. Since $\alpha < \frac{2}{2\kappa_1 + l_r}$, $W - \frac{\alpha l_r}{2} I_m$ is positive definite. Note that $V(X_k, \Lambda_{k+1})$ is non-negative, we conclude that $V(X_k, \Lambda_{k+1})$ converges.

Summing up both sides of (25) from $k = 0$ to t,

$$V(X_{t+1}, \Lambda_{t+2}) - V(X_0, \Lambda_1) \leq -\sum_{k=0}^{t} \|\Lambda_{k+2} - \Lambda_{k+1}\|^2$$

$$-\sum_{k=0}^{t} \langle X_{k+1} - X_k, (W - \frac{\alpha l_r}{2} I_m)(X_{k+1} - X_k)\rangle.$$

We obtain

$$\sum_{k=0}^{\infty} \langle X_{k+1} - X_k, (W - \frac{\alpha l_r}{2} I_m)(X_{k+1} - X_k)\rangle \leq \infty$$

and

$$\sum_{k=0}^{\infty} \|\Lambda_{k+2} - \Lambda_{k+1}\|^2 \leq \infty$$

by letting $t \to \infty$.

Then, we derive that $\lim_{k\to\infty}(X_{k+1} - X_k) = 0$ since $W - \frac{\alpha l_r}{2} I_m$ is positive definite, and $\lim_{k\to\infty}(\Lambda_{k+1} - \Lambda_k) = 0$. According to the converge of $V(X_k, \Lambda_{k+1})$, we conclude that X_k and Λ_k contain convergent subsequence to some limits X^0 and Λ^0, respectively.

By (7) and (8), we obtain

$$X^0 = \mathrm{P}_{D\Omega}\left\{X^0 - \alpha\nabla F(X^0) - \alpha(B^T\Lambda + LX^0)\right\}, \quad (26)$$

$$BX^0 = 0. \quad (27)$$

Thus,

$$X^0 - \alpha\nabla F(X^0) - \alpha(B^T\Lambda + LX^0) - X^0$$
$$= -\alpha\nabla F(X^0) - \alpha B^T\Lambda - \alpha B^T BX^0 = -\alpha\nabla F(X^0) - \alpha B^T\Lambda.$$

By Lemma 3, we conclude that $\forall X \in D\Omega$

$$\left\langle \nabla F(X^0) + B^T\Lambda, X - X^0 \right\rangle \geq 0.$$

Using the convexity of $\Phi(X, \Lambda^0)$, $\forall X \in D\Omega$

$$\Phi(X, \Lambda^0) \geq \Phi(X^0, \Lambda^0) + \left\langle \nabla F(X^0) + B^T\Lambda^0, X - X^0 \right\rangle$$
$$\geq \Phi(X^0, \Lambda^0).$$

From (27) we see $\Phi(X^0, \Lambda^0) = \Phi(X^0, \Lambda) = F(X^0)$, $\forall\Lambda \in \mathbb{R}^p$. Thus (X^0, Λ^0) is a saddle point of $\Phi(X, \Lambda)$. Thus, by Lemma 6, X^0 is an optimal solution to problem (5). From Lemma 5 we conclude that $\{x_{S_i}^{i,k}\}$ converge to some components of ω^* or $-\omega^*$, namely, $\{\sigma_i x_{S_i}^{i,k}\}$ converges to $\omega_{S_i}^*$ for $i \in V$, where $\omega^* \in \Omega^*$ is an optimal solution of problem (2). □

5 Numerical Results

We consider a problem defined on an undirected signed network of 6 agents, the topology is described in Fig 1. The agents are assigned with convex functions as follows.

$$f_1(x_1, x_2) = x_1^2 + x_2^2 - 20x_1 + x_1x_2,$$

$$f_2(x_2, x_3) = x_2^2 + 2^{x_3} + 12x_2, f_3(x_2, x_3) = x_2^2 + 3^{x_3} + 14x_3,$$

$$f_4(x_1, x_2) = x_1^2 + x_2^2 - x_2, f_5(x_1, x_3) = x_1^2 + x_3^2 + x_1,$$

$$f_6(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 + 2x_1,$$

where $x_1, x_2, x_3 \in \mathbb{R}^1$ are the three components of the variable $x \in \mathbb{R}^3$.

The local constraint sets of 6 agents are as follows. $\Omega_1 = \{(x_1, x_2) : x_1 + x_2 \leq 3\}, \Omega_2 = \{(x_2, x_3) : x_2 \geq -2\}$, $\Omega_3 = \{(x_2, x_3) : x_2 \leq 3\}$, $\Omega_4 = \{(x_1, x_2) : x_1 \leq 2\}, \Omega_5 = \{(x_1, x_3) : x_3 \leq 1\}, \Omega_6 = \{(x_1, x_2, x_3) : x_2 + x_3 \leq 3\}$.

The agents are departed into two subsets $V_1 = \{3, 4, 5, 6\}$, $V_2 = \{1, 2\}$. $\sigma_i = 1$ for $i \in V_1$ and $\sigma_i = -1$ for $i \in V_2$. Denote by $x_{S_i}^{i,k}$ the estimates for components of the optimal solution by agent i at time k, S_i is the set of the index of the components that f_i depends on. Let $\{x_{S_i}^{i,k}\}$ be produced by Algorithm 1 with $\alpha = 0.2$, $x_{S_i}^{i,0} = 0$ for $i \in \{1, ..., 6\}$, $\lambda_p^{ij,0} = 0$ for $p \in \{1, ..., n\}$ and $(i, j) \in E_p$. Define the residual of the optimal condition as $r_k = \sum_{i=1}^{N} \|\sigma_i x_{S_i}^{i,k} - x_{S_i}^*\|$ where $x^* = (2.00, -1.30, -2.37)$ is the optimal solution. The evolution of the estimates $x_p^{i,k}$ and the residual r_k is shown in Fig. 2. From Fig. 2(a), it is seen that agents' estimates converge

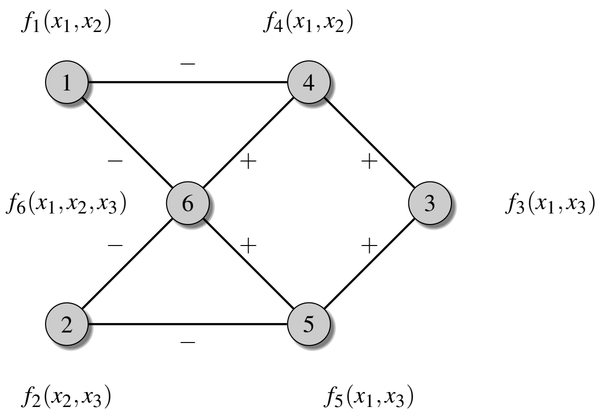


Fig. 1: A signed network consisted of 6 agents, each local function depends a subset of the components of a variable $x = (x_1, x_2, x_3)$.

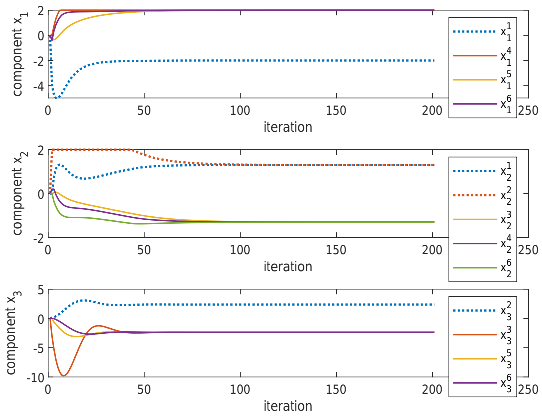
to bipartite consensus and agents of the same subset are in the same side of the bipartition. Moreover, Fig. 2(b) illustrates that $\sigma_i x_{S_i}$, $i \in \{1, ..., 6\}$ converge to the components of the optimal solution x^* that are indexed by S_i, i.e., $x_{S_i}^*$.

6 Conclusion

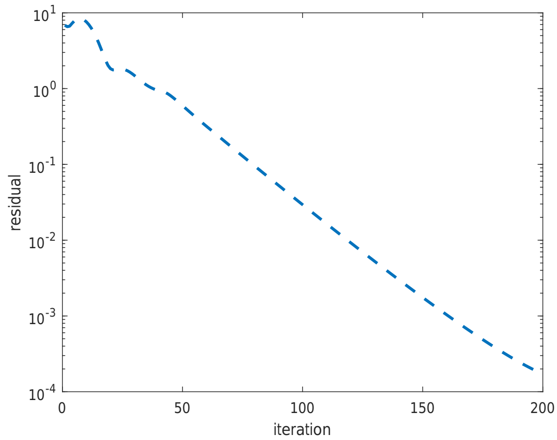
In this paper, we propose a distributed primal-dual algorithm to deal with the multi-agent optimization problem defined on a signed network. The convex function of each agent depends on some components of the variable and is subject to a local constraint set. Our algorithm uses a fixed step size and leads the estimates of agents to converge to the components of an optimal solution x^* or its opposite $-x^*$. The theoretical analysis is justified by a numerical example.

References

[1] L. Ding, L. Y. Wang, G. Yin, W. X. Zheng, and Q.-L. Han, "Distributed energy management for smart grids with an event-triggered communication scheme," *IEEE Transactions on Control Systems Technology*, 2018.



(a)



(b)

Fig. 2: (a) Evolution of the estimates, x_p^i denotes the estimate for x_p^* by agent i. (b)Evolution of the residual r_k.

[2] S. Nabavi, J. Zhang, and A. Chakrabortty, "Distributed optimization algorithms for wide-area oscillation monitoring in power systems using interregional pmu-pdc architectures," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2529–2538, 2015.
[3] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," in *Proceedings of the 3rd international symposium on Information processing in sensor networks*. ACM, 2004, pp. 20–27.
[4] U. A. Khan, S. Kar, and J. M. Moura, "Distributed sensor localization in random environments using minimal number of anchor nodes," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 2000–2016, 2009.
[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "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.
[6] K. I. Tsianos, S. Lawlor, and M. G. Rabbat, "Consensus-based distributed optimization: Practical issues and applications in large-scale machine learning," in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2012, pp. 1543–1550.
[7] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.
[8] A. Nedić and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 601–615, 2015.
[9] D. Jakovetic, D. Bajovic, A. K. Sahu, and S. Kar, "Convergence rates for distributed stochastic optimization over random networks," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 4238–4245.
[10] W. Shi, Q. Ling, G. Wu, and W. Yin, "Extra: An exact first-order algorithm for decentralized consensus optimization," *SIAM Journal on Optimization*, vol. 25, no. 2, pp. 944–966, 2015.
[11] J. Zeng and W. Yin, "Extrapush for convex smooth decentralized optimization over directed networks," *Journal of Computational Mathematics*, vol. 35, no. 4, pp. 383–396, 2017.
[12] C. Xi, R. Xin, and U. Khan, "Add-opt: Accelerated distributed directed optimization," *IEEE Transactions on Automatic Control*, vol. 63, no. 5, pp. 1329–1339, 2018.
[13] A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.
[14] J. Lei, H.-F. Chen, and H.-T. Fang, "Primal–dual algorithm for distributed constrained optimization," *Systems & Control Letters*, vol. 96, pp. 110–117, 2016.
[15] Q. Liu, S. Yang, and Y. Hong, "Constrained consensus algorithms with fixed step size for distributed convex optimization over multiagent networks," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 4259–4265, 2017.
[16] J. F. C. Mota, J. M. F. Xavier, P. M. Q. Aguiar, and M. Puschel, "Distributed optimization with local domains: Applications in mpc and network flows," *IEEE Transactions on Automatic Control*, vol. 60, no. 7, pp. 2004–2009, 2015.
[17] T. Erseghe, "A distributed and maximum-likelihood sensor network localization algorithm based upon a nonconvex problem formulation," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 1, no. 4, pp. 247–258, 2015.
[18] ——, "A distributed and scalable processing method based upon admm," *IEEE Signal Processing Letters*, vol. 19, no. 9, pp. 563–566, 2012.
[19] I. Necoara and D. Clipici, "Parallel random coordinate descent method for composite minimization: Convergence analysis and error bounds," *SIAM Journal on Optimization*, vol. 26, no. 1, pp. 197–226, 2016.
[20] I. Notarnicola, R. Carli, and G. Notarstefano, "Distributed partitioned big-data optimization via asynchronous dual decomposition," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1910–1919, 2018.
[21] C. Altafini, "Consensus problems on networks with antagonistic interactions," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, 2013.
[22] Z. Meng, G. Shi, K. H. Johansson, M. Cao, and Y. Hong, "Modulus consensus over networks with antagonistic interactions and switching topologies," *arXiv preprint arXiv:1402.2766*, 2014.
[23] A. V. Proskurnikov and M. Cao, "Modulus consensus in discrete-time signed networks and properties of special recurrent inequalities," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, Dec 2017, pp. 2003–2008.
[24] G. Paché, "Coopetitive supply chains: Toward a social ties perspective," *Management: journal of contemporary management issues*, vol. 18, no. 2, pp. 125–140, 2013.
[25] S. Wasserman, K. Faust *et al.*, *Social network analysis: Methods and applications*. Cambridge university press, 1994, vol. 8.
[26] D. P. Bertsekas, *Convex optimization theory*. Athena Scientific Belmont, 2009.