Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Distributed optimization based on improved push-sum framework for optimization problem with multiple local constraints and its application in smart grid*

Qian XU[1], Chutian YU[1], Xiang YUAN[1], Mengli WEI[2], Hongzhe LIU[‡2]

[1]*State Grid Zhejiang Economic Research Institute, Hangzhou 310008, China*

[2]*Beijing Yinshan Technology Co., Ltd., Beijing 100871, China*

E-mail: xuqianhangzhou@163.com; yuchutianhangzhou@163.com; yuanxianghangzhou@163.com; wengli33wei@gmail.com; 15150650205@163.com

Received Nov. 25, 2022; Revision accepted Feb. 2, 2023; Crosschecked Mar. 1, 2023; Published online July 28, 2023

Abstract: In this paper, the optimization problem subject to N nonidentical closed convex set constraints is studied. The aim is to design a corresponding distributed optimization algorithm over the fixed unbalanced graph to solve the considered problem. To this end, with the push-sum framework improved, the distributed optimization algorithm is newly designed, and its strict convergence analysis is given under the assumption that the involved graph is strongly connected. Finally, simulation results support the good performance of the proposed algorithm.

Key words: Distributed optimization; Nonidentical constraints; Improved push-sum framework
https://doi.org/10.1631/FITEE.2200596 **CLC number:** O224

1 Introduction

Because the distributed optimization algorithm performs well in solving large-scale optimization problems involved in smart grid systems, smart communication systems, multiple unmanned systems, and intelligent transportation systems, it has attracted increasing attention recently. Many excellent algorithms have been obtained, including the continuous-time algorithms (Wang and Elia, 2011; Gharesifard and Cortés, 2014; Kia et al., 2015; Liu QS and Wang, 2015; Yang et al., 2017; Zhu YN et al., 2019a, 2019b) and the discrete-time algorithms (Nedić and Ozdaglar, 2009; Nedić et al., 2010; Yuan et al., 2011; Zhu MH and Martínez, 2012; Nedić and Olshevsky, 2015; Pu et al., 2018, 2021; Qu and Li, 2018; Xi et al., 2018; Mai and Abed, 2019; Zimmermann et al., 2020; Liu HZ et al., 2021; Yu et al., 2021). In this paper, we focus on the results involving the design of discrete-time algorithms. The distributed gradient (subgradient) method was first designed in Nedić and Ozdaglar (2009), and the unconstrained problem was solved. Then, based on Nedić and Ozdaglar (2009), the distributed projected gradient (subgradient) method was designed in Nedić et al. (2010), and the optimization problem subject to the identical closed convex set constraint was solved. Then, to extend the results with inequality and equality constraints, the primal dual method was introduced in Zhu MH and Martínez (2012).

However, in the above-mentioned results, only balanced graphs were considered. To design distributed optimization algorithms over unbalanced graphs, the method of estimating the left eigenvector with respect to eigenvalue 1 of the row-stochastic adjacent matrix, push-sum method, and push-pull

‡ Corresponding author
* Project supported by the Science and Technology Project from State Grid Zhejiang Electric Power Co., Ltd., China (No. 5211JY20001Q)
ORCID: Hongzhe LIU, https://orcid.org/0000-0001-9021-4755

method were introduced. As for using the method of estimating the left eigenvector with respect to eigenvalue 1 of the row-stochastic adjacent matrix, distributed algorithms were designed to solve the problem with N nonidentical closed convex set constraints in Mai and Abed (2019) and the problem with N nonidentical inequality constraints, N nonidentical equality constraints, and N nonidentical closed convex set constraints in Liu HZ et al. (2021). Concerning the use of the push-sum method, a distributed algorithm was designed in Nedić and Olshevsky (2015), which can solve only the unconstrained optimization problem. A distributed algorithm was developed in Yu et al. (2021) to solve the optimization problem with N nonidentical set constraints and in Zimmermann et al. (2020) to solve the problem with identical global constraints and N nonidentical inequality constraints. However, the structures of these algorithms are complex. As for using the push-pull method, related distributed algorithms were designed in Pu et al. (2018, 2021), and they can solve only unconstrained optimization problems.

In the above-mentioned results, some of the designed distributed algorithms involved the sub-linear convergence rate and others (push-pull based algorithms) involved the linear convergence rate. Most recently, many excellent studies focused on designing distributed algorithms with a linear convergence rate to solve various optimization problems; details can be found in Qu and Li (2018) and Xi et al. (2018).

In this paper, the aim is to use the push-sum method to design a distributed algorithm over the fixed unbalanced graph with a sub-linear convergence rate and to solve the optimization problem with N nonidentical set constraints. Therefore, the main contribution of this paper is the successful development of a distributed optimization algorithm with strict convergence analysis. In detail, the improved push-sum framework was introduced in Zimmermann et al. (2020) to integrate the constraint handling method in the push-sum framework, while the new constraint handling method (gradient decent like method) was employed in Yu et al. (2021) to achieve the same aim. These two methods are combined in this paper to handle the constrained distributed optimization problems for unbalanced graphs. Moreover, compared to Zimmermann et al. (2020), the optimization problem studied in this paper is different and involves N nonidentical set constraints, which is more challenging than the problem with identical set constraints.

2 Preliminaries

The notations and graph theories used in this study are the same as those in Liu HZ et al. (2021). Therefore, we omit them here.

Lemma 1 (Nedić and Olshevsky, 2015) (1) For a scalar sequence $\{\theta(k)\}$, let $\lim_{k\to\infty}\theta(k)=\theta_0$ and $0<\gamma_0<1$ hold. Then, $\lim_{k\to\infty}\sum_{l=0}^{k}\gamma_0^{k-l}\theta(l)=\frac{\theta_0}{1-\gamma_0}$.

(2) For a strictly positive scalar sequence $\{\theta(k)\}$, let $\sum_{k=0}^{\infty}\theta(k)<\infty$ and $0<\gamma_0<1$ hold. Then, $\sum_{k=0}^{\infty}\sum_{l=0}^{k}\gamma_0^{k-l}\theta(l)<\infty$.

Lemma 2 (Nedić and Olshevsky, 2015) For the nonnegative scalar sequences $\{a(k)\}$, $\{b(k)\}$, $\{c(k)\}$, and $\{d(k)\}$ with $\sum_{k=0}^{\infty}b(k)<\infty$ and $\sum_{k=0}^{\infty}c(k)<\infty$, if the following condition holds:

$$a(k+1)\le(1+b(k))a(k)-d(k)+c(k),\ \forall k\ge 0,$$

then the sequence $\{a(k)\}$ converges to a for some $a\ge 0$ and $\sum_{k=0}^{\infty}d(k)<\infty$.

Lemma 3 (Mai and Abed, 2019) The following inequality holds for all $\boldsymbol{x}_1\in\mathbb{R}^n$ and $\boldsymbol{x}_2\in Y$:

$$\|P_Y(\boldsymbol{x}_1)-\boldsymbol{x}_2\|^2\le\|\boldsymbol{x}_1-\boldsymbol{x}_2\|^2-\|P_Y(\boldsymbol{x}_1)-\boldsymbol{x}_1\|^2,$$

where Y is an arbitrarily given closed convex set in $\mathbb{R}^n$ and P_Y denotes the projection operator on Y.

Note that Lemma 1 is the foundation of the directed graph, whereas Lemma 2 is for convergence analysis with diminishing step sizes. Lemma 3 is the projection lemma for the nonidentical set constraints.

3 Main results

3.1 Problem formulation

In this study, we consider the following optimization problem:

$$\min_{\boldsymbol{y}}\ f(\boldsymbol{y})=\sum_{i=1}^{N}f_i(\boldsymbol{y})\ \text{ s.t. }\ \boldsymbol{y}\in Y_i, \tag{1}$$

where $\boldsymbol{y}\in\mathbb{R}^n$, $f_i:\mathbb{R}^n\to\mathbb{R}$ $(i=1,2\cdots,N)$ are convex functions that can be seen as the local objective functions associated with agent i, and

$Y_i \subseteq \mathbb{R}^n$ $(i = 1, 2 \cdots, N)$ are compact sets. Let $Y = \bigcap_{i=1}^{N} Y_i \neq \varnothing$ and $Y_0 = \prod_{i=1}^{N} Y_i$. In this study, we design the distributed optimization algorithm on a static weight-unbalanced digraph $\mathcal{G}$. Moreover, let $\boldsymbol{B} \subseteq \mathbb{R}^{N\times N}$ be the weight matrix associated with $\mathcal{G}$. Assume that $\boldsymbol{B}$ is column-stochastic.

In the following, the assumptions for the optimization problem and digraph $\mathcal{G}$ are introduced for subsequent analysis:

Assumption 1 For $i = 1, 2, \cdots, N$, $f_i(x)$ has a continuous gradient.

Assumption 2 Digraph $\mathcal{G}$ is strongly connected.

Remark 1 Because sets Y_i are compact, $\bigcup_{i=1}^{N} Y_i$ is compact. Thus, there exists a positive constant R such that

$$\operatorname{dist}(\boldsymbol{y}, Y) \leq R \max_{1\leq i\leq N} \operatorname{dist}(\boldsymbol{y}, Y_i),\ \forall \boldsymbol{y} \in \bigcup_{i=1}^{N} Y_i. \quad (2)$$

The measurement $\operatorname{dist}(\boldsymbol{y}, Y)$ and parameter R denote the distance of $\boldsymbol{y}$ to Y and a positive constant related to the diameter of the sets, respectively. Inequality (2) is known as the constraint regularity condition, which is the key to solving optimization problems with multiple nonidentical constraints.

3.2 Distributed algorithms

In this subsection, the distributed algorithm on digraph $\mathcal{G}$ is designed for problem (1), which is given as

$$x_i(t+1) = \sum_{j=1}^{N} B_{ij} x_j(t), \quad (3a)$$

$$\boldsymbol{v}_i(t+1) = \frac{1}{x_i(t+1)} \sum_{j=1}^{N} B_{ij} x_j(t) \boldsymbol{y}_j(t), \quad (3b)$$

$$\boldsymbol{y}_i(t+1) = P_{Y_i}\Big(\boldsymbol{v}_i(t+1) - \frac{\alpha(t)}{x_i(t+1)} \nabla f_i(\boldsymbol{y}_i(t))\Big), \quad (3c)$$

where $x_i(t) \in \mathbb{R}$ is the introduced auxiliary variable, $\boldsymbol{y}_i(t) \in \mathbb{R}^n$ is the decision variable, $\boldsymbol{v}_i(t)$ is the introduced intermediate variable, and $\{\alpha(t)\}$ is the decaying positive step-size sequence satisfying $\sum_{t=0}^{\infty} \alpha(t) = \infty$ and $\sum_{t=0}^{\infty} \alpha^2(t) < \infty$. Moreover, the initial values $\boldsymbol{y}_i(0)$ can be arbitrarily selected and $x_i(0) = 1$, $i = 1, 2, \cdots, N$.

Remark 2 Because Y_i $(i = 1, 2, \cdots, N)$ are compact sets, $\boldsymbol{y}_i(t)$ under iteration (3c) are uniformly bounded with respect to t. Moreover, from Nedić and Olshevsky (2015), we know that $x_i(t)$ under iteration (3a) with the given initial values are strictly positive; in other words, there exists a constant a_0 $(a_0 > 0)$ such that $x_i(t) \geq a_0$, $i = 1, 2, \cdots, N$ and $t \geq 0$. Thus, $\frac{1}{x_i(t)}$ $(i = 1, 2, \cdots, N)$ are uniformly bounded with respect to t. Therefore, it can be assumed that there exists a positive constant M such that

$$\max_{1\leq i\leq N}\left\{x_i(t), \frac{1}{x_i(t)}, \boldsymbol{y}_i(t), \nabla f_i(\boldsymbol{y}_i(t)), \frac{\nabla f_i(\boldsymbol{y}_i(t))}{x_i(t)}\right\} \leq M.$$

Then, without loss of generality, we can assume that $n = 1$ and algorithm (3) can be rewritten in a compact form

$$\boldsymbol{x}(t+1) = \boldsymbol{B}\boldsymbol{x}(t), \quad (4a)$$

$$\boldsymbol{v}(t+1) = \boldsymbol{Q}(t+1)\boldsymbol{y}(t), \quad (4b)$$

$$\boldsymbol{y}(t+1) = P_{Y_0}\Big(\boldsymbol{v}(t+1) - \alpha(t)\boldsymbol{D}(t+1)\nabla F(\boldsymbol{y}(t))\Big), \quad (4c)$$

where

$$\begin{cases} \boldsymbol{x}(t) = (x_1(t), x_2(t), \cdots, x_N(t))^{\mathrm{T}}, \\ \boldsymbol{v}(t) = (\boldsymbol{v}_1(t), \boldsymbol{v}_2(t), \cdots, \boldsymbol{v}_N(t))^{\mathrm{T}}, \\ \boldsymbol{y}(t) = (\boldsymbol{y}_1(t), \boldsymbol{y}_2(t), \cdots, \boldsymbol{y}_N(t))^{\mathrm{T}}, \\ \boldsymbol{Q}(t) = \boldsymbol{D}^{-1}(t)\boldsymbol{B}\boldsymbol{D}(t-1),\ \forall t \geq 1, \\ \boldsymbol{D}(t) = \operatorname{diag}(x_1(t), x_2(t), \cdots, x_N(t)). \end{cases}$$

3.3 Convergence analysis

Clearly, the properties of matrix $\boldsymbol{Q}(t)$ are the key points for the convergence analysis of the proposed algorithm. The roadmap for the convergence analysis is briefly introduced as follows: first, we will discuss the properties of matrix $\boldsymbol{Q}(t)$ in Lemma 4. Then, we will analyze the consensus property of $\boldsymbol{y}_i(t)$ in Lemma 5 and the convergence property of the optimal solution to $\boldsymbol{y}_i(t)$ under algorithm (3) in Lemmas 6 and 7. Finally, we give the convergence properties of $\boldsymbol{y}_i(t)$ in Theorem 1.

Lemma 4 Under Assumption 2, matrix $\boldsymbol{Q}(t)$, induced from algorithm (3), has the following properties:

(1) For all $t \geq 1$, $\boldsymbol{Q}(t)$ are stochastic;

(2) For all $t \geq 0$, $\boldsymbol{x}^{\mathrm{T}}(t+1)\boldsymbol{Q}(t+1) = \boldsymbol{x}^{\mathrm{T}}(t)$.

The proof of Lemma 4 can be found in Zimmermann et al. (2020), and thus it is omitted here.

Lemma 5 Under Assumptions 1 and 2, $\boldsymbol{y}_i(t)$ ($i=1,2,\cdots,N$) generated by algorithm (3) satisfy

$$\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{y}}(t)\| \le C_1\eta^t + C_2\sum_{s=0}^{t-1}\eta^{t-1-s}(\alpha(s)+\gamma(s)), \tag{5}$$

where $\bar{\boldsymbol{y}}(t)=\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{y}_i(t)$, C_1 and C_2 are two positive constants, η is a constant satisfying $0<\eta<1$, and $\gamma(t)=\sum_{i=1}^{N}\|\boldsymbol{z}_i(t)\|$ with

$$\begin{cases} \boldsymbol{z}_i(t)=\boldsymbol{y}_i(t+1)-\boldsymbol{u}_i(t), \\ \boldsymbol{u}_i(t)=\boldsymbol{v}_i(t+1)-\frac{\alpha(t)}{x_i(t+1)}\nabla f_i(\boldsymbol{y}_i(t)). \end{cases}$$

The proof of Lemma 5 can be completed based on the results in Mai and Abed (2019), and thus it is omitted here.

Lemma 6 Let $\boldsymbol{y}^*$ be the optimal solution to problem (1). Under Assumptions 1 and 2, $\boldsymbol{y}_i(t)$ ($i=1,2,\cdots,N$) under algorithm (3) satisfy

$$\begin{aligned} &\sum_{i=1}^{N}x_i(t+1)\|\boldsymbol{y}_i(t+1)-\boldsymbol{y}^*\|^2 \\ &\le \sum_{i=1}^{N}x_i(t)\|\boldsymbol{y}_i(t)-\boldsymbol{y}^*\|^2 - 2\alpha(t)(f(\bar{\boldsymbol{s}}(t))-f^*) \\ &\quad + D_1\alpha(t)\sum_{i=1}^{N}\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{y}}(t)\| - a_0\sum_{i=1}^{N}\|\boldsymbol{z}_i(t)\|^2 \\ &\quad + M^2\alpha^2(t), \end{aligned} \tag{6}$$

where $\bar{\boldsymbol{s}}(t)=P_Y(\bar{\boldsymbol{y}}(t))$ and $D_1=(R+3)M$.

Proof From the definition of $\boldsymbol{y}_i(t+1)$ and Lemma 3, we obtain

$$\begin{aligned} &\sum_{i=1}^{N}x_i(t+1)\|\boldsymbol{y}_i(t+1)-\boldsymbol{y}^*\|^2 \\ &\le \sum_{i=1}^{N}x_i(t+1)\|\boldsymbol{u}_i(t)-\boldsymbol{y}^*\|^2 - \sum_{i=1}^{N}x_i(t+1)\|\boldsymbol{z}_i(t)\|^2. \end{aligned} \tag{7}$$

Furthermore, we have

$$\begin{aligned} &\sum_{i=1}^{N}x_i(t+1)\|\boldsymbol{u}_i(t)-\boldsymbol{y}^*\|^2 \\ &\le \sum_{i=1}^{N}x_i(t+1)\left\|\sum_{j=1}^{N}Q_{ij}(t+1)\boldsymbol{y}_j(t)-\boldsymbol{y}^*\right\|^2 \\ &\quad + \sum_{i=1}^{N}x_i(t+1)\left\|\frac{\alpha(t)}{x_i(t+1)}\nabla f_i(\boldsymbol{y}_i(t))\right\|^2 \\ &\quad -2\sum_{i=1}^{N}x_i(t+1)\left(\sum_{j=1}^{N}Q_{ij}(t+1)\boldsymbol{y}_j(t)-\boldsymbol{y}^*\right)^{\mathrm{T}} \\ &\quad \cdot\left(\frac{\alpha(t)}{x_i(t+1)}\nabla f_i(\boldsymbol{y}_i(t))\right). \end{aligned} \tag{8}$$

Then, for the first term of the right-hand side of inequality (8), based on the property 2 of Lemma 4 we have

$$\begin{aligned} &\sum_{i=1}^{N}x_i(t+1)\left\|\sum_{j=1}^{N}Q_{ij}(t+1)\boldsymbol{y}_j(t)-\boldsymbol{y}^*\right\|^2 \\ &\le \sum_{i=1}^{N}x_i(t)\|\boldsymbol{y}_i(t)-\boldsymbol{y}^*\|^2. \end{aligned} \tag{9}$$

For the second term of the right-hand side of inequality (8), we have

$$\sum_{i=1}^{N}x_i(t+1)\left\|\frac{\alpha(t)}{x_i(t+1)}\nabla f_i(\boldsymbol{y}_i(t))\right\|^2 \le M^2\alpha^2(t). \tag{10}$$

For the third term of the right-hand side of inequality (8), we have

$$\begin{aligned} &\sum_{i=1}^{N}x_i(t+1)\left(\sum_{j=1}^{N}Q_{ij}(t+1)\boldsymbol{y}_j(t)-\boldsymbol{y}^*\right)^{\mathrm{T}} \\ &\cdot\left(\frac{\alpha(t)}{x_i(t+1)}\nabla f_i(\boldsymbol{y}_i(t))\right) \\ &= \sum_{i=1}^{N}\left(\sum_{j=1}^{N}Q_{ij}(t+1)\boldsymbol{y}_j(t)-\boldsymbol{y}_i(t)+\boldsymbol{y}_i(t)-\boldsymbol{y}^*\right)^{\mathrm{T}} \\ &\cdot(\alpha(t)\nabla f_i(\boldsymbol{y}_i(t))). \end{aligned} \tag{11}$$

Furthermore, we can obtain

$$\begin{aligned} &\left(\sum_{j=1}^{N}Q_{ij}(t+1)\boldsymbol{y}_j(t)-\boldsymbol{y}_i(t)\right)^{\mathrm{T}}(\alpha(t)\nabla f_i(\boldsymbol{y}_i(t))) \\ &= \left(\sum_{j=1}^{N}Q_{ij}(t+1)\boldsymbol{y}_j(t)-\bar{\boldsymbol{y}}(t)\right)(\alpha(t)\nabla f_i(\boldsymbol{y}_i(t))) \\ &\quad + (\bar{\boldsymbol{y}}-\boldsymbol{y}_i(t))(\alpha(t)\nabla f_i(\boldsymbol{y}_i(t))) \\ &\ge -M\alpha(t)\sum_{j=1}^{N}Q_{ij}(t+1)\|\boldsymbol{y}_j(t)-\bar{\boldsymbol{y}}(t)\| \\ &\quad - M\alpha(t)\|\bar{\boldsymbol{y}}-\boldsymbol{y}_i(t)\| \\ &\ge -2M\alpha(t)\sum_{i=1}^{N}\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{y}}(t)\|. \end{aligned} \tag{12}$$

Moreover, we have

$$\begin{aligned}&(\boldsymbol{y}_i(t)-\boldsymbol{y}^*)^{\mathrm{T}}(\alpha(t)\nabla f_i(\boldsymbol{y}_i(t)))\\ \geq\ &\alpha(t)(f_i(\boldsymbol{y}_i(t))-f^*)\\ \geq\ &-\alpha(t)|f_i(\boldsymbol{y}_i(t))-f_i(\bar{\boldsymbol{s}}(t))|\\ &+\alpha(t)(f_i(\bar{\boldsymbol{s}}(t))-f_i(\boldsymbol{y}^*))\\ \geq\ &-\alpha(t)M\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{s}}(t)\|+\alpha(t)(f_i(\bar{\boldsymbol{s}}(t))-f_i(\boldsymbol{y}^*)).\end{aligned}\tag{13}$$

Noting the results discussed in Remark 1, we obtain

$$\begin{aligned}&\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{s}}(t)\|\\ \geq&-\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{y}}(t)\|-\|\bar{\boldsymbol{y}}(t)-\bar{\boldsymbol{s}}(t)\|\\ \geq&-\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{y}}(t)\|-R\sum_{i=1}^{N}\|\bar{\boldsymbol{y}}(t)-\boldsymbol{s}_i(t)\|\\ \geq&-\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{y}}(t)\|-R\sum_{i=1}^{N}\|\bar{\boldsymbol{y}}(t)-\boldsymbol{y}_i(t)\|.\end{aligned}\tag{14}$$

Then, it follows from inequalities (13) and (14) that

$$\begin{aligned}&(\boldsymbol{y}_i(t)-\boldsymbol{y}^*)^{\mathrm{T}}(\alpha(t)\nabla f_i(y_i(t)))\\ \geq&-(R+1)M\alpha(t)\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{y}}(t)\|\\ &+\alpha(t)(f_i(\bar{\boldsymbol{s}}(t))-f^*).\end{aligned}\tag{15}$$

Thus, substituting inequalities (12) and (15) into Eq. (11) yields

$$\begin{aligned}&\sum_{i=1}^{N}x_i(t+1)\left(\sum_{j=1}^{N}Q_{ij}(t+1)\boldsymbol{y}_j(t)-\boldsymbol{y}^*\right)^{\mathrm{T}}\\ &\cdot\left(\frac{\alpha(t)}{x_i(t+1)}\nabla f_i(\boldsymbol{y}_i(t))\right)\\ \geq&-D_1\alpha(t)\sum_{i=1}^{N}\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{y}}(t)\|+\alpha(t)(f(\bar{\boldsymbol{s}}(t))-f^*).\end{aligned}\tag{16}$$

Therefore, it follows from inequalities (9), (10), and (16) that

$$\begin{aligned}&\sum_{i=1}^{N}x_i(t+1)\|\boldsymbol{u}_i(t)-\boldsymbol{y}^*\|^2\\ \leq&\sum_{i=1}^{N}x_i(t)\|\boldsymbol{y}_i(t)-\boldsymbol{y}^*\|^2+M^2\alpha^2(t)\\ &+2D_1\alpha(t)\sum_{i=1}^{N}\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{y}}(t)\|-2\alpha(t)(f_i(\bar{\boldsymbol{s}}(t))-f^*),\end{aligned}\tag{17}$$

$$\begin{aligned}&\sum_{i=1}^{N}x_i(t+1)\|\boldsymbol{y}_i(t+1)-\boldsymbol{y}^*\|^2\\ \leq&\sum_{i=1}^{N}x_i(t)\|\boldsymbol{y}_i(t)-\boldsymbol{y}^*\|^2-2\alpha(t)(f_i(\bar{\boldsymbol{s}}(t))-f^*)\\ &+M^2\alpha^2(t)+2D_1\alpha(t)\sum_{i=1}^{N}\|\boldsymbol{y}_i(t)-\bar{\boldsymbol{y}}(t)\|\\ &-\sum_{i=1}^{N}x_i(t+1)\|\boldsymbol{z}_i(t)\|^2.\end{aligned}\tag{18}$$

The proof is completed.

Let $\beta(t)=\alpha(t)\sum\limits_{s=0}^{t-1}\eta^{t-1-s}\gamma(s)$ with $\beta(0)=0$. Furthermore, we have

$$\beta(t+1)\leq\eta\beta(t)+\alpha(t)\gamma(t).\tag{19}$$

Lemma 7 Under Assumptions 1 and 2, $\boldsymbol{y}_i(t)$ $(i=1,2,\cdots,N)$ under algorithm (3) further satisfy

$$\begin{aligned}&\sum_{i=1}^{N}x_i(t+1)\|\boldsymbol{y}_i(t+1)-\boldsymbol{y}^*\|^2+ab\beta(t+1)\\ \leq&\sum_{i=1}^{N}x_i(t)\|\boldsymbol{y}_i(t)-\boldsymbol{y}^*\|^2+ab\beta(t)\\ &-2\alpha(t)(f(\bar{\boldsymbol{s}}(t))-f^*)+\xi(t)+\frac{1}{2}a^2\alpha^2(t),\end{aligned}\tag{20}$$

where $a=\frac{NC_2}{(1-\eta)b}$, $b=\sqrt{\frac{2a_0}{N}}$, and

$$\xi(t)=M^2\alpha^2(t)+D_1C_1\eta^t\alpha(0)+D_1C_2\sum_{s=0}^{t-1}\eta^{t-1-s}\alpha^2(s).$$

Proof It can be obtained from inequalities (5) and (6) that

$$\begin{aligned}&\sum_{i=1}^{N}x_i(t+1)\|\boldsymbol{y}_i(t+1)-\boldsymbol{y}^*\|^2\\ \leq&\sum_{i=1}^{N}x_i(t)\|\boldsymbol{y}_i(t)-\boldsymbol{y}^*\|^2-2\alpha(t)(f(\bar{\boldsymbol{s}}(t))-f^*)\\ &+\xi(t)+NC_2\beta(t)-a_0\sum_{i=1}^{N}\|\boldsymbol{z}_i(t)\|^2.\end{aligned}\tag{21}$$

Then, it follows from inequalities (19) and (21) that

$$\begin{aligned}
&\sum_{i=1}^{N} x_i(t+1)\|\boldsymbol{y}_i(t+1)-\boldsymbol{y}^*\|^2+ab\beta(t+1)\\
\leq&\sum_{i=1}^{N} x_i(t)\|\boldsymbol{y}_i(t)-\boldsymbol{y}^*\|^2+ab\eta\beta(t)+ab\alpha(t)\gamma(t)\\
&-2\alpha(t)(f(\bar{\boldsymbol{s}}(t))-f^*)+\xi(t)\\
&+NC_2\beta(t)-a_0\sum_{i=1}^{N}\|\boldsymbol{z}_i(t)\|^2\\
\leq&\sum_{i=1}^{N} x_i(t)\|\boldsymbol{y}_i(t)-\boldsymbol{y}^*\|^2+ab\beta(t)+ab(\eta-1)\beta(t)\\
&+ab\alpha(t)\gamma(t)-2\alpha(t)(f(\bar{\boldsymbol{s}}(t))-f^*)+\xi(t)\\
&+NC_2\beta(t)-a_0\sum_{i=1}^{N}\|\boldsymbol{z}_i(t)\|^2.
\end{aligned} \tag{22}$$

Thus, with the selected a, we have

$$\begin{aligned}
&\sum_{i=1}^{N} x_i(t+1)\|\boldsymbol{y}_i(t+1)-\boldsymbol{y}^*\|^2+ab\beta(t+1)\\
\leq&\sum_{i=1}^{N} x_i(t)\|\boldsymbol{y}_i(t)-\boldsymbol{y}^*\|^2+ab\beta(t)+ab\alpha(t)\gamma(t)\\
&-2\alpha(t)(f(\bar{\boldsymbol{s}}(t))-f^*)+\xi(t)-a_0\sum_{i=1}^{N}\|\boldsymbol{z}_i(t)\|^2.
\end{aligned} \tag{23}$$

Therefore, noting that

$$\begin{aligned}
ab\alpha(t)\gamma(t)&\leq\frac{1}{2}a^2\alpha^2(t)+\frac{1}{2}b^2\gamma^2(t)\\
&\leq\frac{1}{2}a^2\alpha^2(t)+\frac{1}{2}b^2N\sum_{i=1}^{N}\|\boldsymbol{z}_i(t)\|^2
\end{aligned} \tag{24}$$

and $b=\sqrt{\frac{2a_0}{N}}$, we can obtain inequality (20).

The proof is completed.

Now, we present Theorem 1, which describes the convergence property of the proposed algorithm (3):

Theorem 1 Under Assumptions 1 and 2, all $\boldsymbol{y}_i(t)$ under algorithm (3) converge to a common optimal solution to problem (1).

Proof It can be obtained from Lemma 1 and the properties of $\alpha(t)$ that $\xi(t)$ is summable and thus $\xi(t)+\frac{1}{2}a^2\alpha^2(t)$ is summable. Then, based on the results in Mai and Abed (2019) and Lemma 2, we can easily complete the proof.

4 Simulations

In this section, we consider the economic dispatch problem (EDP) in an IEEE 118-bus system with nine microgrids (nodes), where the topology is given in Fig. 1.

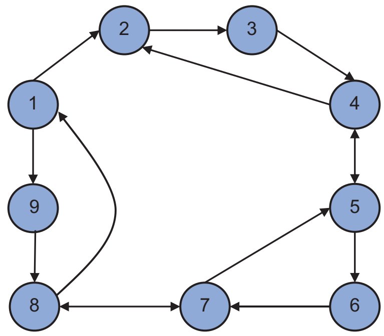

Fig. 1 Communication graph of an IEEE 118-bus example

In general, the distributed EDP can be transformed into problem (1) using the Fenchel dual method. That is, EDP obtains the optimal solution for the output power of all units as long as the optimal solution to problem (1) is achieved. Furthermore, we focus only on the dual problem modeled as

$$\min_{y\in Y} f(y)=\sum_{i=1}^{N} f_i(y),\quad y\in Y_i, \tag{25}$$

where $f_i(y)=a_iy^2+b_iy+c_i$ is a quadratic function with coefficients $a_i\in[100,500]$, $b_i\in[7.5,8]$, $c_i\in[0.001,0.004]$, and $N=9$. Then, we give the column-stochastic matrix $\boldsymbol{B}_{ij}^{\mathrm{c}}$ based on the topology of the fixed digraph as

$$\boldsymbol{B}_{ij}^{\mathrm{c}}=\begin{bmatrix}
\frac{1}{2}&\frac{1}{3}&0&0&0&0&0&0&\frac{1}{2}\\
0&\frac{1}{3}&\frac{1}{2}&0&0&0&0&0&0\\
0&0&\frac{1}{2}&\frac{1}{3}&0&0&0&0&0\\
0&\frac{1}{3}&0&\frac{1}{3}&\frac{1}{3}&0&0&0&0\\
0&0&0&\frac{1}{3}&\frac{1}{3}&\frac{1}{2}&0&0&0\\
0&0&0&0&0&\frac{1}{2}&\frac{1}{3}&0&0\\
0&0&0&0&\frac{1}{3}&0&\frac{1}{3}&\frac{1}{3}&0\\
\frac{1}{2}&0&0&0&0&0&\frac{1}{3}&\frac{1}{3}&0\\
0&0&0&0&0&0&0&\frac{1}{3}&\frac{1}{2}
\end{bmatrix}.$$

Specifically, the behaviors of $y_i(t)$ $(i=1,2,\cdots,9)$ are shown in Fig. 2, and it can be seen that $y_i(t)$ converge to a common optimal solution $y^*=9.684$.

Furthermore, we use the classical distributed projected gradient descent algorithm with row-stochastic adjacent matrix (DPGD-RS) in Mai and

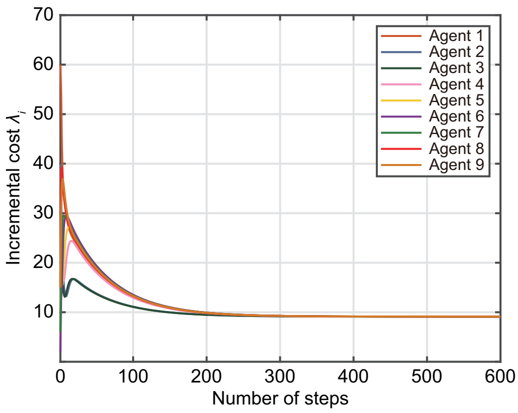

Fig. 2 States for the incremental cost of all generation units with capacity limitations

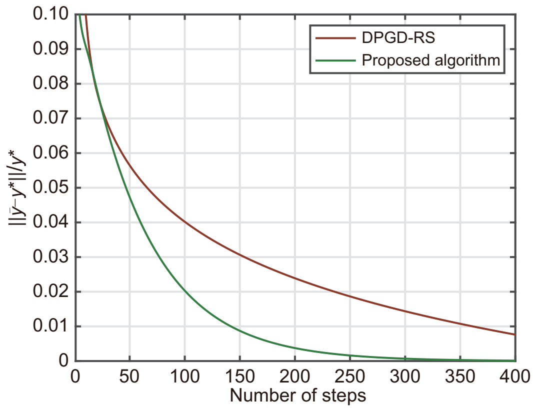

Fig. 3 Accuracy of the proposed algorithm with row-stochastic distributed projected gradient descent

Abed (2019) as a benchmark to compare with the convergence performance of the proposed algorithm in this study. The settings are the same, except the selection of the row-stochastic matrix $\boldsymbol{B}_{ij}^{\mathrm{r}}$ for DPGD-RS as

$$\boldsymbol{B}_{ij}^{\mathrm{r}}=\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

All behaviors of $y_i(t)$ under the DPGD-RS as proposed in Mai and Abed (2019) are shown in Fig. 3 for comparison. Comparing the results from the convergence analysis, the algorithm in this study has some advantages in terms of the convergence rate. Clearly, it can be shown that the algorithm presented in this study has a better convergence performance.

5 Conclusions

In this paper, we investigated the optimization problem with N nonidentical closed convex set constraints. Furthermore, we introduced the improved push-sum framework, based on which we designed the distributed algorithm over the fixed unbalanced graph. We presented strict convergence analysis of the proposed algorithm. Additionally, we verified the good performance of the proposed algorithm using simulations. Future works will focus on optimization problems with more complicated constraints and on distributed optimization over time-varying unbalanced graphs.

Contributors

Qian XU and Chutian YU designed the research. Qian XU and Xiang YUAN processed the data. Qian XU and Mengli WEI drafted the paper. Hongzhe LIU helped organize the paper. Hongzhe LIU revised and finalized the paper.

Compliance with ethics guidelines

Qian XU, Chutian YU, Xiang YUAN, Mengli WEI, and Hongzhe LIU declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Gharesifard B, Cortés J, 2014. Distributed continuous-time convex optimization on weight-balanced digraphs. *IEEE Trans Autom Contr*, 59(3):781-786. https://doi.org/10.1109/TAC.2013.2278132

Kia SS, Cortés J, Martínez S, 2015. Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication. *Automatica*, 55:254-264. https://doi.org/10.1016/j.automatica.2015.03.001

Liu HZ, Zheng WX, Yu W, 2021. Distributed discrete-time algorithms for convex optimization with general local constraints on weight-unbalanced digraph. *IEEE Trans Contr Netw Syst*, 8(1):51-64. https://doi.org/10.1109/TCNS.2020.3029996

Liu QS, Wang J, 2015. A second-order multi-agent network for bound-constrained distributed optimization. *IEEE Trans Autom Contr*, 60(12):3310-3315. https://doi.org/10.1109/TAC.2015.2416927

Mai VS, Abed EH, 2019. Distributed optimization over directed graphs with row stochasticity and constraint regularity. *Automatica*, 102:94-104. https://doi.org/10.1016/j.automatica.2018.07.020

Nedić A, Ozdaglar A, 2009. Distributed subgradient methods for multi-agent optimization. *IEEE Trans Autom Contr*, 54(1):48-61.
https://doi.org/10.1109/TAC.2008.2009515

Nedić A, Olshevsky A, 2015. Distributed optimization over time-varying directed graphs. *IEEE Trans Autom Contr*, 60(3):601-615.
https://doi.org/10.1109/TAC.2014.2364096

Nedić A, Ozdaglar A, Parrilo PA, 2010. Constrained consensus and optimization in multi-agent networks. *IEEE Trans Autom Contr*, 55(4):922-938.
https://doi.org/10.1109/TAC.2010.2041686

Pu S, Shi W, Xu JM, et al., 2018. A push-pull gradient method for distributed optimization in networks. IEEE Conf on Decision and Control, p.3385-3390.
https://doi.org/10.1109/CDC.2018.8619047

Pu S, Shi W, Xu J, et al., 2021. Push pull gradient methods for distributed optimization in networks. *IEEE Trans Autom Contr*, 66(1):1-16.
https://doi.org/10.1109/TAC.2020.2972824

Qu GN, Li N, 2018. Harnessing smoothness to accelerate distributed optimization. *IEEE Trans Contr Netw Syst*, 5(3):1245-1260.
https://doi.org/10.1109/TCNS.2017.2698261

Wang J, Elia N, 2011. A control perspective for centralized and distributed convex optimization. Proc 50th IEEE Conf on Decision and Control and European Control Conf, p.3800-3805.
https://doi.org/10.1109/CDC.2011.6161503

Xi CG, Xin R, Khan UA, 2018. ADD-OPT: accelerated distributed directed optimization. *IEEE Trans Autom Contr*, 63(5):1329-1339.
https://doi.org/10.1109/TAC.2017.2737582

Yang SF, Liu QS, Wang J, 2017. A multi-agent system with a proportional-integral protocol for distributed constrained optimization. *IEEE Trans Autom Contr*, 62(7):3461-3467.
https://doi.org/10.1109/TAC.2016.2610945

Yu WW, Liu HZ, Zheng WX, et al., 2021. Distributed discrete-time convex optimization with nonidentical local constraints over time-varying unbalanced directed graphs. *Automatica*, 134:109899.
https://doi.org/10.1016/j.automatica.2021.109899

Yuan DM, Xu SY, Zhao HY, 2011. Distributed primal-dual subgradient method for multiagent optimization via consensus algorithms. *IEEE Trans Syst Man Cybern Part B Cybern*, 41(6):1715-1724.
https://doi.org/10.1109/TSMCB.2011.2160394

Zhu MH, Martínez S, 2012. On distributed convex optimization under inequality and equality constraints. *IEEE Trans Autom Contr*, 57(1):151-164.
https://doi.org/10.1109/TAC.2011.2167817

Zhu YN, Yu WW, Wen GH, et al., 2019a. Continuous-time coordination algorithm for distributed convex optimization over weight–unbalanced directed networks. *IEEE Trans Circ Syst II Express Briefs,* 66(7):1202-1206.
https://doi.org/10.1109/TCSII.2018.2878250

Zhu YN, Yu WW, Wen GH, et al., 2019b. Continuous-time distributed subgradient algorithm for convex optimization with general constraints. *IEEE Trans Autom Contr*, 64(4):1694-1701.
https://doi.org/10.1109/TAC.2018.2852602

Zimmermann J, Tatarenko T, Willert V, et al., 2020. Projected push-sum gradient descent-ascent for convex optimization with application to economic dispatch problems. Proc 59th IEEE Conf on Decision and Control, p.542-547.
https://doi.org/10.1109/CDC42340.2020.9304360