2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Surplus-Based ADMM for Distributed Constrained Optimization over Directed Graphs

Qiutong Ji, Guanghui Wen, Tao Yang

Abstract— **This paper introduces a distributed parallel Alternating Direction Method of Multipliers (ADMM) algorithm for solving the distributed constrained optimization problem over directed graphs. To effectively handle the effect of asymmetric information communication on the convergence of the optimization algorithm, a surplus-based averaging consensus algorithm is integrated into the ADMM-based optimization algorithm. Unlike existing distributed ADMM algorithms over directed graphs that focus on the case with solely local constraints, the proposed algorithm can deal with both local constraints and coupling constraints. Under the assumption that the objective function is convex and the underlying graph is strongly connected, the convergence of the surplus-based ADMM to an optimal solution of the distributed constrained problem is proved. Finally, numerical simulations are conducted to validate the effectiveness of the proposed algorithm.**

I. INTRODUCTION

Distributed optimization has garnered increasing attention from various scientific communities, partly due to its wide-ranging applications in areas such as smart grid [1], intelligent transportation systems [2], industrial systems [3], among others. In tackling distributed optimization problems, existing approaches can be generally categorized into two classes: primal [4] and dual-based methods [2]. It is worth noting that dual-based algorithms are particularly well-suited for solving optimization problems with constraints. As one of the most widely used dual-based methods, the Alternating Direction Method of Multipliers (ADMM) has been extensively studied in the literature due to its notable advantages in implementation, convergence, and flexibility [5], [6].

Recently, various distributed ADMM algorithms have been proposed in the literature, see e.g., [7]–[10] and some references therein. For instance, a distributed ADMM algorithm was developed in [7] for solving optimization problems with local polyhedron constraints. Subsequently, a parallel ADMM was designed in [8] to address distributed optimization problems with a global linear constraint and local box constraints over undirected graphs. Furthermore, a kind of parallel ADMM algorithms incorporating a tracking mechanism was introduced in [9] tailored for constraint-coupled convex optimization problems. Building upon this, the tracking-ADMM method was further employed in [10] to address distributed optimization problems with various constraints, including equality constraints, unbounded local constraint sets, and nonlinear inequality coupling constraints. These advancements have significantly contributed to our understanding of utilizing ADMM algorithms for solving distributed constrained optimization problems. However, it is noteworthy that the aforementioned distributed ADMM algorithms are specifically tailored for distributed-constrained optimization over undirected graphs.

In practical scenarios, communication between networking systems often exhibits directed characteristics. For instance, due to resource constraints, sensors typically collect data and transmit it to specific nodes in networks rather than engaging in undirected communication [11]. Additionally, factors like packet loss and communication interference unavoidably lead to directed communication [12]. Partly motivated by the above observation, much effort has been dedicated to developing distributed ADMM algorithms for solving distributed optimization problems over digraphs. Specifically, the distributed ADMM algorithms were proposed in [13] and [14] for strongly convex and smooth objective functions, respectively, where the balancing weights and dynamic average consensus strategy are utilized. A new kind of ADMM algorithm was presented in [15] for general distributed convex optimization problems, which incorporates the push-sum technique to accommodate the effect of directed communication links on the convergence of the algorithm. The approach developed in [15] was further extended in [16] to handle distributed optimization problems with local constraints. Nevertheless, it should be noted that the above-mentioned ADMM-based algorithms over digraphs cannot be directly employed to address the distributed optimization problems with coupled constraints. Constraint-coupled optimization problems are prevalent in practical applications, such as resource allocation constraints [17] and collaboration constraints. Thus, it is crucial to develop distributed ADMM-based algorithms for constraint-coupled optimization over digraphs.

Motivated by the aforementioned discussions, this paper aims to devise a parallel ADMM algorithm for tackling distributed optimization problems with both local and coupling constraints over digraphs. The main contributions are delineated as follows. This paper proposes a new kind of surplus-based ADMM algorithms for solving distributed constrained optimization problems over digraphs. Herein, a

This work was supported in part by the National Natural Science Foundation of China through Grant Nos. 62325304, U22B2046, 62073079, 62088101, and in part by the Jiangsu Provincial Scientific Research Center of Applied Mathematics under Grant No. BK20233002.

Qiutong Ji is with the School of Cyber Science and Engineering, Southeast University, Nanjing 211189, China. 230219294@seu.edu.cn

Guanghui Wen is with the Department of Systems Science, School of Mathematics, Southeast University, Nanjing 211189, China. ghwen@seu.edu.cn

Tao Yang is with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China. yangtao@mail.neu.edu.cn

979-8-3503-1632-2/24/$31.00 ©2024 IEEE

surplus-based consensus method is successfully integrated into the proposed parallel ADMM algorithm, which effectively overcomes the impact of asymmetric information communication over digraphs on the convergence of the algorithm. In contrast to existing algorithms in [13]–[16], which primarily address distributed optimization problems with local constraints, the approach developed in the present paper demonstrates proficiency in addressing distributed optimization with local constraints and global coupling constraints over digraphs.

The rest of this paper is organized as follows. Section II introduces the constrained optimization problem. Section III presents the surplus-based ADMM algorithm. The convergence analysis of the present algorithm is provided in Section IV. A simulation example is presented in Section V. Section VI concludes the paper.

II. PRELIMINARIES AND PROBLEM FORMULATION

A. Graph Theory

A directed graph (digraph) is described as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{\nu_1, \nu_2, \ldots, \nu_N\}$ is a nonempty set of nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of directed edges. The sets of out-neighbours and in-neighbours of node ν_i are denoted by $\mathcal{N}_i^- = \{\nu_j \in \mathcal{V} \,|(\nu_i, \nu_j) \in \mathcal{E}, j \neq i\}$ and $\mathcal{N}_i^+ = \{\nu_j \in \mathcal{V} \,|(\nu_j, \nu_i) \in \mathcal{E}, j \neq i\}$, respectively. The digraph $\mathcal{G}$ is strongly connected if there is a directed path between any two nodes.

B. Parallel ADMM Algorithm

The parallel ADMM is an advanced variant of the ADMM algorithm, which is designed for multi-block convex problems [18]. Specifically, a concise review of the parallel ADMM is provided as follows:

$$\begin{aligned} \underset{x \in C_1, z \in C_2}{\text{minimize}} \quad & G_1(x) + G_2(z) \\ \text{subject to:} \quad & Bx = z, \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^{m_1}$, $z \in \mathbb{R}^{m_2}$, $B \in \mathbb{R}^{m_2 \times m_1}$, $G_1 : \mathbb{R}^{m_1} \to \mathbb{R}$ and $G_2 : \mathbb{R}^{m_2} \to \mathbb{R}$ are convex functions, C_1 and C_2 are convex sets. The augmented Lagrangian function is defined as $L_c(x, z, \lambda) = G_1(x) + G_2(z) + \lambda^{\mathrm{T}}(Bx - z) + (c/2)\|Bx - z\|_2^2$, where $\lambda \in \mathbb{R}^{m_2}$ is a Lagrange multiplier vector. The parallel ADMM is as follows:

$$x_{t+1} = \underset{x \in C_1}{\arg\min}\, L_c(x, z_t, \lambda_t), \tag{2}$$

$$z_{t+1} = \underset{z \in C_2}{\arg\min}\, L_c(x_{t+1}, z, \lambda_t), \tag{3}$$

$$\lambda_{t+1} = \lambda_t + c(Bx_{t+1} - z_{t+1}), \tag{4}$$

where $c > 0$ is a penalty parameter.

C. Problem Formulation

Consider a multi-agent system consisting of N agents and the communication topology of the agents is depicted as a digraph $\mathcal{G}$. The agents aim to collaboratively tackle the following constraint-coupled optimization problem:

$$\begin{aligned} \underset{x_1, \ldots, x_N}{\text{minimize}} \quad & \sum_{i=1}^{N} f_i(x_i) \\ \text{subject to:} \quad & x_i \in \chi_i, \\ & \sum_{i=1}^{N} A_i x_i = b, \end{aligned} \tag{5}$$

where for all $i = 1, \ldots, N$, $x_i \in \mathbb{R}^n$ are decision variables and $f_i : \mathbb{R}^n \to \mathbb{R}$ are the objective functions. The variable x_i is subject to a local set constraint $\chi_i \subseteq \mathbb{R}^n$ and a coupling constraint $\sum_{i=1}^{N} A_i x_i = b$ with $A_i^{\mathrm{T}} \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

To solve the problem (5) using ADMM method, the additional variables $z_i \in \mathbb{R}$ are introduced, $i = 1, \ldots, N$. Through an indicator function $I(x_i)$, the local constraint $x_i \in \chi_i$ can be integrated into the objective function. Then the problem (5) can be equivalently transformed as:

$$\begin{aligned} \underset{x_i \in \mathbb{R}^n, z_i \in \mathbb{R}}{\text{minimize}} \quad & \sum_{i=1}^{N} F_i(x_i) \\ \text{subject to:} \quad & A_i x_i = z_i, \\ & \sum_{i=1}^{N} z_i = b, \end{aligned} \tag{6}$$

where $F_i(x_i) = f_i(x_i) + I(x_i)$ and the indicator function $I(x_i)$ is defined as

$$I(x_i) = \begin{cases} 0 & \text{if } x_i \in \chi_i, \\ \infty & \text{otherwise.} \end{cases} \tag{7}$$

Let $\boldsymbol{x} = \left[x_1^{\mathrm{T}}, \ldots, x_N^{\mathrm{T}}\right]^{\mathrm{T}}$, $\boldsymbol{z} = [z_1, \ldots, z_N]^{\mathrm{T}}$, and $A = diag\{A_1, \ldots, A_N\}$. The Lagrangian function associated with problem (6) is established as follows:

$$L(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\rho}) = \sum_{i=1}^{N} F_i(x_i) + \boldsymbol{\rho}^{\mathrm{T}}(A\boldsymbol{x} - \boldsymbol{z}), \tag{8}$$

where $\boldsymbol{\rho} = [\rho_1, \ldots, \rho_N]^{\mathrm{T}} \in \mathbb{R}^N$ is the Lagrange multiplier vector.

To make problem (5) well-posed, the following assumptions are made.

Assumption 1. *The digraph $\mathcal{G}$ is strongly connected.*

Assumption 2. *For any $i = 1, \ldots, N$, the function f_i is convex. Moreover, the set χ_i is convex and compact.*

Assumption 3. *The saddle point $(\boldsymbol{x}^*, \boldsymbol{z}^*, \boldsymbol{\rho}^*)$ for the Lagrangian function L defined in (8) exists.*

III. SURPLUS-BASED ADMM ALGORITHM

A. Distributed Computation Framework

Regarding $\boldsymbol{x}$ and $\boldsymbol{z}$ as two blocks, let $G_1(\boldsymbol{x}) = \sum_{i=1}^{N} F_i(x_i)$, $G_2(\boldsymbol{z}) = 0$, $C_1 = \mathbb{R}^{Nn}$, and $C_2 = \left\{\boldsymbol{z} \,\middle|\, \sum_{i=1}^{N} z_i = b\right\}$. The augmented Lagrangian function correlated with (6) is

$$L_c(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\rho}) = G_1(\boldsymbol{x}) + \boldsymbol{\rho}^{\mathrm{T}}(A\boldsymbol{x} - \boldsymbol{z}) + \frac{c}{2}\|A\boldsymbol{x} - \boldsymbol{z}\|_2^2. \tag{9}$$

Based on the parallel ADMM method (2)-(4), the algorithm is presented as follows:

$$\boldsymbol{x}_{t+1}=\underset{\boldsymbol{x}\in C_1}{\arg\min}\left\{G_1\left(\boldsymbol{x}\right)+\boldsymbol{\rho}_t^{\mathrm{T}}A\boldsymbol{x}+\frac{c}{2}\left\|A\boldsymbol{x}-\boldsymbol{z}_t\right\|_2^2\right\}, \tag{10}$$

$$\boldsymbol{z}_{t+1}=\underset{\boldsymbol{z}\in C_2}{\arg\min}\left\{-\boldsymbol{\rho}_t^{\mathrm{T}}\boldsymbol{z}+\frac{c}{2}\left\|A\boldsymbol{x}_{t+1}-\boldsymbol{z}\right\|_2^2\right\}, \tag{11}$$

$$\boldsymbol{\rho}_{t+1}=\boldsymbol{\rho}_t+c\left(A\boldsymbol{x}_{t+1}-\boldsymbol{z}_{t+1}\right). \tag{12}$$

Note that the minimization with respect to $\boldsymbol{z}$ in (11) involves a separable quadratic term and an equality constraint. To simplify this update, an optimal solution is derived analytically in [18] as

$$\bar{z}_{i,t+1}=A_ix_{i,t+1}-\bar{h}_{t+1}, \tag{13}$$

where $\bar{h}_{t+1}$ is the average violation of coupling constraints at iteration $t+1$ defined as

$$\bar{h}_{t+1}=\frac{1}{N}\sum_{i=1}^{N}\left(A_ix_{i,t+1}-b_i\right), \tag{14}$$

with $\sum_{i=1}^{N}b_i=b$. It is worth mentioning that the computation of $\bar{h}_{t+1}$ needs to receive quantities $A_ix_{i,t}-b_i$ from all agents over the network. To address this challenge and facilitate the distributed implementation of the algorithm, a surplus-based consensus technique is successfully integrated into the variant ADMM algorithm.

Specifically, given initial values $y_{i,0}=A_ix_{i,t+1}-b_i-s_{i,0}$ with $s_{i,0}\geq 1$, each agent i estimates the average $\bar{h}_{t+1}$ at the iteration $t\in\mathbb{Z}$ as

$$\begin{bmatrix}\boldsymbol{y}_{k+1}\\ \boldsymbol{s}_{k+1}\end{bmatrix}=\begin{bmatrix}\boldsymbol{I}_N-\mathcal{L}(k) & E\\ \mathcal{L}(k) & S-E\end{bmatrix}\begin{bmatrix}\boldsymbol{y}_k\\ \boldsymbol{s}_k\end{bmatrix}, \tag{15}$$

where $\boldsymbol{y}_{k+1}=\left[y_{1,k+1},\ldots,y_{N,k+1}\right]$ denotes the estimation of $\bar{h}_{t+1}$, $\boldsymbol{s}_{k+1}=\left[s_{1,k+1},\ldots,s_{N,k+1}\right]^{\mathrm{T}}\in\mathbb{R}^N$ represents the surplus variable, and $E=diag\{\varepsilon_1,\ldots,\varepsilon_N\}$ with the constant parameter $\varepsilon_i\in(0,1)$, $i=1,\ldots,N$, respectively. Moreover, $\mathcal{L}(k)$ is a matrix with zero row sums and S is a column stochastic matrix. Then, it can be straightforward to obtain that

$$\mathbf{1}_{2N}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{y}_{k+1}\\ \boldsymbol{s}_{k+1}\end{bmatrix}=\mathbf{1}_{2N}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{y}_0\\ \boldsymbol{s}_0\end{bmatrix}=N\bar{h}_{t+1}, \tag{16}$$

which implies that $\mathbf{1}_{2N}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{y}_k^{\mathrm{T}} & \boldsymbol{s}_k^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$ is time-invariant for all k. According to [19], the above surplus-based algorithm ensures that the estimates $y_{i,k+1}, i=1,\ldots,N$, converge to the average $\bar{h}_{t+1}$ as $k\to\infty$ under Assumption 1. Moreover, the following lemma holds.

Lemma 1 [19]. *Suppose that Assumption 1 holds. Define the minimum state $\underline{m}(\boldsymbol{y}_k)=\min_{i\in\mathcal{V}}y_{i,k}$, then the non-decreasing state $\underline{m}(\boldsymbol{y}_k)\leq\bar{h}_{t+1}$, $\forall k\in\mathbb{Z}^+$, $\forall t\geq 0$. Moreover, $\underline{m}(\boldsymbol{y}_k)=\bar{h}_{t+1}$ implies that $y_{i,k}=\bar{h}_{t+1}$ and the nonnegative surplus $s_{i,k}=0$, $\forall i\in\mathcal{V}$, i.e., an average consensus is achieved.*

Denote the estimation of $\bar{h}_{t+1}$ as $h_{i,t+1}$ for agent i. Then, the approximation of $\bar{z}_{i,t+1}$ can be defined as

$$z_{i,t+1}=A_ix_{i,t+1}-h_{i,t+1}. \tag{17}$$

With the matrices defined in (15), it can be observed that the estimation $z_{i,t+1}$ is achieved through a distributed algorithm, where each agent only requires local knowledge from itself and its neighbors rather than global information.

The update of $\boldsymbol{x}_{t+1}$ in (10) is naturally decomposable across x_i when the iteration of $z_{i,t}$ is implemented in a distributed scheme. Thus, each agent computes a minimizer of the following optimization problem in parallel as

$$x_{i,t+1}=\underset{x_i\in\mathbb{R}^n}{\arg\min}\left\{F_i\left(x_i\right)+\rho_{i,t}A_ix_i+\frac{c}{2}\left\|A_ix_i-z_{i,t}\right\|_2^2\right\}. \tag{18}$$

Then, each agent can also update the dual variable by utilizing its own variable $x_{i,t+1}$ and the estimation $z_{i,t+1}$ as

$$\rho_{i,t+1}=\rho_{i,t}+c\left(A_ix_{i,t+1}-z_{i,t+1}\right). \tag{19}$$

B. Algorithm Description

The proposed surplus-based ADMM algorithm is summarized in Algorithm 1.

Algorithm 1: Surplus-based ADMM algorithm

Initialization: Set $x_{i,0}\in\chi_i$, $h_{i,0}=A_ix_{i,0}-b_i$, $\rho_{i,0}=0$.
Repeat for $t=0,1,2,\ldots$ **do**
  for $i=1,\ldots,N$ **do**
    Update $x_{i,t+1}$ based on (18);
    Let $y_{i,0}=A_ix_{i,t+1}-b_i-s_{i,0}$, $s_{i,0}\geq 1$, $D_{i,0}=0$, $flag_{i,0}=0$, $m=1$;
    repeat for $k=0,1,2,\ldots$ **do**
      Update $y_{i,k+1}$ and $s_{i,k+1}$ based on (15);
      $D_{i,k+1}=\max\limits_{j\in\mathcal{N}_i^+}\left\{\left|y_{i,k+1}-y_{j,k}\right|+D_{j,k}\right\}$;
      if $k=md$, $m=1,2,\ldots$ **then**
        if $D_{i,k+1}\leq\frac{1}{2}\sigma_{t+1}$, $s_{i,k+1}\leq\frac{1}{2}\sigma_{t+1}$ **then**
          $flag_{i,k+1}=1$;
        else
          $D_{i,k+1}=0$;
        end
      end
    until $flag_{i,md+1}=1$
    Let $h_{i,t+1}=y_{i,k+1}$, update $z_{i,t+1}$ using (17);
    Update $\rho_{i,t+1}$ based on (19);
  end
Until the stopping criterion is satisfied

In Algorithm 1, to obtain the estimation of $\bar{z}_{i,t+1}$, agents execute an inner loop where they exchange information and update their estimate $y_{i,k+1}$ of $\bar{h}_{t+1}$, $\forall t\geq 0$. Subsequently, the vector $\boldsymbol{z}_{t+1}$ is updated in a distributed manner using (17). Moreover, the variable $D_{i,k}$ represents the radius of the smallest ball enclosing all agent states. It is devised for the purpose of monitoring the convergence of system states. According to Lemma 1, $\underline{m}(\boldsymbol{y}_k)=\bar{h}_{t+1}$, as $k\to\infty$, $\forall t\geq 0$, which indicates $D_{i,k}=0$. At $k=md$, $m\in\mathbb{Z}$, each agent generates a signal $flag_{i,k+1}$ utilized for consensus detection, where d denotes the diameter of $\mathcal{G}$. Define a prescribed tolerance σ_t satisfying $\sigma_{t+1} \leq \sigma_t$ with

$$\sum_{t=1}^{\infty} t\sigma_t < \infty. \tag{20}$$

If $D_{i,md+1} \leq (1/2)\sigma_{t+1}$ and $s_{i,md+1} \leq (1/2)\sigma_{t+1}$, then $flag_{i,md+1} = 1$ which indicates consensus with tolerance is achieved and $flag_{i,md+1} = 0$, otherwise. Without loss of generality, we assume that there exists a finite K such that $|y_{i,K} - y_{j,K}| \leq (1/2)\sigma_{t+1}$, $\forall i, j \in \mathcal{V}$, due to $D_{i,K} \leq (1/2)\sigma_{t+1}$, $s_{i,K} \leq (1/2)\sigma_{t+1}$, and $flag_{i,K} = 1$. Then, from (16), the maximum state $\bar{m}(\boldsymbol{y}_K) \geq \bar{h}_{t+1} - (1/2)\sigma_{t+1}$ can be obtained, where $\bar{m}(\boldsymbol{y}_K) = \max_{i\in\mathcal{V}} y_{i,K}$. Furthermore, as $|y_{i,K} - y_{j,K}| \leq (1/2)\sigma_{t+1}$, $\forall i, j \in \mathcal{V}$, it follows that

$$\underline{m}(\boldsymbol{y}_K) \geq \bar{m}(\boldsymbol{y}_K) - \frac{1}{2}\sigma_{t+1} \geq \bar{h}_{t+1} - \sigma_{t+1}, \tag{21}$$

Then, from Lemma 1, one has

$$\begin{aligned} y_{j,K} &\leq N\bar{h}_{t+1} - (N-1)\underline{m}(\boldsymbol{y}_K) \\ &\leq \bar{h}_{t+1} + (N-1)\sigma_{t+1}, \forall j \in \mathcal{V} \end{aligned} \tag{22}$$

Hence, $\left|\bar{h}_{t+1} - y_{i,K}\right| \leq (N-1)\sigma_{t+1}$. Let $h_{i,t+1} = y_{i,K}$. Then, the obtained vector $\boldsymbol{h}_{t+1} = [h_{1,t+1}, \ldots, h_{N,t+1}]^{\mathrm{T}} \in \mathbb{R}^N$ is an inexact solution to the update (11) satisfying $\left\|\boldsymbol{h}_{t+1} - \bar{h}_{t+1}\mathbf{1}_N\right\| \leq \sqrt{N}(N-1)\sigma_{t+1}$. Furthermore, it can be yielded from (17) that

$$\bar{\boldsymbol{z}}_{t+1} = \boldsymbol{z}_{t+1} - \boldsymbol{e}_{t+1}, \tag{23}$$

where $\bar{\boldsymbol{z}}_{t+1} = [\bar{z}_{1,t+1}, \ldots, \bar{z}_{N,t+1}]^{\mathrm{T}}$, $\boldsymbol{z}_{t+1} = [z_{1,t+1}, \ldots, z_{N,t+1}]^{\mathrm{T}}$, and $\|\boldsymbol{e}_{t+1}\| \leq \sqrt{N}(N-1)\sigma_{t+1}$.

IV. CONVERGENCE ANALYSIS

Before presenting the convergence result and analysis for Algorithm 1, the following lemma is introduced.

Lemma 2. *Let ϕ be a convex function. Given $\tilde{\theta}$ and $\epsilon > 0$, if $\hat{\theta} = \arg\min_{\theta} \phi(\theta) + \frac{\epsilon}{2}\|\theta - \tilde{\theta}\|^2$, there holds that*

$$\frac{2}{\epsilon}\left[\phi(\hat{\theta}) - \phi(\theta)\right] \leq \left\|\theta - \tilde{\theta}\right\|^2 - \left\|\hat{\theta} - \tilde{\theta}\right\|^2 - \left\|\theta - \hat{\theta}\right\|^2. \tag{24}$$

Proof. Define $\psi(\theta) := \phi(\theta) + (\epsilon/2)\left\|\theta - \hat{\theta}\right\|^2$ to be ϵ-strongly convex. From the strongly convex property of objective functions, it can be easily deduced that $\psi(\theta) \geq \psi(\hat{\theta}) + (\epsilon/2)\|\theta - \hat{\theta}\|^2$. According to the definition of $\psi(\theta)$, it is straightforward to obtain that $(2/\epsilon)[\phi(\hat{\theta}) - \phi(\theta)] \leq \|\theta - \tilde{\theta}\|^2 - \|\hat{\theta} - \tilde{\theta}\|^2 - \|\theta - \hat{\theta}\|^2$. □

Theorem 1. *If the consensus tolerance sequence $\{\sigma_t\}_{t\geq1}$ satisfies* (20)*, the sequence $\{(\boldsymbol{x}_t, \boldsymbol{z}_t)\}_{t\geq1}$ generated by Algorithm 1 can converge to an optimal solution $(\boldsymbol{x}^*, \boldsymbol{z}^*)$ of* (6)*, and the sequence $\{\boldsymbol{\rho}_t\}_{t\geq1}$ converges to the dual optimal point $\boldsymbol{\rho}^*$ under Assumptions 1-3.*

Proof. From Lemma 4.1 in [18] and (10) , (11), for $\forall \boldsymbol{x} \in C_1$ and $\forall \bar{\boldsymbol{z}} \in C_2$, one obtains

$$\begin{aligned} &G_1(\boldsymbol{x}_{t+1}) + [\boldsymbol{\rho}_t + c(A\boldsymbol{x}_{t+1} - \boldsymbol{z}_t)]^{\mathrm{T}} A\boldsymbol{x}_{t+1} \\ &\quad \leq G_1(\boldsymbol{x}) + [\boldsymbol{\rho}_t + c(A\boldsymbol{x}_{t+1} - \boldsymbol{z}_t)]^{\mathrm{T}} A\boldsymbol{x}, \end{aligned} \tag{25}$$

and

$$\begin{aligned} &-[\boldsymbol{\rho}_t + c(A\boldsymbol{x}_{t+1} - \bar{\boldsymbol{z}}_{t+1})]^{\mathrm{T}} \bar{\boldsymbol{z}}_{t+1} \\ &\leq -[\boldsymbol{\rho}_t + c(A\boldsymbol{x}_{t+1} - \bar{\boldsymbol{z}}_{t+1})]^{\mathrm{T}} \bar{\boldsymbol{z}}. \end{aligned} \tag{26}$$

Combining (19) and replacing $\boldsymbol{x}$ with $\boldsymbol{x}^*$ in (25) yields

$$\begin{aligned} &G_1(\boldsymbol{x}_{t+1}) + \boldsymbol{\rho}_{t+1}^{\mathrm{T}} A\boldsymbol{x}_{t+1} + c(\boldsymbol{z}_{t+1} - \boldsymbol{z}_t)^{\mathrm{T}} A\boldsymbol{x}_{t+1} \\ \leq &G_1(\boldsymbol{x}^*) + \boldsymbol{\rho}_{t+1}^{\mathrm{T}} A\boldsymbol{x}^* + c(\boldsymbol{z}_{t+1} - \boldsymbol{z}_t)^{\mathrm{T}} A\boldsymbol{x}^*. \end{aligned} \tag{27}$$

Substituting $\bar{\boldsymbol{z}}$ with $\boldsymbol{z}^*$ in (26) produces

$$-\bar{\boldsymbol{\rho}}_{t+1}^{\mathrm{T}} \bar{\boldsymbol{z}}_{t+1} \leq -\bar{\boldsymbol{\rho}}_{t+1}^{\mathrm{T}} \boldsymbol{z}^*, \tag{28}$$

where

$$\bar{\boldsymbol{\rho}}_{t+1} = \boldsymbol{\rho}_t + c(A\boldsymbol{x}_{t+1} - \bar{\boldsymbol{z}}_{t+1}) = \boldsymbol{\rho}_{t+1} + c\boldsymbol{e}_{t+1}. \tag{29}$$

According to (27) and (28), it follows that

$$\begin{aligned} &G_1(\boldsymbol{x}_{t+1}) + \boldsymbol{\rho}_{t+1}^{\mathrm{T}}(A\boldsymbol{x}_{t+1} - A\boldsymbol{x}^*) - \bar{\boldsymbol{\rho}}_{t+1}^{\mathrm{T}}(\bar{\boldsymbol{z}}_{t+1} - \boldsymbol{z}^*) \\ \leq &G_1(\boldsymbol{x}^*) - c(\boldsymbol{z}_{t+1} - \boldsymbol{z}_t)^{\mathrm{T}}(A\boldsymbol{x}_{t+1} - A\boldsymbol{x}^*). \end{aligned} \tag{30}$$

Through Assumption 3 and the saddle point theorem in [18], one has

$$G_1(\boldsymbol{x}^*) \leq G_1(\boldsymbol{x}_{t+1}) + \boldsymbol{\rho}^{*\mathrm{T}}(A\boldsymbol{x}_{t+1} - \bar{\boldsymbol{z}}_{t+1}). \tag{31}$$

According to (31) and the fact $A\boldsymbol{x}^* = \boldsymbol{z}^*$, (30) can be rewritten as

$$\begin{aligned} &(\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}^*)^{\mathrm{T}}(A\boldsymbol{x}_{t+1} - A\boldsymbol{x}^*) - (\bar{\boldsymbol{\rho}}_{t+1} - \boldsymbol{\rho}^*)^{\mathrm{T}}(\bar{\boldsymbol{z}}_{t+1} - \boldsymbol{z}^*) \\ &+ c(\boldsymbol{z}_{t+1} - \boldsymbol{z}_t)^{\mathrm{T}}(A\boldsymbol{x}_{t+1} - A\boldsymbol{x}^*) \leq 0. \end{aligned} \tag{32}$$

Substituting (23) and (29) into (32) yields

$$\begin{aligned} &(\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}^*)^{\mathrm{T}}(A\boldsymbol{x}_{t+1} - \boldsymbol{z}_{t+1}) - c\boldsymbol{e}_{t+1}^{\mathrm{T}}(\boldsymbol{z}_{t+1} - \boldsymbol{e}_{t+1} - \boldsymbol{z}^*) \\ &+ c(\boldsymbol{z}_{t+1} - \boldsymbol{z}_t)^{\mathrm{T}}(A\boldsymbol{x}_{t+1} - A\boldsymbol{x}^*) + \boldsymbol{e}_{t+1}^{\mathrm{T}}(\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}^*) \leq 0. \end{aligned} \tag{33}$$

According to (19) and the fact $A\boldsymbol{x}^* = \boldsymbol{z}^*$, one attains that $A\boldsymbol{x}_{t+1} - \boldsymbol{z}_{t+1} = (\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}_t)/c$ and $A\boldsymbol{x}_{t+1} - A\boldsymbol{x}^* = (\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}_t)/c + (\boldsymbol{z}_{t+1} - \boldsymbol{z}^*)$. Then it follows from (33), (23), and (29) that

$$\begin{aligned} &\frac{1}{c}(\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}^*)^{\mathrm{T}}(\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}_t) + c(\boldsymbol{z}_{t+1} - \boldsymbol{z}_t)^{\mathrm{T}}(\boldsymbol{z}_{t+1} - \boldsymbol{z}^*) \\ &- (\boldsymbol{e}_{t+1} - \boldsymbol{e}_t)^{\mathrm{T}}(c\boldsymbol{z}_{t+1} - c\boldsymbol{z}_t - \boldsymbol{\rho}_{t+1} + \boldsymbol{\rho}_t) + c\|\boldsymbol{e}_t - \boldsymbol{e}_{t+1}\|^2 \\ &+ (\bar{\boldsymbol{z}}_{t+1} - \bar{\boldsymbol{z}}_t)^{\mathrm{T}}(\bar{\boldsymbol{\rho}}_{t+1} - \bar{\boldsymbol{\rho}}_t) + \boldsymbol{e}_{t+1}^{\mathrm{T}}(\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}^*) \\ &- c\boldsymbol{e}_{t+1}^{\mathrm{T}}(\boldsymbol{z}_{t+1} - \boldsymbol{z}^*) \leq 0. \end{aligned} \tag{34}$$

Furthermore, by substituting $\bar{\boldsymbol{z}} = \bar{\boldsymbol{z}}_t$ into (26) and setting $\bar{\boldsymbol{z}} = \bar{\boldsymbol{z}}_{t+1}$ at t, one obtains $-\bar{\boldsymbol{\rho}}_{t+1}^{\mathrm{T}}\bar{\boldsymbol{z}}_{t+1} \leq -\bar{\boldsymbol{\rho}}_{t+1}^{\mathrm{T}}\bar{\boldsymbol{z}}_t$ and $-\bar{\boldsymbol{\rho}}_t^{\mathrm{T}}\bar{\boldsymbol{z}}_t \leq -\bar{\boldsymbol{\rho}}_t^{\mathrm{T}}\bar{\boldsymbol{z}}_{t+1}$. Combining these inequalities yields $(\bar{\boldsymbol{z}}_{t+1} - \bar{\boldsymbol{z}}_t)^{\mathrm{T}}(\bar{\boldsymbol{\rho}}_{t+1} - \bar{\boldsymbol{\rho}}_t) \geq 0$. Thus, from (34) one obtains

$$\begin{aligned} &\frac{1}{2c}\|\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}^*\|^2 + \frac{c}{2}\|\boldsymbol{z}_{t+1} - \boldsymbol{z}^*\|^2 \\ \leq &\frac{1}{2c}\|\boldsymbol{\rho}_t - \boldsymbol{\rho}^*\|^2 + \frac{c}{2}\|\boldsymbol{z}_t - \boldsymbol{z}^*\|^2 - \frac{1}{2c}\|\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}_t\|^2 \\ &- \frac{c}{2}\|\boldsymbol{z}_{t+1} - \boldsymbol{z}_t\|^2 + \|\boldsymbol{e}_{t+1}\|\,\|c\boldsymbol{z}_{t+1} - c\boldsymbol{z}_t - \boldsymbol{\rho}_{t+1} + \boldsymbol{\rho}_t\| \\ &+ \|\boldsymbol{e}_t\|\,\|c\boldsymbol{z}_{t+1} - c\boldsymbol{z}_t - \boldsymbol{\rho}_{t+1} + \boldsymbol{\rho}_t\| + \|\boldsymbol{e}_{t+1}\|\,\|\boldsymbol{\rho}_{t+1} - \boldsymbol{\rho}^*\| \\ &+ c\|\boldsymbol{e}_{t+1}\|\,\|\boldsymbol{z}_{t+1} - \boldsymbol{z}^*\|, \end{aligned} \tag{35}$$

where the relation $2(b_1-b_2)^{\mathrm{T}}(b_1-b_3)=\|b_1-b_2\|^2+\|b_1-b_3\|^2-\|b_2-b_3\|^2$ is employed. With $a_t=\frac{1}{2c}\|\boldsymbol{\rho}_t-\boldsymbol{\rho}^*\|^2+\frac{c}{2}\|\boldsymbol{z}_t-\boldsymbol{z}^*\|^2$, it follows that

$$\begin{aligned}a_{t+1}\leq& a_t+3c\|\boldsymbol{e}_t\|\|\boldsymbol{z}_{t+1}\|+2c\|\boldsymbol{e}_t\|\|\boldsymbol{z}_t\|+3\|\boldsymbol{e}_t\|\|\boldsymbol{\rho}_{t+1}\|\\&+2\|\boldsymbol{e}_t\|\|\boldsymbol{\rho}_t\|+c\|\boldsymbol{e}_t\|\|\boldsymbol{z}^*\|+\|\boldsymbol{e}_t\|\|\boldsymbol{\rho}^*\|.\end{aligned}\tag{36}$$

Under Assumption 2, one can easily know that $x_{i,t}\in\chi_i$ is bounded for all t. Then one knows that $\{\bar{h}_t\}_{t\geq1}$ is also bounded from (14). Based on the inequality $\|\boldsymbol{h}_{t+1}-\bar{h}_{t+1}\mathbf{1}_N\|\leq\sqrt{N}(N-1)\sigma_{t+1}$, it is evident that $\{\boldsymbol{h}_t\}_{t\geq1}$ is bounded as well. Moreover, since $\boldsymbol{z}_t=A\boldsymbol{x}_t-\boldsymbol{h}_t$, the sequence $\{\boldsymbol{z}_t\}_{t\geq1}$ is bounded with an upper bound z_{max}. Additionally, noting the initialization $\boldsymbol{\rho}_0=\mathbf{0}$ in Algorithm 1, $\boldsymbol{\rho}_{t+1}=\boldsymbol{\rho}_t+c\boldsymbol{h}_{t+1}=c\sum_{\tau=0}^{t}\boldsymbol{h}_{\tau+1}$ can be derived from (19). Thus, define $\|\boldsymbol{\rho}_{t+1}\|\leq(t+1)Q$, where the constant Q is relevant to the bound of $\boldsymbol{x}_t$, the maximum of $\boldsymbol{e}_t$, and constant b given in (6). Then, (36) can be further simplified as

$$a_{t+1}\leq a_0+R\sum_{\tau=0}^{t}\sigma_\tau+5Q\sqrt{N}(N-1)\sum_{\tau=0}^{t}\tau\sigma_\tau,$$

where $R=\sqrt{N}(N-1)(5cz_{max}+c\|\boldsymbol{z}^*\|+\|\boldsymbol{\rho}^*\|+3Q)$. By virtue of (20), one knows that a_t is bounded. Based on (37), (35) can be reformulated as

$$\begin{aligned}&\frac{1}{2c}\|\boldsymbol{\rho}_{t+1}-\boldsymbol{\rho}_t\|^2+\frac{c}{2}\|\boldsymbol{z}_{t+1}-\boldsymbol{z}_t\|^2\\&\quad\leq a_t-a_{t+1}+R\sigma_t+5Q\sqrt{N}(N-1)t\sigma_t.\end{aligned}\tag{37}$$

Summing (37) from $t=0$ to t, one attains

$$\begin{aligned}\sum_{\tau=0}^{t}&\left(\frac{1}{2c}\|\boldsymbol{\rho}_{\tau+1}-\boldsymbol{\rho}_\tau\|^2+\frac{c}{2}\|\boldsymbol{z}_{\tau+1}-\boldsymbol{z}_\tau\|^2\right)\leq a_0-a_{t+1}\\&+R\sum_{\tau=0}^{t}\sigma_\tau+5Q\sqrt{N}(N-1)\sum_{\tau=0}^{t}\tau\sigma_\tau.\end{aligned}\tag{38}$$

As $t\to\infty$ for (20) along with the boundedness of a_t, it is easy to get

$$\sum_{\tau=0}^{\infty}\left(\frac{1}{2c}\|\boldsymbol{\rho}_{\tau+1}-\boldsymbol{\rho}_\tau\|^2+\frac{c}{2}\|\boldsymbol{z}_{\tau+1}-\boldsymbol{z}_\tau\|^2\right)<\infty,\tag{39}$$

which implies $\lim_{t\to\infty}\|\boldsymbol{\rho}_{t+1}-\boldsymbol{\rho}_t\|=0$ and $\lim_{t\to\infty}\|\boldsymbol{z}_{t+1}-\boldsymbol{z}_t\|=0$. Hence, based on (19) and (17), $\lim_{t\to\infty}\|A\boldsymbol{x}_t-\boldsymbol{z}_t\|=0$ and $\lim_{t\to\infty}\|\boldsymbol{h}_t\|=0$ hold. Moreover, $\|\bar{\boldsymbol{\rho}}_{t+1}-\boldsymbol{\rho}_t\|\leq\|\bar{\boldsymbol{\rho}}_{t+1}-\boldsymbol{\rho}_{t+1}\|+\|\boldsymbol{\rho}_{t+1}-\boldsymbol{\rho}_t\|\leq c\sqrt{N}(N-1)\sigma_{t+1}+\|\boldsymbol{\rho}_{t+1}-\boldsymbol{\rho}_t\|\to0$, $\|A\boldsymbol{x}_{t+1}-A\boldsymbol{x}_t\|\leq\|A\boldsymbol{x}_{t+1}-\boldsymbol{z}_{t+1}\|+\|\boldsymbol{z}_{t+1}-\boldsymbol{z}_t\|+\|\boldsymbol{z}_t-A\boldsymbol{x}_t\|\to0$, and $\|A\boldsymbol{x}_{t+1}-\bar{\boldsymbol{z}}_{t+1}\|\leq\|A\boldsymbol{x}_{t+1}-\boldsymbol{z}_{t+1}\|+\|\boldsymbol{z}_{t+1}-\bar{\boldsymbol{z}}_{t+1}\|\leq\|A\boldsymbol{x}_{t+1}-\boldsymbol{z}_{t+1}\|+\sqrt{N}(N-1)\sigma_{t+1}\to0$ as $t\to\infty$. Thus, it can be yielded from (13) that $\lim_{t\to\infty}\|\bar{h}_{t+1}\|=0$.

With $\hat{\theta}=\boldsymbol{x}_{t+1}$, $\tilde{\theta}=\boldsymbol{z}_t$, and $\theta=\boldsymbol{x}$, it follows from Lemma 2 and (10) that

$$\begin{aligned}&\frac{2\gamma}{c}\Big[L(\boldsymbol{x}_{t+1},\boldsymbol{z}_t,\boldsymbol{\rho}_t)-L(\boldsymbol{x},\boldsymbol{z}_t,\boldsymbol{\rho}_t)\Big]\leq\|A\boldsymbol{x}-\boldsymbol{z}_t\|^2\\&\quad-\|A\boldsymbol{x}_{t+1}-\boldsymbol{z}_t\|^2-\|A\boldsymbol{x}-A\boldsymbol{x}_{t+1}\|^2.\end{aligned}\tag{40}$$

where γ represents the largest eigenvalue of $A^{\mathrm{T}}A$. Similarly, according to Lemma 2, one can derive that

$$\begin{aligned}&\frac{2}{c}\Big[L(\boldsymbol{x}_{t+1},\bar{\boldsymbol{z}}_{t+1},\boldsymbol{\rho}_t)-L(\boldsymbol{x}_{t+1},\boldsymbol{z},\boldsymbol{\rho}_t)\Big]\leq\|\boldsymbol{z}-A\boldsymbol{x}_{t+1}\|^2\\&\quad-\|\bar{\boldsymbol{z}}_{t+1}-A\boldsymbol{x}_{t+1}\|^2-\|\boldsymbol{z}-\bar{\boldsymbol{z}}_{t+1}\|^2.\end{aligned}\tag{41}$$

Summing the inequalities (40) and (41) yields

$$\begin{aligned}&L(\boldsymbol{x}_{t+1},\bar{\boldsymbol{z}}_{t+1},\boldsymbol{\rho}_t)-L(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\rho}_t)\\\leq&\frac{c}{2\gamma}\left(\|A\boldsymbol{x}-\boldsymbol{z}_t\|^2-\|A\boldsymbol{x}_{t+1}-\boldsymbol{z}_t\|^2-\|A\boldsymbol{x}-A\boldsymbol{x}_{t+1}\|^2\right)\\&+\frac{c}{2}\left(\|\boldsymbol{z}-A\boldsymbol{x}_{t+1}\|^2-\|\bar{\boldsymbol{z}}_{t+1}-A\boldsymbol{x}_{t+1}\|^2-\|\boldsymbol{z}-\bar{\boldsymbol{z}}_{t+1}\|^2\right).\end{aligned}\tag{42}$$

Considering the limit on both sides of the above inequality, it can be derived that

$$L(\boldsymbol{x}_\infty,\boldsymbol{z}_\infty,\boldsymbol{\rho}_\infty)\leq L(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\rho}_\infty).\tag{43}$$

Through Lemma 2 and the equality $\bar{\rho}_{i,t+1}=\rho_{i,t}+c(A_ix_{i,t+1}-\bar{z}_{i,t+1})$, one has

$$\begin{aligned}&2c\big[L(\boldsymbol{x}_{t+1},\bar{\boldsymbol{z}}_{t+1},\boldsymbol{\rho})-L(\boldsymbol{x}_{t+1},\bar{\boldsymbol{z}}_{t+1},\bar{\boldsymbol{\rho}}_{t+1})\big]\\&\quad\leq\|\boldsymbol{\rho}-\boldsymbol{\rho}_t\|^2-\|\bar{\boldsymbol{\rho}}_{t+1}-\boldsymbol{\rho}_t\|^2-\|\boldsymbol{\rho}-\bar{\boldsymbol{\rho}}_{t+1}\|^2.\end{aligned}\tag{44}$$

As $t\to\infty$, it can be inferred that

$$L(\boldsymbol{x}_\infty,\boldsymbol{z}_\infty,\boldsymbol{\rho})\leq L(\boldsymbol{x}_\infty,\boldsymbol{z}_\infty,\boldsymbol{\rho}_\infty).\tag{45}$$

Under Assumption 3, by replacing $\boldsymbol{x}=\boldsymbol{x}^*$ and $\boldsymbol{z}=\boldsymbol{z}^*$ in (43), and $\boldsymbol{\rho}=\boldsymbol{\rho}^*$ in (45), we have

$$\begin{aligned}&L(\boldsymbol{x}^*,\boldsymbol{z}^*,\boldsymbol{\rho}^*)\leq L(\boldsymbol{x}_\infty,\boldsymbol{z}_\infty,\boldsymbol{\rho}^*)\leq L(\boldsymbol{x}_\infty,\boldsymbol{z}_\infty,\boldsymbol{\rho}_\infty)\\&\leq L(\boldsymbol{x}^*,\boldsymbol{z}^*,\boldsymbol{\rho}_\infty)\leq L(\boldsymbol{x}^*,\boldsymbol{z}^*,\boldsymbol{\rho}^*).\end{aligned}\tag{46}$$

Hence, one can attain $L(\boldsymbol{x}^*,\boldsymbol{z}^*,\boldsymbol{\rho}^*)=L(\boldsymbol{x}_\infty,\boldsymbol{z}_\infty,\boldsymbol{\rho}_\infty)$, which implies that $G_1(\boldsymbol{x}_\infty)=G_1(\boldsymbol{x}^*)$. □

V. NUMERICAL SIMULATIONS

In this section, consider the following constrained least squares problem:

$$\begin{aligned}\underset{x_i\in\mathbb{R}^n}{\text{minimize}}\ &\frac{1}{2}\sum_{i=1}^{4}\|\omega_ix_i-\delta_i\|^2\\\text{subject to}:\ &x_i^{\min}\leq x_i\leq x_i^{\max},\\&\sum\nolimits_{i=1}^{4}A_ix_i=b,\end{aligned}\tag{47}$$

where $x_i=[x_{i1},x_{i2}]^{\mathrm{T}}\in\mathbb{R}^2$ is the unknown quantity to be estimated. The entries of the measurement matrix $\omega_i\in\mathbb{R}^{q\times2}$ and the measured data $\delta_i\in\mathbb{R}^q$ with $q=5$ are selected from i.i.d standard normal distribution $\mathcal{N}(0,1)$. Set the upper and lower bounds of $\boldsymbol{x}$ as $\boldsymbol{x}^{\min}=[0.1,0.2,0.3,0.1,0.5,0.1,0.2,0.3]^{\mathrm{T}}$ and $\boldsymbol{x}^{\max}=[1.1,2,1,1,1.5,1.5,1.3,1.8]^{\mathrm{T}}$, and the parameters as $A_1=[1,1]$, $A_2=[2,2]$, $A_3=[1,2]$, $A_4=[2,1]$, and $b=4$. Moreover, the communication digraph between $N=4$ agents is depicted in Fig. 1.

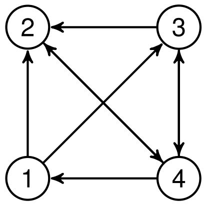

Fig. 1. The digraph $\mathcal{G}$ between four agents.

Under the given parameter settings, the simulation results are illustrated in Figs. 2 and 3. Specifically, from Fig. 2, it can be observed that the states x_i reach the optimal solution $[0.1, 0.2, 0.3, 0.1, 0.5, 0.49, 0.56, 0.3]^{\mathrm{T}}$ for all $i = 1, \cdots, 4$ under Algorithm 1. Additionally, the states x_i consistently remain within the local set constraints. Moreover, from Fig. 3, it is evident that the coupling constraint $\sum_{i=1}^{4} A_i x_i = 4$ is satisfied after approximately 30 iterations.

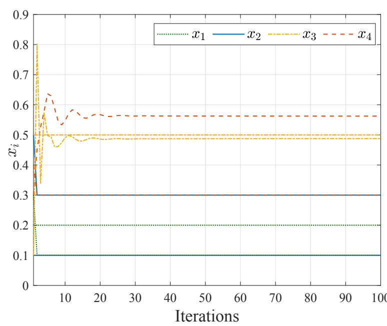

Fig. 2. The evolution of states x_i for $i = 1, \cdots, 4$.

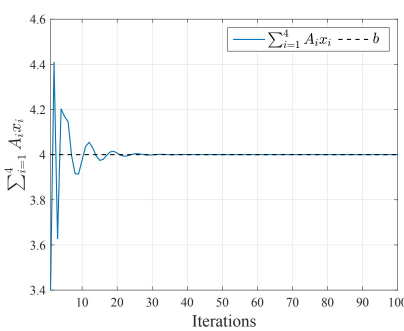

Fig. 3. The evolution of the coupling constraint $\sum_{i=1}^{4} A_i x_i = b$.

VI. CONCLUSIONS

In this paper, a new kind of surplus-based ADMM algorithm has been proposed to solve the distributed optimization problems with both local constraints and coupling constraints. By integrating a surplus-based averaging consensus algorithm into the newly designed ADMM-based algorithm, the effect of the asymmetric information communication on the convergence of the optimization algorithm has been effectively eliminated. Moreover, the update of primal variables with constraints can be implemented in a distributed manner by only utilizing the local information. Furthermore, the convergence to the optimal solution under the mild convex assumption for the objective functions has been proved. Future work will focus on designing a parallel ADMM algorithm suitable for time-varying graphs.

REFERENCES

[1] G. Wen, X. Yu, Z. Liu, and W. Yu, "Adaptive consensus-based robust strategy for economic dispatch of smart grids subject to communication uncertainties," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 6, pp. 2484-2496, 2018.

[2] Y. Zhang, S. Xie, and S. Shu, "Decentralized optimization of multiarea interconnected traffic-power systems with wind power uncertainty," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 133-143, 2023.

[3] S. Meng, W. Huang, X. Yin, M. R. Khosravi, Q. Li, S. Wan, and L. Qi, "Security-aware dynamic scheduling for real-time optimization in cloud-based industrial applications," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4219-4228, 2021.

[4] G. Wen, W. Zheng, and Y. Wan, "Distributed robust optimization for networked agent systems with unknown nonlinearities," *IEEE Transactions on Automatic Control*, vol. 68, no. 9, pp. 5230-5244, 2023.

[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Now Foundations and Trends*, 2011.

[6] J. Eckstein and W. Yao, "Understanding the convergence of the alternating direction method of multipliers: Theoretical and computational perspectives," *Pacific Journal of Optimization*, vol. 11, no. 4, pp. 619-644, 2015.

[7] T.-H. Chang, "A proximal dual consensus ADMM method for multi-agent constrained optimization," *IEEE Transactions on Signal Processing*, vol. 64, no. 14, pp. 3719-3734, 2016.

[8] M. Doostmohammadian, W. Jiang, and T. Charalambous, "DTAC-ADMM: Delay-tolerant augmented consensus ADMM-based algorithm for distributed resource allocation," in *2022 61st IEEE Conference on Decision and Control (CDC)*, pp. 308-315, 2022.

[9] A. Falsone, I. Notarnicola, G. Notarstefano, and M. Prandini, "Tracking-ADMM for distributed constraint-coupled optimization," *Automatica*, vol. 117, pp. 108962, 2020.

[10] A. Falsone and M. Prandini, "Augmented lagrangian tracking for distributed optimization with equality and inequality coupling constraints," *Automatica*, vol. 157, pp. 111269, 2023.

[11] Z. Zhao, W. Gao, Y. Wang, and M. Wei, "Differentially private distributed online optimization via push-sum one-point bandit dual averaging," *Neurocomputing*, vol. 572, pp. 127184, 2024.

[12] M. Wei, W. Yu, H. Liu, and Q. Xu, "Distributed weakly convex optimization under random time-delay interference," *IEEE Transactions on Network Science and Engineering*, vol. 11, no. 1, pp. 212-224, 2024.

[13] K. Rokade and R. K. Kalaimani, "Distributed ADMM over directed networks," *arXiv e-prints, arXiv:2010.10421*, 2020.

[14] D. Yi and N. M. Freris, Distributed optimization on directed graphs based on inexact ADMM with partial participation," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, pp. 1138-1143, 2023.

[15] V. Khatana and M. V. Salapaka, "D-DistADMM: A O(1/k) distributed ADMM for distributed optimization in directed graph topologies," in *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 2992-2997, 2020.

[16] V. Khatana and M. V. Salapaka, "DC-DistADMM: ADMM algorithm for constrained optimization over directed graphs," *IEEE Transactions on Automatic Control*, vol. 68, no. 9, pp. 5365-5380, 2023.

[17] Y. Xu, T. Han, K. Cai, Z. Lin, G. Yan, and M. Fu, "A distributed algorithm for resource allocation over dynamic digraphs," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2600-2612, 2017.

[18] D. Bertsekas and J. Tsitsiklis, "Parallel and Distributed Computation: Numerical Methods," *Athena Scientific*, 1989.

[19] K. Cai and H. Ishii, "Average consensus on arbitrary strongly connected digraphs with time-varying topologies," *IEEE Transactions on Automatic Control*, vol. 59, no. 4, pp. 1066-1071, 2014.