Letter

Push-Sum Based Algorithm for Constrained Convex Optimization Problem and Its Potential Application in Smart Grid

Qian Xu, Zao Fu, Bo Zou, Hongzhe Liu, and Lei Wang

Dear Editor,

This letter aims to investigate the optimization problem where the decision variable is contained in a closed convex set. By combing the gradient decent-like method with the push-sum algorithm framework, we design the distributed iterative formulas under the condition that the considered graphs sequence is time-varying and unbalanced. Under some standard assumptions on the problem setting and the decaying step-sizes, we analyze the convergence property of the generated variables sequence under the proposed iterative formulas.

The distributed optimization has been widely and deeply researched during the recent decades due to its great potential application value in various practical fields. Accordingly, a great deal of excellent results with both the continuous-time [1]–[5] and discrete-time algorithms [6]–[16] developed have been obtained. Here, we only care about the results with discrete-time algorithms developed.

References [6] and [7] respectively developed the basic frameworks of the distributed gradient descent algorithms for unconstrained and constrained problems. Then, based on [6] and [7], researchers solved the constrained optimization problems with different constraints by various distributed algorithms designed on balanced graphs [8], [9] or unbalanced graphs [10], [11]. Specially, with the time-varying and unbalanced graphs considered, the distributed algorithm was designed in [12], which was only effective on the unconstrained optimization problem. Most recently, the distributed algorithms under the gradient tracking framework [13]–[16] became more popular since those algorithms can possess the linear convergence rates. Particularly, the unbalanced and time-varying graphs were involved in [16]. However, those algorithms can only solve the unconstrained optimization problems. In a conclusion, it is still very hard to design the distributed algorithms for the constrained optimization problems on the time-varying and unbalanced graphs and only few related works exist. Specially, [17] studied the constrained optimization problem with a closed convex set constraint $x \in \Omega$ and considered the time-varying and unbalanced graphs. However, the operation $\operatorname{argmin}_{x\in\Omega}$ was employed in the designed algorithm to deal with the involved constraint, which will make the solving process more complex and will demand more numerical task.

As for this paper, the main contributions are that a novel gradient decent-like method handling the closed convex set constraint is introduced and it can be successfully integrated into the push-sum based algorithm. Accordingly, the new distributed iterative formulas are designed without using the operation $\operatorname{argmin}_{x\in\Omega}$, which can successfully address a kind of constrained optimization problems although the time-varying and unbalanced graphs are involved. Furthermore, under the suitable assumptions on the decaying step-sizes, we provide the convergence analysis.

Corresponding author: Hongzhe Liu.

Citation: Q. Xu, Z. Fu, B. Zou, H. Z. Liu, and L. Wang, "Push-sum based algorithm for constrained convex optimization problem and its potential application in smart grid," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 10, pp. 1889–1891, Oct. 2022.

Q. Xu, B. Zou, and L. Wang are with the State Grid Zhejiang Economic Research Institute, Hangzhou 310008, China (e-mail: qianxuhangzhou@gmail.com; eebzou@zj.sgcc.com.cn; wangleihangzhou10@163.com).

Z. Fu and H. Z. Liu are with Beijing Yinshan Technology Co., LTD., Beijing 100871, China (e-mail: 278437639@qq.com; 15150650205@163.com).

Digital Object Identifier 10.1109/JAS.2022.105890

Preliminaries: In this part, a general description on graph theory and the notations employed are shown.

Graph theory: Let the index set of the nodes be denoted as $\mathcal{N} \triangleq \{1,\dots,n\}$. Specially, assume the communication network among n agents in this letter consists of the graphs sequence $\{\mathcal{G}(r)\}$ with $r \in \{0,1,2\dots\} \triangleq \mathcal{R}$, which is time-varying and unbalanced. Additionally, we assume that $\{B(r)\}$ is the adjacency matrices sequence and all $B(r)$ are nonnegative column stochastic. Furthermore, for the given $r \in \mathbb{R}$, we define $B_{ji}(r)$ as $B_{ji}(r) > 0$, if $i \in \mathcal{N}^j(r)$ or $i = j$; $B_{ji}(r) > 0$, otherwise, where $\mathcal{N}^j(r)$ denotes the neighbours set of the agent $j \in \mathcal{N}$ at step r. For the digraphs $\mathcal{G}(r), \mathcal{G}(r+1), \dots, \mathcal{G}(r+H_0)$ with H_0 being a positive integer, their joint graph is depicted by the node set $\mathcal{N}$ and the edge set $\mathcal{E}(r)\bigcup\mathcal{E}(r+1)\bigcup\cdots\bigcup\mathcal{E}(r+H_0)$. Furthermore, we say $\{\mathcal{G}(r)\}$ is uniformly jointly strongly connected if a positive integer H exists such that for all $r \in \mathcal{R}$, the joint graph of $\mathcal{G}(r), \mathcal{G}(r+1), \dots, \mathcal{G}(r+H-1)$ is strongly connected.

Assumption 1: $\{\mathcal{G}(r)\}$ considered in this letter is uniformly jointly strongly connected.

Notations: Let $\mathbb{R}^p$ stand for Euclidean vectors space and $\mathbb{R}^{q\times p}$ represent the $q \times p$-dimensional Euclidean matrices space. Furthermore, for a given $y \in \mathbb{R}^p$ and $M \in \mathbb{R}^{q\times p}$, $\|y\|$ and $\|M\|$ respectively stand for the 2-norm of y and the Frobenius norm of M. Furthermore, M_{ij} represents the ij-th entry of M. Additionally, for a convex function g defined on $\mathbb{R}^p$, $\partial g(y)$ represents its gradient at y. For a closed convex set Y, $P_Y(\cdot)$ represents the projection operator on Y.

Main results: First, we formulate the optimization problem as

$$\min_{x\in X} g(x) = \frac{1}{n}\sum_{i=1}^{n} g_i(x) \tag{1}$$

where $X \subseteq \mathbb{R}^p$ is a closed convex set; g_i are convex local objective functions defined on $\mathbb{R}^p$ for all $i \in \mathcal{N}$. Next, to solve the problem (1), the following distributed iterative (2) is designed:

$$z_i(r+1) = \sum_{j=1}^{n} B_{ij}(r) y_j(r) \tag{2a}$$

$$w_i(r+1) = \sum_{j=1}^{n} B_{ij}(r) w_j(r) \tag{2b}$$

$$x_i(r+1) = \frac{z_i(r+1)}{w_i(r+1)} \tag{2c}$$

$$\begin{aligned} y_i(r+1) = z_i(r+1) + \frac{\beta(r)}{\|d_i(r+1)\|}(s_i(r+1) \\ - x_i(r+1)) - \alpha(r)\partial g_i(s_i(r+1)) \end{aligned} \tag{2d}$$

where $s_i(r+1) = P_X(x_i(r+1))$, and

$$d_i(r+1) = \begin{cases} s_i(r+1) - x_i(r+1), & \|s_i(r+1) - x_i(r+1)\| > d \\ s, & \|s_i(r+1) - x_i(r+1)\| \le d \end{cases}$$

with d being a positive constant and s being a nonzero vector satisfying $\|s\| = 1$. Furthermore, we set $d = 1$ for convenience. The step-sizes sequences $\{\alpha(r)\}$ and $\{\beta(r)\}$ are assumed to be positive and non-increasing. The initial values $y_i(0)$ can be arbitrarily selected and $w_i(0) = 1$, $\forall i \in \mathcal{N}$.

Assumption 2: It holds

$$\max\{\|\partial g_i(s_i(r+1))\|, 1\} \le D, \quad \forall i \in \mathcal{N},\ r \in \mathcal{R} \tag{3}$$

where D is a positive constant.

Clearly, we can deduce that

$$\left\| \frac{s_i(r+1) - x_i(r+1)}{\|d_i(r+1)\|} \right\| \le 1.$$

Thus, under Assumptions 1 and 2, we can directly get the following result based on the result in [12]: $\forall r \in \mathcal{R}$,

$$\|x_i(r+1)-\bar{y}(r)\| \le D_1\rho^r + D_2\sum_{l=0}^{r-1}\rho^{r-1-l}(\alpha(l)+\beta(l)) \tag{4}$$

where $\bar{y}(r)=(1/n)\sum_{i=1}^{n} y_i(r)$; D_1 and D_2 are positive constants and $0<\rho<1$. For all $r\in\mathbb{R}$, we define $\phi_{i1}(r)=\beta(r)\|x_i(r+1)-\bar{y}(r)\|$, $\phi_{i2}(r)=\alpha(r)\|x_i(r+1)-\bar{y}(r)\|$, $\mathcal{N}_1(r)=\{i\in\mathcal{N}\mid d_i(r+1)=s\}$ and $\mathcal{N}_2(r)=\{i\in\mathcal{N}\mid d_i(r+1)\ne s\}$.

Lemma 1: $\forall r\in\mathcal{R}$, the following inequality holds for the $\bar{y}(r)$ under the algorithm (2) and Assumptions 1 and 2:

$$\begin{aligned}\|\bar{y}(r+1)-s^*\|^2 \le &\Big(1+\frac{2}{n}\sum_{i=1}^{n}\phi_{i1}(r)\Big)\|\bar{y}(r)-s^*\|^2-2\alpha(r)(g(\bar{s}(r))\\ &-g(s^*))-\Big(\frac{|\mathcal{N}_1(r)|}{n}\beta(r)\|\bar{y}(r)-\bar{s}(r)\|^2\\ &+2\frac{\underline{l}|\mathcal{N}_2(r)|}{n}\gamma(r)\|\bar{y}(r)-\bar{s}(r)\|\Big)+\varphi(r)\end{aligned} \tag{5}$$

where s^* is an arbitrary vector in the optimal solutions set, $\bar{s}(r)=P_X(\bar{y}(r))$, $\gamma(r)=\beta(r)-(D/\underline{l})\alpha(r)$ with $\underline{l}$ being a positive constant satisfying $0<\underline{l}\le 1$, and

$$\begin{aligned}\varphi(r)=&\frac{4D}{n}\sum_{i=1}^{n}\phi_{i2}(r)+\frac{2}{n}\sum_{i=1}^{n}\phi_{i1}(r)+2D^2\alpha^2(r)\\ &+2\beta^2(r)+\frac{D^2|\mathcal{N}_1(r)|}{n}\frac{\alpha^2(r)}{\beta(r)}.\end{aligned}$$

Proof 1: For maintaining a smooth presentation flow, the detailed proof of Lemma 1 is provided in the supplementary file. ■

Clearly, the summable condition of $\varphi(r)$ should be obtained for completing the convergence analysis. To this end, an assumption is shown blow on $\alpha(r)$ and $\beta(r)$ for guaranteeing the summable condition of $\varphi(r)$.

Assumption 3: The step-sizes $\{\alpha(r)\}$ and $\{\beta(r)\}$ satisfy the following conditions:

$$\sum_{r=0}^{\infty}\alpha(r)=\infty,\ \sum_{r=0}^{\infty}\alpha^2(r)<\infty,\ \beta(r)=\frac{D}{\underline{l}}\alpha(r)+\gamma(r)$$

with $\{\gamma(r)\}$ satisfying $\gamma(r)\ge\alpha(r)$, $\forall r\in\mathcal{R}$ and

$$\sum_{r=0}^{\infty}\gamma(r)=\infty,\ \sum_{r=0}^{\infty}\gamma^2(r)<\infty,\ \sum_{r=0}^{\infty}\frac{\alpha^2(r)}{\gamma(r)}<\infty.$$

Remark 1: In fact, the step-sizes satisfying the above conditions can be easily found. For example, we can choose $\alpha(r)=1/(r+1)$ and $\beta(r)=(D/\underline{l})(1/(r+1))+1/(r+1)^{0.6}$, implying that $\gamma(r)=1/(r+1)^{0.6}$. Easily, we can confirm that $\sum_{r=0}^{\infty}\alpha(r)=\infty$, $\sum_{r=0}^{\infty}\alpha^2(r)<\infty$, $\sum_{r=0}^{\infty}\gamma(r)=\infty$ and $\sum_{r=0}^{\infty}\gamma^2(r)<\infty$. Furthermore, we have

$$\sum_{r=0}^{\infty}\frac{\alpha^2(r)}{\gamma(r)}\le\sum_{r=0}^{\infty}\frac{1}{(r+1)^{1.4}}<\infty. \tag{6}$$

Additionally, under Assumption 3, noting that $\beta(r)\ge\gamma(r)$, $\forall r\in\mathcal{R}$, we have

$$\sum_{r=0}^{\infty}\beta(r)\ge\sum_{r=0}^{\infty}\gamma(r)>\infty,\ \sum_{r=0}^{\infty}\frac{\alpha^2(r)}{\beta(r)}\le\sum_{r=0}^{\infty}\frac{\alpha^2(r)}{\gamma(r)}<\infty. \tag{7}$$

Moreover, under Assumption 3, we can also get

$$\sum_{r=0}^{\infty}\beta^2(r)\le\sum_{r=0}^{\infty}2\Big(\frac{D^2}{\underline{l}^2}\alpha^2(r)+\gamma^2(r)\Big)<\infty. \tag{8}$$

Remark 2: One may concern that the definition of $\beta(r)$ involves the coefficient $D/\underline{l}$ whose denominator can not be exacted defined. In fact, as we state in the proof of Lemma 1, $\underline{l}$ is near to one, which implies that we can choose $\underline{l}$ a small positive constant.

Based on Lemma 1, we have enough foundations to conclude the main theorem below.

Theorem 1: All $x_i(r)$ $(i\in\mathcal{N})$ converge to a common optimal solution of (1) under Assumptions 1, 2 and 3, and the iterative rules (2).

Proof 2: $\forall r\in\mathcal{R}$, let

$$a(r)=\|\bar{y}(r)-s^*\|^2,\ b(r)=\frac{2}{n}\sum_{i=1}^{n}\phi_{i1}(r),\ c(r)=\varphi(r)$$
$$d(r)=2\alpha(r)(g(\bar{s}(r))-g^*)+h(\|\bar{y}(r)-\bar{s}(r)\|,r) \tag{9}$$

where g^* represents the optimal value and $h(y,r)$ denotes the following function $h(y,r)=|\mathcal{N}_1(r)|/n\beta(r)y^2+2\underline{l}|\mathcal{N}_2(r)|/n\gamma(r)y$. Next, we will verify that $\varphi(r)$ is summable. Recall the discussion in Remark 1, $\beta^2(r)$ and $\alpha^2(r)/\beta(r)$ are summable. Furthermore, $\alpha(r)\beta(r)\le(1/2)\times(\alpha^2(r)+\beta^2(r))$ indicates that $\alpha(r)\beta(r)$ is summable. Thus, based on [11, Lemma 1], we deduce that $\phi_{i1}(r)$ and $\phi_{i2}(r)$ are summable. Thus, the summable property of $\varphi(r)$ is established. Hence, noting that all the required conditions in [11, Lemma 2] holds, we have

$$\lim_{r\to\infty}\|\bar{y}(r)-s^*\|^2=c \tag{10}$$

$$\sum_{r=0}^{\infty}\alpha(r)(g(\bar{s}(r))-f^*)<\infty \tag{11}$$

$$\sum_{r=0}^{\infty}h(\|\bar{y}(r)-\bar{s}(r)\|,r)<\infty \tag{12}$$

where c is a positive constant.Noting that $\sum_{r=0}^{\infty}\alpha(r)$ and $\sum_{r=0}^{\infty}\beta(r)$ tend to infinity, we have

$$\liminf_{r\to\infty}g(\bar{s}(r))-g^*=0 \tag{13}$$

$$\liminf_{r\to\infty}h(\|\bar{x}(r)-\bar{s}(r)\|,r)=0. \tag{14}$$

Next, we will confirm that an optimal solution s_0^* exists such that $c=0$. Based on (13), a subsequence $\{\bar{s}(r_t)\}$ satisfying $\lim_{t\to\infty}g(\bar{s}(r_t))-g^*=0$ exists. Furthermore, taking (10) into consideration, we can deduce that $\{\bar{y}(r)\}$ is contained in a bounded set, which implies that $\{\bar{s}(r_t)\}$ is also contained in a bounded set. Consequently, there exists a convergent subsequence $\{\bar{s}(r_{t_l})\}$, and we assume that $\lim_{l\to\infty}\bar{s}(r_{t_l})=s_0^*$. Since the function g is continuous, we have $g^*=\lim_{l\to\infty}g(\bar{s}(r_{t_l}))=g(\lim_{l\to\infty}\bar{s}(r_{t_l}))=g(s_0^*)$, which indicates that s_0^* is an optimal solution of (1). Then, by replacing s^* by s_0^* in (9), we can directly obtain $\|\bar{y}(r)-s_0^*\|^2\le2(\|\bar{y}(r)-\bar{s}(r)\|^2+\|\bar{s}(r)-s_0^*\|^2)$, which yields

$$\begin{aligned}\frac{c}{2}&\le\liminf_{r\to\infty}\|\bar{y}(r)-\bar{s}(r)\|^2+\liminf_{r\to\infty}\|\bar{s}(r)-s_0^*\|^2\\ &\le\liminf_{r\to\infty}\|\bar{y}(r)-\bar{s}(r)\|^2.\end{aligned} \tag{15}$$

Additionally, considering the definition of the function h and based on (14), we can obtain

$$\liminf_{r\to\infty}\|\bar{y}(r)-\bar{s}(r)\|=0 \tag{16}$$

which implies that (10) holds. ■

Computer simulations: For verifying the theoretical result above, a simple simulation example which may be involved in the fields of smart grid and machine learning is given. Select $n=4$ and the following four local objective functions defined on $\mathbb{R}^2$ are considered in our simulation:

$$g_i(x)=\ln[1+e^{-a_i(b_ix^1+x^2)}]+|x^1|$$

with $i=1,2,3,4$, $a_i=(-1)^i$ and $b_i=0.01i$. Moreover, X is selected as $[0,2]\times[0,2]$. The unbalanced graphs sequence is selected to describe the communication topologies, which is depicted in Fig. 1. Furthermore, it is stipulated that Figs. 1(a)−1(d) are sequentially and repeatedly selected as $\mathcal{G}(r)$ for $r\in\mathcal{R}$. Moreover, according to the conditions given in Assumption 3, we select $\alpha(r)=1/(r+1)$ and $\beta(r)=1/(r+1)+1/(r+1)^{0.6}$, for all $r\in\mathcal{R}$. Finally, Fig. 2 clearly shows that all $x_i(r)$ converge to the optimal solution $x^*=(0,0.0325)$ under iterative rules (2).

Conclusion: This paper has investigated an optimization problem where the decision variable is contained in a closed convex set. Accordingly, we have designed a distributed algorithm over the time-varying and unbalanced graphs sequence. Furthermore, the favorable convergence of the designed algorithm has been shown based on the

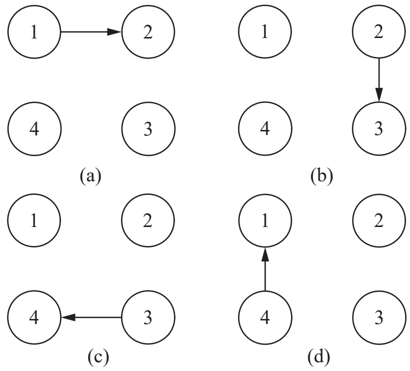

Fig. 1. Unbalanced graphs sequence with four agents.

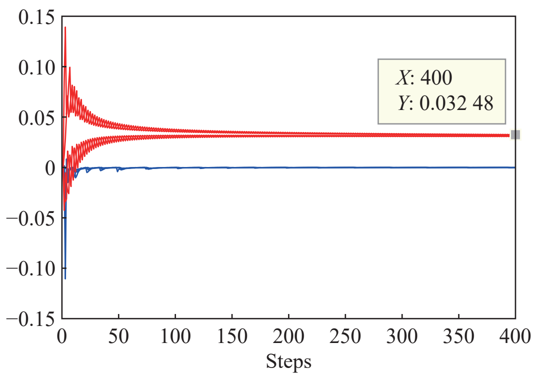

Fig. 2. Behaviors of the variables $x_i(k)$ under iterative rules (2).

convergence analysis. The advantage is that we address a challenging problem in distributed optimization field based on a novel distributed algorithm designed.

Acknowledgment: This work was supported by the Science and Technology Project of State Grid Zhejiang Electric Power CO., LTD. (5211JY19000X)

References

[1] Y. Zhu, W. Yu, G. Wen, and W. Ren, "Continuous-time coordination algorithm for distributed convex optimization over weight-unbalanced directed networks," *IEEE Trans. Circuits Syst Ⅱ: Exp Briefs*, vol. 66, no. 7, pp. 1202–1206, 2019.

[2] Q. Liu and J. Wang, "A second-order multi-agent network for bound-constrained distributed optimization," *IEEE Trans. Autom. Control*, vol. 60, no. 12, pp. 3310–3315, 2015.

[3] P. Lin, W. Ren, and J. A. Farrell, "Distributed continuous-time optimization: Nonuniform gradient gains, finite-time convergence, and convex constraint set," *IEEE Trans. Autom. Control*, vol. 62, no. 5, pp. 2239–2253, 2017.

[4] P. Yi, Y. Hong, and F. Liu, "Distributed gradient algorithm for constrained optimization with application to load sharing in power systems," *Syst. Control Lett.*, vol. 83, pp. 45–52, 2015.

[5] S. Yang, Q. Liu, and J. Wang, "A multi-agent system with a proportional-integral protocol for distributed constrained optimization," *IEEE Trans. Autom. Control*, vol. 62, no. 7, pp. 3461–3467, 2017.

[6] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Trans. Autom. Control*, vol. 54, no. 1, pp. 48–61, 2009.

[7] A. Nedić, A. Ozdaglar, and A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922–938, 2010.

[8] M. Zhu and S. Martínez, "On distributed convex optimization under inequality and equality constraints," *IEEE Trans. Autom. Control*, vol. 57, no. 1, pp. 151–164, 2012.

[9] J. Lei, H.-F. Chen, and H.-T. Fang, "Primal-dual algorithm for distributed constrained optimization," *Syst. Control Lett.*, vol. 96, pp. 110–117, 2016.

[10] V. Mai and E. Abed, "Distributed optimization over directed graphs with row stochasticity and constraint regularity," *Automatica*, vol. 102, pp. 94–104, 2019.

[11] H. Liu, W. X. Zheng, and W. Yu, "Distributed discrete-time algorithms for convex optimization with general local constraints on weight-unbalanced digraph," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 1, pp. 51–64, 2021.

[12] A. Nedić and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Trans. Autom. Control*, vol. 60, no. 3, pp. 601–615, 2015.

[13] A. Nedić, A. Ozdaglar, and W. Shi, "Achieving geometric convergence for distributed optimization over time-varying graphs," *SIAM J. Optim.*, vol. 27, no. 4, pp. 2597–2633, 2017.

[14] G. Qu and N. Li, "Harnessing smoothness to accelerate distributed optimization," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 3, pp. 1245–1260, 2018.

[15] C. Xi, V. S. Mai, R. Xin, E. H. Abed, and U. A. Khan, "Linear convergence in optimization over directed graphs with row-stochastic matrices," *IEEE Trans. Autom. Control*, vol. 63, no. 10, pp. 3558–3565, 2018.

[16] F. Saadatniaki, R. Xin, and U. A. Khan, "Decentralized optimization over time-varying directed graphs with row and column-stochastic matrices," *IEEE Trans. Autom. Control*, vol. 65, no. 11, pp. 4769–4780, 2020.

[17] S. Liang, L. Y. Wang, and G. Yin, "Dual averaging push for distributed convex optimization over time-varying directed graph," *IEEE Trans. Autom. Control*, vol. 65, no. 4, pp. 1785–1791, 2020.