Distributed Inertial Continuous and Discrete Time Algorithms for Solving Resource Allocation Problem

You Zhao, Xiaofeng Liao, *Fellow, IEEE*, and Xing He, *Member, IEEE*

***Abstract*— In this article, we investigate several distributed inertial algorithms in continuous and discrete time for solving resource allocation problem (RAP), where its objective function is convex or strongly convex. First, the original RAP is equivalently transformed into a distributed unconstrained optimization problem by introducing an auxiliary variable. Then, two distributed inertial continuous time algorithms and two discrete time algorithms are proposed and the rates of their convergence based on the gap between the objective function and their optimal function are determined. Our first distributed damped inertial continuous time algorithm is designed for RAP with a convex function, it achieves convergence rate at $O(\frac{1}{t^2})$ based on Lyapunov analysis method, and then we design a rate-matching distributed damped inertial discrete time algorithm by exploiting implicit and Nesterov's discretization scheme. Our second distributed fixed inertial discrete time algorithm is designed to deal with the RAP with a strongly convex objective function. Noteworthy, the transformed distributed problem is no longer strongly convex even though the original objective function is strongly convex, but it satisfies the Polyak-Łjasiewicz (PL) and quadratic growth (QG) conditions. Inspired by the Heavy-Ball method, a distributed fixed inertial continuous time algorithm is proposed, it has an explicit and accelerated exponential convergence rate. Later, a rate-matching accelerated distributed fixed inertial discrete time algorithm is also obtained by applying explicit, semi-implicit Euler discretization and sufficient decrease update schemes. Finally, the effectiveness of the proposed distributed inertial algorithms is verified by simulation.**

***Index Terms*—Distributed inertial algorithms, resource allocation, rate-matching, accelerated convergence, linear convergence rate.**

Manuscript received 25 April 2022; revised 12 February 2023; accepted 19 February 2023. Date of publication 23 February 2023; date of current version 25 October 2023. This work was supported in part by the National Key R&D Program of China under Grant 2018AAA0100101, in part by the National Natural Science Foundation of China under Grants 61932006 and 62176027, and in part by the Fundamental Research Funds for the Central Universities under Grant XDJK2020TY003. Recommended for acceptance by Prof. Michael Small. *(Corresponding author: Xiaofeng Liao.)*

You Zhao and Xiaofeng Liao are with the College of Computer Science, Chongqing University, Chongqing 400044, China (e-mail: zhaoyou1991sdtz@163.com; xfliao@cqu.edu.cn).

Xing He is with the Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, School of Electronics and Information Engineering, Southwest University, Chongqing 400715, China (e-mail: hexingdoc@swu.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSE.2023.3248267, provided by the authors.

Digital Object Identifier 10.1109/TNSE.2023.3248267

I. Introduction

RESOURCE allocation problem (RAP) is one of the important problems of network optimization that has been extensively studied in power systems [1], antenna wireless edge computing [2], cloud computing [3], machine learning [4], etc. Resource allocation means assigning available resources to gain maximum benefit. This issue has been modeled as an optimization problem with the objective function being the sum of local cost functions and the constraints being coupled to achieve supply and demand equilibrium in [5] and [6]. To solve RAP, many centralized algorithms have been investigated, for instance, the Newton-Raphson method [7] and particle swarm algorithm [8]. The centralized algorithms require a central node to process and transmit the global variable information. However, as the network scales up, this central node will endure heavy communication and computational burdens, unexpected single points of failure, weaker scalability and robustness, and privacy issues.

A. Literature Review

Recently, many distributed algorithms have been proposed for RAP, which can be divided into two categories: **distributed discrete time algorithms** and **distributed continuous time algorithms**.

For distributed discrete time algorithms: Yang et al. [9] proposed an incremental cost consensus algorithm for RAP and provided its global convergence analysis. Necoara [10] proposed a series of random coordinate descent algorithms that had a convergence rate $O(\frac{1}{k})$ and a linear convergence rate for RAP with convex and strongly convex objective functions, respectively. Doan et al. [11] investigated a distributed algorithm for RAP over time-varying networks, and it had a convergence rate $O(\frac{LBn^2}{k})$ for a convex objective function and a linear convergence rate $O((1-\frac{\mu}{4Ln^2})^{\frac{k}{B}})$ for the strongly convex objective function. Based on the Lagrangian function and an ADMM-like method, Aybat and Hamedani [12] proposed a class of distributed algorithms to solve RAP with a convex objective function, and they had a convergence rate at $O(\frac{1}{k})$. Nedić et al. [13] investigated a series of Mirror-EXTRA algorithms for RAP and derived an improved convergence rate, that is, an $o(\frac{1}{k})$ when the objective function was convex. Lü et al. [14] developed a distributed Lagrangian momentum algorithm for RAP under a directed graph. Liu et al. [15] studied a distributed algorithm to deal with the combined economic environmental dispatch problem of Microgrids. Liu et al. [48] presented a distributed optimization algorithm for economic dispatch with region partitioning. Wang et al. [16] presented a transactive energy sharing (TES) approach in a distributed manner for a microgrid operator (MGO) and multiple distributed energy resources (DER) aggregators to minimize the total social cost. Li et al. [17] presented a distributed market clearing algorithm based on the alternating direction method of multipliers (ADMM) and the best response seeking is developed to solve the optimal bidding problem of MW-level microgrids aggregator (MGAs) and the optimal dispatch problem of the subtransmission network operator (SNO).

For distributed continuous time algorithms: In [18], Cherukuri and Cortés proposed Laplacian-gradient dynamics for RAP with an exponential convergence rate when the objective function was strongly convex. Later, under a strongly connected and weight-balanced digraph, Cherukuri and Cortés [19] designed a distributed coordination algorithm (DAC+LMG) for RAP. Based on a θ-logarithmic barrier function, a distributed incremental cost consensus (first-order) algorithm [20] and a distributed primal-dual algorithm [21] for RAP were investigated, they had exponential convergence rates when the objective function was strongly convex. Based on the primal-dual dynamical method and projection operators, Yi et al. [22] investigated an initialization-free distributed continuous time algorithm for RAP over an undirected network, and it had an exponential convergence rate for the strongly convex objective function but only provided the guarantee of convergence property when the objective function was convex. Zeng et al. [23] investigated a distributed derivative feedback algorithm for an extended monotropic optimization problem (a generalized version of RAP) with a convergence rate at $O(\frac{1}{t})$ when the objective function was convex. Bai et al. [24] presented distributed algorithms based on dynamic average consensus, leader-following consensus, and the saddle point dynamics to solve the economic dispatch problem. Wang et al. [25] proposed distributed dynamical systems over state-dependent communication networks to deal with the nonsmooth resource allocation problem with a convergence rate of $O(\frac{1}{t})$. Second-order continuous time algorithms (based on the inertial method) for RAP have been investigated in [26], [27], [28]. In [26] and [28], the authors only provided the convergence property of the proposed distributed algorithm. In [27], the proposed second-order continuous time algorithm that had an exponential convergence rate for RAP with a strongly convex function, while only the convergence property was analyzed when the objective function was convex. Zhu et al. [29] investigated distributed proximal gradient continuous time algorithms for RAP with a nonsmooth objective function. Guo et al. [30] proposed an adaptive distributed continuous time algorithm with duplex control laws to deal with nonsmooth resource allocation problems, and only provided convergence analysis. Zeng et al. [31] investigated a distributed primal-dual accelerated method based on the dynamical primal-dual method and Nesterov's accelerated scheme to solve the distributed extended monotropic optimization problem with a convergence rate of $O(\frac{1}{t^2})$.

B. Motivations

As for the RAP, the most existing distributed continuous and discrete time algorithms were studied independently. In fact, continuous time and discrete time algorithms are closely related, and the researches to build bridges between accelerated continuous time algorithms and corresponding rate-matching discrete time algorithms are in full swing in centralized optimization [32], [33], [34], [35], [36], [37]. To the authors' knowledge, there are very little works to investigate the relationship between distributed continuous time algorithms and rate-matching distributed discrete time algorithms, except for the following recent works: Chen et al. [38] revealed the DIGing algorithm was a discretization of a distributed second-order continuous time algorithm. To solve RAP with strongly convex functions, Shi et al. [39] first proposed a continuous time distributed algorithm with an exponential convergence, then presented a rate-matching distributed discrete time algorithm. In addition, the convergence rates of distributed continuous time and discrete time algorithms mentioned above for solving RAP with a convex objective function are slower than the optimal convergence rates $O(\frac{1}{t^2})$ and $O(\frac{1}{k^2})$. In addition, solving RAP with a strongly convex objective function, the distributed discrete time algorithm in [11] had a convergence rate $O(p^k), p \approx 1 - C\frac{\mu}{L}$ (C is a positive constant, $\frac{\mu}{L}$ is the condition number) that is less than $O(q^k)$, $q \approx 1 - C\sqrt{\frac{\mu}{L}}$ in [40]. Based on the above discussion, it is of great significance to establish the relationship between distributed continuous time and distributed discrete time algorithms for RAP, and to enable the proposed distributed algorithms with faster convergence rates, i.e., for RAP with a convex objective function, the distributed continuous and discrete time algorithms respectively had convergence rates $O(\frac{1}{t^2})$ and $O(\frac{1}{k^2})$. For RAP with a strongly convex objective function, the distributed continuous time algorithms had faster exponential convergence rates, and the distributed discrete time algorithms had a convergence rate of $O(q^k)$, $q \approx 1 - C\sqrt{\frac{\mu}{L}}$.

C. Statement of Contributions

The main interest of this article is to investigate distributed inertial continuous time algorithms and their corresponding rate-matching discrete time algorithms for RAP over an undirected network, and the work of this article is motivated by the centralized accelerated algorithms [33], [41], [42], [43], [44] together with distributed continuous time algorithms [20], [26] and [45]. To deal with this issue, the original RAP is first equivalently transformed into a distributed unconstrained optimization problem by introducing an auxiliary variable, then a distributed damped inertial continuous time algorithm and its rate-matching distributed discrete time algorithm are also studied for RAP with a convex objective function. Moreover, a distributed fixed inertial continuous time algorithm and its rate-matching discrete time algorithm are also presented to address RAP with a strongly convex function. We wish to extend the theory of distributed accelerated continuous and discrete time optimization, and to provide a new perspective for designing distributed continuous time and discrete time algorithms. In summary, the main

TABLE I
COMPARISONS OF DIFFERENT DISTRIBUTED CONTINUOUS AND DISCRETE TIME ALGORITHMS

Algorithm	Continuous/ Discrete	Objective function	Convergence rate
Incremental cost consensus (first-order) algorithm [20]	Continuous	Convex	✗
DAC+LMG [19]	Continuous	Convex	✗
DDFA [23]	Continuous	Convex	$O\left(\frac{1}{t}\right)$
DDS [25]	Continuous	Convex	$O\left(\frac{1}{t}\right)$
Incremental cost consensus (second-order) algorithm [26]	Continuous	Convex	✗
DDICA (23)	Continuous	Convex	$O\left(\frac{1}{t^2}\right)$
RCDA [10]	Discrete	Convex	$O\left(\frac{1}{k}\right)$
GBP [11]	Discrete	Convex	$O\left(\frac{1}{k}\right)$
Distributed ADMM-like method [12]	Discrete	Convex	$O\left(\frac{1}{k}\right)$
Mirror-EXTRA [13]	Discrete	Convex	$o\left(\frac{1}{k}\right)$
DDIDA (27)	Discrete	Convex	$O\left(\frac{1}{k^2}\right)$
DRAA [27]	Continuous	Strongly convex	Exponential convergence (⊠)
SOCA-DE [28]	Continuous	Strongly convex	Exponential convergence (⊠)
DCTI [39]	Continuous	Strongly convex	Exponential convergence (⊠)
DFICA (29)	Continuous	Strongly convex	$O\left(e^{-\left(2-\sqrt{2}\right)\sqrt{\mu}\lambda_2 t}\right)$
GBP [11]	Discrete	Strongly convex	$O\left(\left(1-\frac{\mu}{4Ln^2}\right)^{\frac{k}{B}}\right)$
D-DLM [14]	Discrete	Strongly convex	Linear convergence (⊠)
DTDA [39]	Discrete	Strongly convex	Linear convergence (⊠)
DICC-BOA [47]	Discrete	Strongly convex	Linear convergence (⊠)
DFIDA (33)	Discrete	Strongly convex	$O\left(\left(1-\frac{(2-\sqrt{2})\sqrt{\mu}\lambda_2}{\sqrt{l\varpi_{\max}(L)}}+o\left(\frac{\sqrt{\mu}\lambda_2}{\sqrt{l\varpi_{\max}(L)}}\right)\right)^k\right)$

✗means that convergence rate analysis is not provided, ⊠denotes that it cannot be exactly written as $O(\bullet)$, because the value of $\bullet$ depends on the complex coupling of many non-exact parameters.

contributions of our paper can be summarised in the following (see Table I):

- For RAP with a convex objective function, a distributed damped inertial continuous time algorithm (DDICA) is designed. In contrast to works in [19], [20], [21], [22] and [26], [27], [28], DDICA has an explicit and accelerated convergence rate $O(\frac{1}{t^2})$, which is faster than the algorithm in [23]. Specifically, the convergence rate of DDICA can be obtained easily and straightforwardly by using the Lyapunov method.
- By the Nesterov's discretization scheme, a novel rate-matching distributed damped inertial discrete time algorithm (DDIDA) with fixed step size is proposed. Compared with [10], [11], [12], [13], DDIDA has faster and optimal convergence rate $O(\frac{1}{k^2})$, which is easily obtained by a discrete Lyapunov method.
- For RAP with a strongly convex objective function, we propose a distributed fixed inertial continuous time algorithm (DFICA) that has an explicit exponential convergence rate. Compared with the second-order distributed continuous time algorithms [26], [27], [28] (based on the inertial method), we present the optimal inertial parameters of RAP with a strongly convex objective function.
- On account of explicit discretization scheme and sufficient decrease update scheme, a rate-matching distributed fixed inertial discrete time algorithm (DFIDA) is designed, and it has a faster linear convergence rate $O(q^k)$, $q \approx 1 - C\sqrt{\frac{\mu}{L}}$.
- In addition, we have found that the transformed distributed unconstrained optimization problem satisfies Polyak-Łojasiewicz (**PL**) and quadratic growth (**QG**) conditions when the objective function in RAP is strongly convex. In reality, some non-convex functions satisfy **PL** condition [44], so the results of DFICA and DFIDA can also be extended to solve some distributed non-convex optimization problems.

The remainder of this article is organized as follows. In Section II, graph theory, the resource allocation problem, and its equivalently transformed problem are introduced. Section III, distributed damped inertial continuous and discrete time algorithms for RAP with a convex function are provided. In Section IV, distributed fixed inertial continuous and discrete time algorithms for RAP with a strongly convex function are discussed. Two examples are presented to illustrate the superior and effectiveness of our proposed distributed accelerated algorithms in Section V. Finally, we conclude the paper in Section VI. To better understand the structure of our manuscript

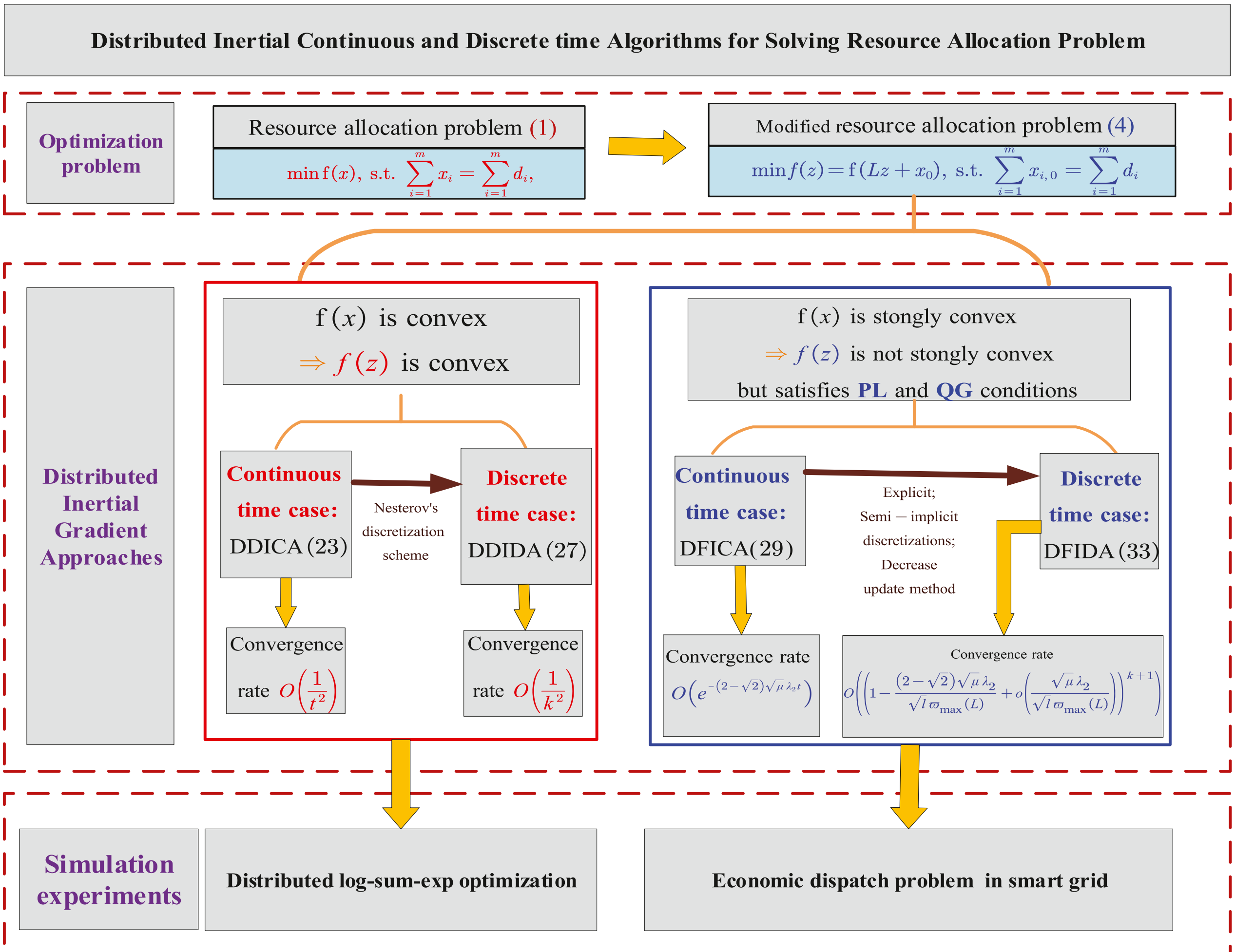

Fig. 1. Overall block diagram of this article.

and the relationship between distributed inertial continuous and discrete time algorithms, we present an overall block diagram in Fig. 1.

Notations: Let column vectors $x = (x_1,\ x_2, \{\ldots,\ x_n)^T$ and $y = (y_1,\ y_2, \{\ldots,\ y_n)^T, x^T y = \sum_{i=1}^{n} x_i y_i$ is the inner product of x and y, x_i is the i-th element of x. $\|x\| = (\sum_{i=1}^{n} x_i^2)^{\frac{1}{2}}$ denotes the Euclidean norm. Let l_i and μ_i be Lipschitz constant and strongly convex constant of $\nabla \mathrm{f}_i$, respectively, then set $l = \max\{l_1, \ldots, l_n\}, \mu = \max\{\mu_1, \ldots, \mu_n\}$. $L \in R^{n\times n}$ is a Laplacian matrix of an undirected graph, then λ_2 and $\varpi_{\max}(L)$ are the second largest and the largest eigenvalue of Laplacian matrix L, respectively. $\mathcal{L}^1$ represents a set of Lebesgue integrable functions. In addition, the symbols z, x, y denote the variables in DDICA and DDIDA respectively, while z, x, y, w, v represent the variables in algorithms DFICA and DFIDA, respectively. The original optimization objective function in RAP is $\mathrm{f}(x)$ and the modified objective function is $f(z) = \mathrm{f}(Lz + x_0)$ by using an auxiliary variable z. The functions $\varGamma^k, \varGamma(t)$ have convergence rate $O(\frac{1}{k^\rho})$, $O(\frac{1}{t^\rho})$ with $\rho = 1, 2$ if $\lim\limits_{k\to+\infty} \sup\ k^\rho \varGamma^k < +\infty$, $\lim\limits_{t\to+\infty} \sup\ t^\rho \varGamma(t) < +\infty$ hold respectively. In addition, it is said to have a convergence rate $o(\frac{1}{k})$ if $\lim\limits_{k\to+\infty} \sup\ k\varGamma^k = 0$.

II. Graph Theory, Resource Allocation Problem and Its Equivalent Reformulation

This section introduces the preliminaries of algebraic graph theory. Subsequently, the resource allocation problem (RAP) and its equivalent reformulation are presented.

A. Graph Theory

A weighted undirected communication topology among agents is denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ with the set of nodes $\mathcal{V} = \{1, 2, \ldots, N\}$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denoting the edge set of pairs of agents, i.e., the communication links. The couples of agents in an undirected graph are unordered, where the (i, j) means that there exists an information exchange between both agents i and agents j. A path in undirected graph between agent i and agent j is a sequence of edges of the form (i, i_1), $(i_1, i_2), \ldots, (i_s, j)$, where the $i_1, \cdots, i_s, j$ are distinct agents. The weighted adjacency matrix is defined by A, where A is a $N \times N$ nonnegative matrix with $a_{ij} = a_{ji}$ if $(i, j) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. The Laplacian matrix of graph $\mathcal{G}$ is defined as L with $L_{ij} = \sum_{j=1, j\neq i}^{N} a_{ij}$ and $L_{ij} = -a_{ij}$, which indicates $\sum_{j=1}^{N} a_{ij} = 0$.

Consider a RAP over an undirected network with n agents as follows:

$$\min \mathrm{f}(x) = \sum_{i=1}^{n} \mathrm{f}_i(x_i), \tag{1a}$$

$$\text{s.t.} \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} d_i, \tag{1b}$$

where f_i is continuously differentiable and convex, x_i, d_i are the local decision variable and resource demand variable of agent i respectively. $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} d_i$ represents the total demand of loads.

Assumption 1: The undirected graph of the multi-agents network is connected.

Assumption 2: Each function $\mathrm{f}_i(x_i)$, $i = 1, \ldots, n$ is convex and l_i-smooth, that is, it satisfies $\mathrm{f}_i(x_i) - \mathrm{f}_i(\mathrm{x}_i) \geqslant \nabla \mathrm{f}_i(\mathrm{x}_i)^T(x_i - \mathrm{x}_i)\ \forall x_i \neq \mathrm{x}_i$, and $\nabla^2 \mathrm{f}_i(x_i) \leqslant l_i$.

Assumption 3: The function $\mathrm{f}_i(x_i)$, $i = 1, \ldots, n$ is $\mu_i > 0$-strongly convex and l_i-smooth such that $\mathrm{f}_i(x_i) - \mathrm{f}_i(\mathrm{x}_i) \geqslant \nabla \mathrm{f}_i(\mathrm{x}_i)^T(x_i - \mathrm{x}_i) + \frac{\mu_i}{2}\|x_i - \mathrm{x}_i\|^2\ \forall x_i \neq \mathrm{x}_i$ and $l_i \geqslant \nabla^2 \mathrm{f}_i(x_i) \geqslant \mu_i > 0$.

It can be derived from Assumptions 1 and 2 that the strongly convex assumption condition serves as a special case of the convexity condition and is a stronger condition than the convexity assumption.

B. Optimality and Reformulation

Lemma 1: Suppose Assumptions 1, 2 or 3 hold, $x^* = (x_1^*, \ldots, x_n^*)^T \in R^n$ is an optimal solution to the problem (1) if and only if it satisfies

$$\begin{cases} \nabla \mathrm{f}_1(x_1^*) = \nabla \mathrm{f}_2(x_2^*) = \ldots = \nabla \mathrm{f}_n(x_n^*), \\ \sum_{i=1}^{n} x_i^* = \sum_{i=1}^{n} d_i. \end{cases} \tag{2}$$

Proof: From Karush-Kuhn-Tucker conditions, $x_i^* \in R$ is an optimal solution of problem (1) if there exists Υ^* such that

$$\sum_{i=1}^{n} x_i^* = \sum_{i=1}^{n} d_i, \nabla \mathrm{f}_i(x_i^*) = \Upsilon^*, i = 1, \ldots, n,$$

which implies $\nabla \mathrm{f}_1(x_1^*) = \nabla \mathrm{f}_2(x_2^*) = \ldots = \nabla \mathrm{f}_n(x_n^*)$. Thus, the proof is completed. □

If *Assumption 1* and $\sum_{i=1}^{n} x_{i,0} = \sum_{i=1}^{n} d_i$ are satisfied, then the sum of $x_i, i = 1, \ldots, n$ in problem (1) satisfies

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \sum_{j \in \mathcal{N}_i} a_{ij}(z_i - z_j) + \sum_{i=1}^{n} x_{i,0} = \sum_{i=1}^{n} d_i, \tag{3}$$

where z_i is an auxiliary variable. The global equality constraint is satisfied since $x_i = \sum_{j \in \mathcal{N}_i} a_{ij}(z_i - z_j) + x_{i,0}$, $(i = 1, \ldots, n)$. Thus, the problem (1) can be equivalently converted into the following problem [45]:

$$\begin{aligned} \min f(z) &= \mathrm{f}(Lz + x_0) \\ &= \sum_{i=1}^{n} \left(f_i(z_i) = \mathrm{f}_i \left(\sum_{j \in \mathcal{N}_i} a_{ij}(z_i - z_j) + x_{i,0} \right) \right), \\ \text{s.t.} &\sum_{i=1}^{n} x_{i,0} = \sum_{i=1}^{n} d_i, \end{aligned} \tag{4}$$

where $z = (z_1, \ldots, z_n)^T \in R^n$.

C. The Gradient and Hessian Matrix Properties of $f(Z)$

For problem (4), the derivative of $x_i, i = 1, \ldots, n$ with respect to $z_j, j = 1, \ldots, n$ is given by

$$\frac{\mathrm{d}x_i}{\mathrm{d}z_j} = \begin{cases} \sum_{j \in \mathcal{N}_i} a_{ij}, & i = j, \\ -a_{ij}, & i \neq j. \end{cases} \tag{5}$$

Thus, we obtain

$$\frac{\mathrm{d}f_i}{\mathrm{d}z_j} = \frac{\mathrm{d}f_i}{\mathrm{d}x_i}\frac{\mathrm{d}x_i}{\mathrm{d}z_j} = \begin{cases} \sum_{j \in \mathcal{N}_i} a_{ij} \nabla \mathrm{f}_i(x_i), & i = j, \\ -a_{ij} \nabla \mathrm{f}_i(x_i), & i \neq j. \end{cases} \tag{6}$$

Note that the gradient of function $\nabla \mathrm{f}(x)$ is shown as $\nabla \mathrm{f}(x) = (\nabla \mathrm{f}_1(x_1), \ldots, \nabla \mathrm{f}_n(x_n))^T$. Applying (6) and the Laplacian matrix L, one has

$$\nabla f(z) = L \nabla \mathrm{f}(x). \tag{7}$$

Furthermore, we get the Hessian matrix of $f(z)$ as follows:

$$\nabla^2 f(z) = L \nabla^2 \mathrm{f}(x) L. \tag{8}$$

It follows from (7) that the function $f(z)$ is Lipschitz continuous, such that $\|\nabla f(z^1) - \nabla f(z^2)\| \leqslant l\varpi_{\max}(L)^2\|z^1 - z^2\|, \forall z^1, z^2 \in R^n$, since $\|\nabla \mathrm{f}(x^1) - \nabla \mathrm{f}(x^2)\| \leqslant l\|x^1 - x^2\|, \forall x^1, x^2 \in R^n$. While, on the basis of (8), the Hessian matrix $\nabla^2 f(z)$ is semi-positive definite even though function $\mathrm{f}(x)$ is a μ-strongly convex function since the Laplacian matrix L is semi-positive. It's worth noting that if $\mathrm{f}(x)$ is a μ-strongly convex function, $f(z)$ satisfies the Polyak-Łjasiewicz (**PL**) and quadratic growth (**QG**) conditions.

Lemma 2: The $f(z)$ satisfies the following conditions

$$\begin{aligned} &\frac{\mu\lambda_2^2}{4}(f(z) - f(z^*)) \leqslant \frac{1}{2}\|\nabla f(z)\|^2, \quad (\mathbf{PL}) \\ &f(z) - f(z^*) \geqslant \frac{\mu\lambda_2^2}{2}\|z(t) - z^*\|^2, \quad (\mathbf{QG}) \end{aligned} \tag{9}$$

if $\mathrm{f}(x)$ is a μ-strongly convex function with $\mu = \min\{\mu_1, \ldots, \mu_n\}$.

Proof: 1) For PL [45]: Let $z^* = (z_1^*, \ldots, z_n^*)^T \in R^n$ be an optimal solution to problem (4), such that it is a trivial solution if $z^* \in R^n \backslash b\mathbf{1}$ for any constant b. In this proof, we mainly focus on the nontrivial case, i.e., the convex and compact set $z^* \in R^n \backslash b\mathbf{1}$.

It follows from the convex property that

$$f\left(\mathrm{z}\right)=f(z)+\nabla f(z)^{T}\left(\mathrm{z}-z\right)+\frac{1}{2}\left(\mathrm{z}-z\right)^{T}\nabla^{2}f\left(\hat{\mathrm{z}}\right)\left(\mathrm{z}-z\right), \tag{10}$$

where $\hat{\mathrm{z}}=z+\eta(\mathrm{z}-z)$ with $\eta\in(0,1)$.

Since the undirected graph communication is connected, then its Laplacian matrix has the following eigenvalues $0=\lambda_1<\lambda_2\leqslant\lambda_3\leqslant\cdots\leqslant\lambda_n$ and their corresponding eigenvectors $\mathbf{1},\zeta_2,\zeta_3,\ldots,\zeta_n$ with $\|\zeta_i\|=1, i=2,\ldots,n$. Thus, $z-z^*$ can be represented as follows

$$z-z^{*}=\mathfrak{a}_1\mathbf{1}+\mathfrak{a}_2\zeta_2+\cdots+\mathfrak{a}_n\zeta_n, \tag{11}$$

where $\mathfrak{a}_i$, $(i=1,\ldots,n)$ are constants. Let $\zeta=\mathfrak{a}_2\zeta_2+\cdots+\mathfrak{a}_n\zeta_n$, then $\|\zeta\|^2=\mathfrak{a}_2^2+\cdots+\mathfrak{a}_n^2$. From (10), one has

$$f\left(z^{*}\right)=f(z)+\nabla f(z)^{T}\left(z^{*}-z\right)+\int_{0}^{1}\int_{0}^{\tau}\left(z^{*}-z\right)^{T}\nabla^{2}f\left(z+\eta\left(z^{*}-z\right)\right)\left(z^{*}-z\right)d\eta d\tau, \tag{12}$$

and

$$f(z)-f\left(z^{*}\right)\leqslant\nabla f(z)^{T}\left(z-z^{*}\right). \tag{13}$$

From (11) and (13) one has

$$\begin{aligned}f(z)-f\left(z^{*}\right)&\leqslant\nabla f(z)^{T}\left(\mathfrak{a}_1\mathbf{1}+\zeta\right)\\&=\nabla f(x)^{T}L\mathfrak{a}_1\mathbf{1}+\nabla f(z)^{T}\zeta\\&\leqslant||\nabla f(z)||||\zeta||,\end{aligned} \tag{14}$$

the last inequality holds due to $\mathfrak{a}_1L\mathbf{1}=\mathbf{0}$.

Recalling again (12) and the strongly convex property of $\mathrm{f}(x)$, we have

$$\begin{aligned}&f(z)-f\left(z^{*}\right)\\&\geqslant\frac{\mu}{2}\|\lambda_2\mathfrak{a}_2\zeta_2+\cdots+\lambda_n\mathfrak{a}_n\zeta_n\|^2\\&\geqslant\frac{\mu\lambda_2^2}{2}\left(\mathfrak{a}_2^2+\cdots+\mathfrak{a}_n^2\right)=\frac{\mu\lambda_2^2}{2}\|\zeta\|^2,\end{aligned} \tag{15}$$

Combining (14) and (15), we have

$$\frac{\mu\lambda_2^2}{2}\|\zeta\|^2\left(f(z)-f\left(z^{*}\right)\right)\leqslant\|\nabla f(z)\|^2\|\zeta\|^2. \tag{16}$$

The vector $\zeta\neq\mathbf{0}$ since the solution is nontrivial. Dividing (16) by $\|\zeta\|^2$ yields

$$\frac{\mu\lambda_2^2}{4}\left(f(z)-f\left(z^{*}\right)\right)\leqslant\frac{1}{2}\|\nabla f(z)\|^2. \tag{17}$$

2) For QG: The proof is inspired by the work in [49]. Define a function $h(z)=\sqrt{f(z)-f(z^*)}$. From (17) with $z_0\neq z^*$, one has $\|\nabla h(z)\|^2=\frac{\|\nabla f(z)\|^2}{f(z)-f(z^*)}=\frac{\mu\lambda_2^2}{2}$, i.e., $\|\nabla g(z)\|^2\geqslant\sqrt{\frac{\mu\lambda_2^2}{2}}$. Note that $f(z)$ satisfies **PL** condition (17) and we have that $f(z)$ is an invex function (i.e., every equilibrium point of f is the global minimiser of f), thus, $h(z)$ is a positive invex function with a closed optimal solution set S^* with $h(\hat{z})=0$, if $z\in S^*$. By solving the following differential equation:

$$\dot{z}=-\nabla h\left(z(t)\right), z_0\notin S^{*}. \tag{18}$$

Since $\nabla g(z)$ is bounded from below, i.e., $\|\nabla g(z)\|^2\geqslant\sqrt{\frac{\mu\lambda_2^2}{2}}$ and $h(z)$ is a positive invex function, we know that function $g(z)$ is bounded below. By running the trajectory of (18), the function $h(z)$ is sufficiently reduced to 0, and $z(t)$ tends to the optimal solution set S^*. Therefore, there exists a constant T_e, such that $z(\mathrm{T}_e)\in S^*$ (at this point the differential equation is well-posed). To obtain the results above mentioned, we use the following condition

$$\begin{aligned}h\left(z_0\right)-h\left(z(t)\right)&=-\int_0^{\mathrm{T}}\nabla h\left(z(t)\right)^{T}\dot{z}(t)dt\\&=\int_0^{\mathrm{T}}\|\nabla h\left(z(t)\right)\|^2dt\geqslant\frac{\mu\lambda_2^2}{2}\mathrm{T}.\end{aligned} \tag{19}$$

The nonnegative property of $h(z(t))$ implies that there exists a $\mathrm{T}\leqslant\frac{2h(z_0)}{\mu\lambda_2^2}$, such that $z(\mathrm{T})\in S^*$.

Next, we will show that the length of orbit $z(t)$ from initial point z_0 by $\mathbb{L}(z_0)$, which is derived from the following inequality:

$$\begin{aligned}\mathbb{L}\left(z_0\right)&=\int_0^{\mathrm{T}}\|\dot{z}(t)\|dt=\int_0^{\mathrm{T}}\|\nabla h\left(z(t)\right)\|dt\\&\geqslant\|z_0-z^*\|,\end{aligned} \tag{20}$$

where the above inequality holds since the orbit is a path from z_0 to an optimal solution in optimal set S^*.

Using (19) again, one has

$$\begin{aligned}&h\left(z_0\right)-h\left(z\left(\mathrm{T}_e\right)\right)=\int_0^{\mathrm{T}_e}\|\nabla h\left(z(t)\right)\|^2dt\\&\geqslant\sqrt{\frac{\mu\lambda_2^2}{2}}\int_0^{\mathrm{T}_e}\|\nabla h\left(z(t)\right)\|dt\geqslant\sqrt{\frac{\mu\lambda_2^2}{2}}\|z_0-z^*\|.\end{aligned} \tag{21}$$

Note that $h(z(\mathrm{T}_e))=0$, then, $h(z_0)\geqslant\sqrt{\frac{\mu\lambda_2^2}{2}}\|z_0-z^*\|$. In addition, let $z_0=z$, we obtain $h(z(t))\geqslant\sqrt{\frac{\mu\lambda_2^2}{2}}\|z(t)-z^*\|$, i.e., $h(z(t))^2=f(z)-f(z^*)\geqslant\frac{\mu\lambda_2^2}{2}\|z(t)-z^*\|^2$. Thus, the proof is completed. □

III. Distributed Damped Inertial Algorithms for RAP With Convex Functions

A. Distributed Damped Inertial Continuous Time Algorithm (DDICA)

In this section, to solve problem (4) with a convex objective function and get an accelerated convergence rate $O(\frac{1}{t^2})$, we propose a distributed damped inertial continuous time algorithm (DDICA) of agent i as follows:

$$\begin{cases}\ddot{z}_i(t)+\frac{\alpha}{t}\dot{z}_i(t)=-\sum_{j\in\mathcal{N}_i}\left(\nabla \mathrm{f}_i\left(x_i(t)\right)-\nabla \mathrm{f}_j\left(x_j(t)\right)\right)\\x_i(t)=\sum_{j\in\mathcal{N}_i}a_{ij}\left(z_i-z_j\right)+x_{i,0},\end{cases} \tag{22}$$

where $\alpha\geq 3$, $t>0$. The compact form of DDICA (22) is

$$\begin{cases}\ddot{z}(t)+\frac{\alpha}{t}\dot{z}(t)=-L\nabla \mathrm{f}\left(x(t)\right),\\x(t)=Lz(t)+x_0.\end{cases} \tag{23}$$

In what follows, we will demonstrate the existence and uniqueness of the solution to DDICA (23). Firstly, the definition of its strong global solution needs to be given.

Definition 1: We call $z(t) : [t_0, +\infty) \to R^n$ is a strong global solution of DDICA (23) if it satisfies:

(i) : $z(t), \dot{z}(t) : [t_0, +\infty) \to R^n$ are locally absolutely continuous;

(ii) : $\ddot{z}(t) + \frac{\alpha}{t}\dot{z}(t) + L\nabla \mathrm{f}(Lz + x_0) = 0$ for almost every $t \geq t_0$;

(iii) : $z(t_0) = u_0$ and $\dot{z}(t_0) = v_0$.

In order to prove the existence and uniqueness of the trajectories of DDICA (23), the following Lemma is needed.

Lemma 3: When $\mathrm{f}(x)$ is a convex and l-smooth function, i.e., it satisfies Assumption 2, then for any $u_0, v_0 \in R^n$, there exists a unique strong global solution of DDICA (23).

Proof: The DDICA (23) can be equivalently reformulated into a first order dynamical system:

$$\begin{cases} \dot{Y}(t) = F\left(t, Y(t)\right), \\ Y\left(t_0\right) = \left(u_0, v_0\right), \end{cases} \tag{24}$$

where $Y(t) = (y(t), \dot{y}(t))$, $F : [t_0, +\infty) \times R^n \times R^n \to R^n \times R^n$, $F(t, u, v) = (v, -\frac{\alpha}{t}v - L\nabla \mathrm{f}(Lu + x_0))$.

Applying the Cauchy-Lipschitz-Picard Theorem to the first-order dynamics in (24) with the following (i) and (ii) conditions to guarantee the existence and uniqueness of strong global solution to (24).

i) For every $t \in [t_0, +\infty)$, the mapping $F(t, \cdot, \cdot)$ is $\mathfrak{l}(t)$-Lipschitz continuous and $\mathfrak{l}(t) \in \mathcal{L}^1_{\mathrm{loc}}$.

Proof: Let $t \in [t_0, +\infty)$ be fixed and (u, v), $(\bar{u}, \bar{v}) \in R^n \times R^n$. Taking advantage of the Lipschitz continuity of f, and DDICA (23), we have

$$\begin{aligned} &\|F(t, u, v) - F\left(t, \bar{u}, \bar{v}\right)\| \\ &= \left(\|v - \bar{v}\|^2 + \left\|\frac{\alpha}{t}\left(\bar{v} - v\right)\right.\right. \\ &\quad \left.\left. + L\left(\nabla \mathrm{f}\left(L\bar{u} + x_0\right) - \nabla \mathrm{f}\left(Lu + x_0\right)\right)\right\|^2\right)^{\frac{1}{2}} \\ &\leqslant \left(\left(1 + \frac{2\alpha^2}{t^2}\right)\|v - \bar{v}\|^2 + 2\left(l\varpi_{\max}(L)^2\right)^2\|u - \bar{u}\|^2\right)^{\frac{1}{2}} \\ &\leqslant \left(1 + \frac{\sqrt{2}\alpha}{t} + \sqrt{2}l\varpi_{\max}(L)^2\right)\|(u, v) - (\bar{u}, \bar{v})\|. \end{aligned}$$

□

Letting $\mathfrak{l}(t) = 1 + \frac{\sqrt{2}\alpha}{t} + \sqrt{2}l\varpi_{\max}(L)^2$, we have $\|F(t, u, v) - F(t, \bar{u}, \bar{v})\| \leq \mathfrak{l}(t)\|(u, v) - (\bar{u}, \bar{v})\|$. Obviously the function $\mathfrak{l}(t)$ is integrable on $[t_0, +\infty)$, hence $\mathfrak{l}(t) \in \mathcal{L}^1_{\mathrm{loc}}$.

ii) For all u, $v \in R^n$, one has $F(\cdot, u, v) \in \mathcal{L}^1([t_0, +\infty), R^n \times R^n)$.

Proof:

$$\begin{aligned} &\int_{t_0}^{\mathrm{T}} \|F(t, u, v)\| dt \\ &= \int_{t_0}^{\mathrm{T}} \left(\|v\|^2 + \left\|\frac{\alpha}{t}v + L\nabla \mathrm{f}\left(Lu + x_0\right)\right\|^2\right)^{\frac{1}{2}} dt \\ &\leq \int_{t_0}^{\mathrm{T}} \left(\left(1 + \tfrac{2\alpha^2}{t^2}\right)\|v\|^2 + 2\varpi_{\max}(L)^2\|\nabla \mathrm{f}\left(Lu + x_0\right)\|^2\right)^{\frac{1}{2}} dt \\ &\leq \int_{t_0}^{\mathrm{T}} \left(1 + \tfrac{\sqrt{2}\alpha}{t} + \sqrt{2}\varpi_{\max}(L)\right) dt \sqrt{\|v\|^2 + \|\nabla \mathrm{f}\left(Lu + x_0\right)\|^2}, \end{aligned}$$

and the conclusion to (ii) follows by the continuity of the function $t \to (1 + \frac{\sqrt{2}\alpha}{t} + \sqrt{2}\varpi_{\max}(L))$ on $[t_0, \mathrm{T}]$. □

In view of the statements (i) and (ii), the existence and uniqueness of a strong global solution to first-order dynamical system (24) are hold which comes from the Cauchy-Lipschitz-Picard Theorem. Therefore, the same conclusions are also applicable to the DDICA (23). □

Theorem 1: Suppose Assumptions 1 and 2, $\alpha \geqslant 3$, and $\sum_{i=1}^m x_{0,i} = \sum_{i=1}^m d_i$ hold, DDICA (23) has a convergence rate at $O(\frac{1}{t^2})$, i.e.,

$$\begin{aligned} &\mathrm{f}\left(Lz(t) + x_0\right) - \mathrm{f}\left(Lz^* + x_0\right) \\ &= \mathrm{f}\left(x(t)\right) - \mathrm{f}\left(x^*\right) = O\left(\frac{1}{t^2}\right). \end{aligned} \tag{25}$$

Proof: See the Appendix A, which can be found on the Computer Society Digital Library at https://doi.org/10.1109/TNSE.2023.3248267. □

B. Distributed Damped Inertial Discrete Time Algorithm (DDIDA)

In the following, we derive a 'rate-matching' distributed discrete time algorithm of DDICA (23), named as distributed damped inertial discrete time algorithm (DDIDA) with a convergence rate of $O(\frac{1}{k^2})$ when selecting a fixed positive step-size.

Introducing an auxiliary function $\gamma^k = \frac{k+\alpha-1}{k(l\varpi_{\max}(L)^2)}$, $\lim_{k\to+\infty} \gamma^k = \frac{k+\alpha-1}{k(l\varpi_{\max}(L)^2)} = 1$ in f. Then, applying explicit Euler discretization to f with an auxiliary variable ξ, and semi-implicit Euler discretization with respect to the $\frac{\alpha}{t}$ with step-size is 1 to (23), we obtain

$$\begin{aligned} &z^{k+1} - 2z^k + z^{k-1} + \frac{\alpha - 1}{k}\left(z^{k+1} - z^k\right) \\ &+ \frac{1}{k}\left(z^k - z^{k-1}\right) + \gamma^k L\nabla \mathrm{f}\left(L\xi^k + x_0\right) = 0, \end{aligned} \tag{26}$$

and setting $\xi^k = y^k$ by using the Nesterov's discretization scheme in Fig. 2, we obtain the DDIDA as follows:

$$\begin{cases} y^k = z^k + \frac{k-1}{k+\alpha-1}\left(z^k - z^{k-1}\right), \\ z^{k+1} = y^k - \frac{1}{l\varpi_{\max}(L)^2}L\nabla \mathrm{f}\left(Ly^k + x_0\right), \\ x^k = Lz^k + x_0, \end{cases} \tag{27}$$

Theorem 2: Under Assumptions 1–2, $\sum_{i=1}^m x_{0,i} = \sum_{i=1}^m d_i$ and $\alpha \geqslant 3$, the DDIDA (27) has a convergence rate $O(\frac{1}{k^2})$, i.e.,

$$\begin{aligned} &\mathrm{f}\left(Lz^k + x_0\right) - \mathrm{f}\left(Lz^* + x_0\right) \\ &= \mathrm{f}\left(x^k\right) - \mathrm{f}\left(x^*\right) = O\left(\frac{1}{k^2}\right). \end{aligned}$$

Proof: Please see the Appendix B, available in the online supplemental material. □

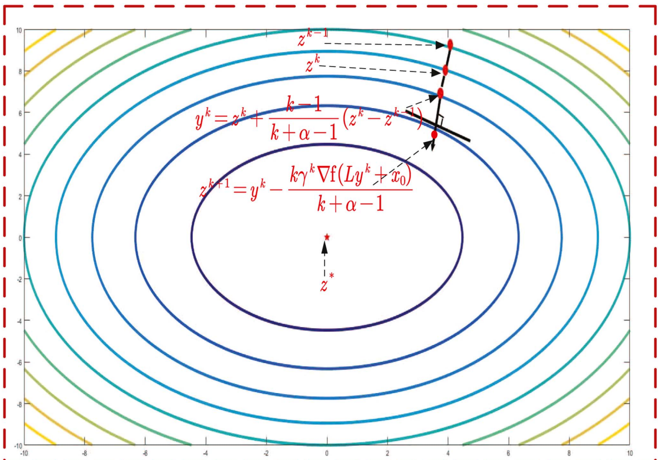

Fig. 2. Nesterov's discretization scheme.

IV. Distributed Fixed Inertial Algorithms for RAP With Strongly Convex Function

For strongly convex optimization problems, the centralized fixed inertial algorithms [41], [42], [43], [44] have faster exponential and linear convergence rates. In this section, we want to design distributed fixed inertial algorithms for RAP with a strongly convex function, which have faster exponential and linear convergence rates. It is worth noting that $f(z)$ does not satisfy the strong convexity property, even if the original $\mathrm{f}(x)$ is strongly convex, so it is necessary to introduce some new techniques to design fixed inertial distributed algorithms to obtain faster exponential convergence rate (or linear convergence rate of distributed discrete time algorithms).

A. Distributed Fixed Inertial Continuous Time Algorithm (DFICA)

Note that $f(\mathrm{z})$, i.e., $\mathrm{f}(L\mathrm{z}+\mathrm{x}_0)$ is not a strongly convex function even if f(x) is strongly convex, but $f(z)$ satisfies a quadratic growth (QG) condition from Lemma 2. Inspired by the centralized fixed inertial algorithms [41], [42], [43], [44], we propose the following distributed fixed inertial continuous time algorithm (DFICA) of agent i:

$$\begin{cases} \ddot{\mathrm{z}}_i(t) + \theta\dot{\mathrm{z}}_i(t) \\ = \sum_{j\in\mathcal{N}_i}\left(\nabla \mathrm{f}_i\left(\mathrm{x}_i(t)\right) - \nabla \mathrm{f}_j\left(\mathrm{x}_j(t)\right)\right), \\ \mathrm{x}_i(t) = \sum_{j\in\mathcal{N}_i} \mathrm{z}_i(t) + \mathrm{x}_{i,0}, \end{cases} \tag{28}$$

where $\theta = (2 - \frac{\sqrt{2}}{2})\sqrt{\mu}\lambda_2$, and $\mu = \max_{1\leqslant i\leqslant n}\{\mu_i\}$ in Assumption 3. The compact form of DFICA (28) yields:

$$\begin{cases} \ddot{\mathrm{z}}(t) + \theta\dot{\mathrm{z}}(t) = L\nabla \mathrm{f}\left(\mathrm{x}(t)\right), \\ \mathrm{x}(t) = L\mathrm{z}(t) + \mathrm{x}_0, \end{cases} \tag{29}$$

Before proving the existence and uniqueness of the strong global solution of DFICA (29), it is needed to provide the definition of a strong global solution similar to *Definition 1*, which is

Definition 2: A solution $\mathrm{z}(t): [t_0, +\infty) \to R^n$ is called a strong global solution to DFICA (29) if the following conditions hold.

(i) : $\mathrm{z}(t), \dot{\mathrm{z}}(t) : [t_0, +\infty) \to R^n$ are locally absolutely continuous, that is, they are absolutely continuous in every interval $[0, q]$, with $0 < q < +\infty$;

(ii) : $\ddot{\mathrm{z}}(t) + \theta\dot{\mathrm{z}}(t) + L\nabla\mathrm{f}(L\mathrm{z} + \mathrm{x}_0) = 0$, for almost every $t \geq t_0 > 0$;

(iii) : $\mathrm{z}(t_0) = \mathrm{u}_0$ and $\dot{\mathrm{z}}(t_0) = \mathrm{r}_0$.

Lemma 4: If $\mathrm{f}(x)$ satisfies Assumption 3, i.e, $\mathrm{f}(x)$ is a μ-strongly convex and l-smooth function, then, there exists a unique strong global solution of DFICA (29) for any $\mathrm{u}_0, \mathrm{r}_0 \in R^n$.

Proof: The DDICA (23) can be equivalently reformulated into a first order dynamical system:

$$\begin{cases} \dot{\mathrm{Y}}(t) = \mathrm{F}\left(t, \mathrm{Y}(t)\right), \\ \mathrm{Y}\left(t_0\right) = \left(\mathrm{u}_0, \mathrm{r}_0\right), \end{cases} \tag{30}$$

where $\mathrm{Y}(t) = (\mathrm{y}(t), \dot{\mathrm{y}}(t))$, $\mathrm{F}: [t_0, +\infty) \times R^n \times R^n \to R^n \times R^n$, $\mathrm{F}(t, \mathrm{u}, \mathrm{r}) = (\mathrm{r}, -\theta\mathrm{r} - L\nabla\mathrm{f}(L\mathrm{u} + \mathrm{x}_0))$.

According to Cauchy-Lipschitz-Picard Theorem, the existence and uniqueness of strong global solution to (30) can be guaranteed if the following (i) and (ii) conditions hold.

i) For every $t \in [t_0, +\infty)$, the mapping $F(t, \cdot, \cdot)$ is ι -Lipschitz continuous and $\iota \in \mathcal{L}^1_{\mathrm{loc}}$.

Proof: Let $t \in [t_0, +\infty)$ be fixed and $(u, v), (\bar{u}, \bar{v}) \in R^n \times R^n$. Taking advantage of the Lipschitz continuity of f, and DDICA (23), we have

$$\begin{aligned} &\left\|\mathrm{F}\left(t, \mathrm{u}, \mathrm{r}\right) - \mathrm{F}\left(t, \bar{\mathrm{u}}, \bar{\mathrm{r}}\right)\right\| \\ &= \Big(\left\|\mathrm{r} - \bar{\mathrm{r}}\right\|^2 + \left\|\theta\left(\mathrm{r} - \bar{\mathrm{r}}\right)\right. \\ &\quad \left. + L\left(\nabla\mathrm{f}\left(L\mathrm{u} + \mathrm{x}_0\right) - \nabla\mathrm{f}\left(L\bar{\mathrm{u}} + \mathrm{x}_0\right)\right)\right\|^2\Big)^{\frac{1}{2}} \\ &\leqslant \left(\left(1 + 2\theta^2\right)\left\|\mathrm{r} - \bar{\mathrm{r}}\right\|^2 + 2\left(l\varpi_{\max}(L)^2\right)^2\left\|\mathrm{u} - \bar{\mathrm{u}}\right\|^2\right)^{\frac{1}{2}} \\ &\leqslant \left(1 + \sqrt{2}\theta + \sqrt{2}l\varpi_{\max}(L)^2\right)\left\|(\mathrm{u}, \mathrm{r}) - (\bar{\mathrm{u}}, \bar{\mathrm{r}})\right\|. \end{aligned}$$

□

Let $\iota = 1 + \sqrt{2}\theta + \sqrt{2}l\varpi_{\max}(L)^2$, we have $\|\mathrm{F}(t, \mathrm{u}, \mathrm{r}) - \mathrm{F}(t, \bar{\mathrm{u}}, \bar{\mathrm{r}})\| \leq \iota\|(\mathrm{u}, \mathrm{r}) - (\bar{\mathrm{u}}, \bar{\mathrm{r}})\|$. Since $\iota : (t_0, +\infty) \to R$ is integrable and is continuous on $[t_0, +\infty)$, hence $\iota \in \mathcal{L}^1_{\mathrm{loc}}$.

ii) For all $\mathrm{u}, \mathrm{r} \in R^n$, one has $\mathrm{F}(\cdot, \mathrm{u}, \mathrm{r}) \in \mathcal{L}^1([t_0, +\infty), R^n \times R^n)$.

Proof:

$$\begin{aligned} &\int_{t_0}^{\mathrm{T}} \left\|\mathrm{F}\left(t, \mathrm{u}, \mathrm{r}\right)\right\| dt \\ &= \int_{t_0}^{\mathrm{T}} \left(\left\|\mathrm{r}\right\|^2 + \left\|\theta\mathrm{r} + L\nabla\mathrm{f}\left(L\mathrm{u} + \mathrm{x}_0\right)\right\|^2\right)^{\frac{1}{2}} dt \\ &\leq \int_{t_0}^{\mathrm{T}} \left(\left(1 + 2\theta^2\right)\left\|\mathrm{r}\right\|^2 + 2\varpi_{\max}(L)^2\left\|\nabla\mathrm{f}\left(L\mathrm{u} + \mathrm{x}_0\right)\right\|^2\right)^{\frac{1}{2}} dt \\ &\leq \int_{t_0}^{\mathrm{T}} \left(1 + \sqrt{2}\theta + \sqrt{2}\varpi_{\max}(L)\right) dt \sqrt{\left\|\mathrm{r}\right\|^2 + \left\|\nabla\mathrm{f}\left(L\mathrm{u} + \mathrm{x}_0\right)\right\|^2}, \end{aligned}$$

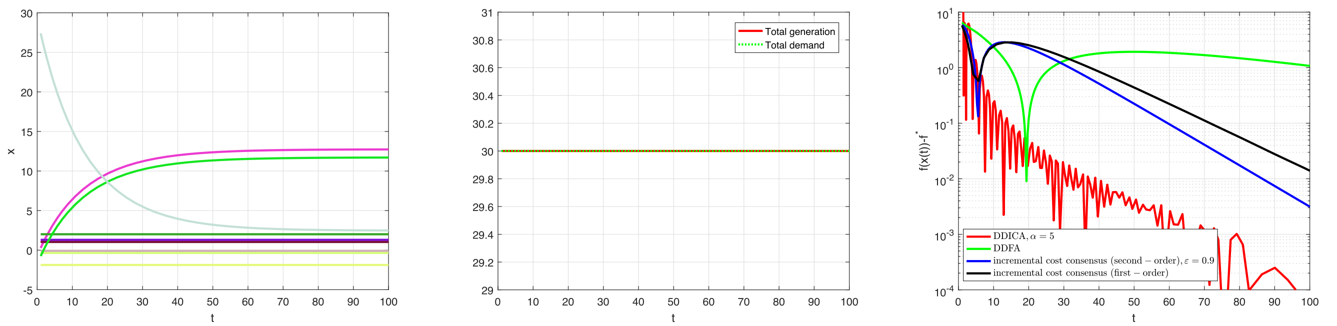

Fig. 3. Trajectories of state x of DDICA (23) (left). The trajectories of supply demand balance (middle). Comparison of convergence rates among DDICA (23), DDFA [23], incremental cost consensus (second-order) algorithm [26] and incremental cost consensus (first-order) algorithm [20] (right).

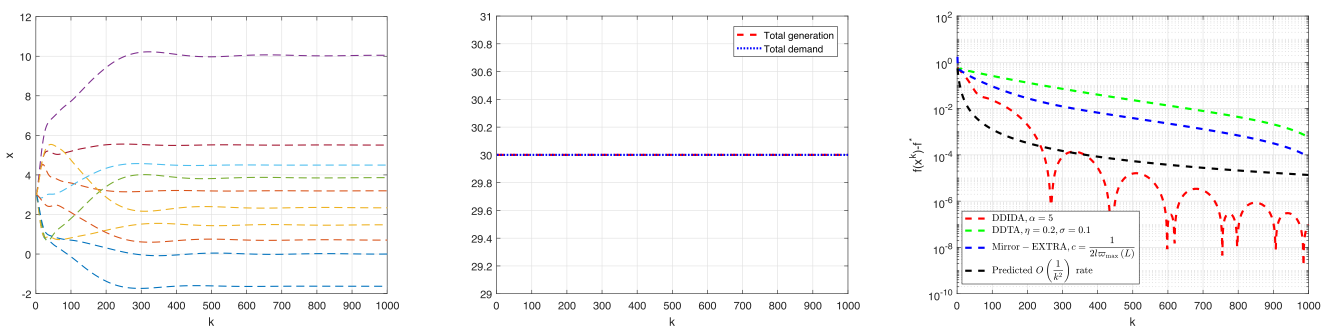

Fig. 4. Trajectories of state x of DDIDA (27) (left). The trajectories of supply demand balance (middle). Comparison of convergence rates among DDIDA (27), DDTA [15], Mirror-EXTRA [13], Predicted $O(\frac{1}{k^2})$ rate (right).

and the conclusion to (ii) follows by the continuity of the function $t \to (1+\sqrt{2}\theta+\sqrt{2}\varpi_{\max}(L))$. □

In view of the statements of (i) and (ii), the existence and uniqueness of strong global solutions of the first-order dynamical system (30) holds, which comes from the Cauchy-Lipschitz-Picard Theorem. Therefore, the same conclusion is also applied to DFICA (29). □

Theorem 3: Under Assumption 1 and 3 and $\sum_{i=1}^n x_{0,i} = \sum_{i=1}^n d_i$, the proposed DFICA (29) has a convergence rate of $O(e^{-(2-\sqrt{2})\sqrt{\mu}\lambda_2 t})$, i.e.,

$$\begin{aligned} f\left(\mathrm{z}(t)\right)-f\left(\mathrm{z}^*\right) &= \mathrm{f}\left(\mathrm{x}(t)\right)-\mathrm{f}\left(\mathrm{x}^*\right)=O\left(e^{-\left(2-\sqrt{2}\right)\sqrt{\mu}\lambda_2 t}\right), \\ \|\mathrm{z}(t)-\mathrm{z}^*\| &= O\left(e^{-\left(1-\frac{\sqrt{2}}{2}\right)\sqrt{\mu}\lambda_2 t}\right). \end{aligned} \tag{31}$$

Proof: The proof is given in the Appendix C, available in the online supplemental material. □

B. Distributed Fixed Inertial Discrete Time Algorithm (DFIDA)

Note that (29) can be rewritten as:

$$\dot{\mathrm{z}}(t)=\mathrm{v}(t), \tag{32a}$$

$$\dot{\mathrm{v}}(t)=-\theta\mathrm{v}(t)-L\nabla\mathrm{f}\left(\mathrm{x}(t)\right), \tag{32b}$$

$$\mathrm{x}(t)=L\mathrm{z}(t)+\mathrm{x}_0. \tag{32c}$$

According to the discretization schemes in works [43], [44], i.e., applying explicit discretization scheme to variable z and semi-implicit Euler discretization scheme to v and following by a sufficient decrease update, one has

1) For (32a), let $\frac{\mathrm{y}^k-\mathrm{z}^k}{\mathrm{s}}=\mathrm{v}^k$, $\mathrm{z}^{k+1}=\mathrm{y}^k-\mathrm{s}^2L\nabla\mathrm{f}(L\mathrm{y}^k+x_0)$ on account of explicit Euler discretization scheme and a sufficient decrease update step with step-size $\mathrm{s}=\frac{1}{\sqrt{l\varpi_{\max}(L)}}$.
2) For (32b), using semi-implicit discretization scheme with $\theta=(2-\frac{\sqrt{2}}{2})\sqrt{\mu}\lambda_2$, $\mathrm{s}=\frac{1}{\sqrt{l\varpi_{\max}(L)}}$, $a=\sqrt{\mu}\lambda_2$ yields
$$\frac{\mathrm{v}^{k+1}-\mathrm{v}^k}{s}=-\frac{\theta\mathrm{v}^k}{1+\theta\mathrm{s}}-\frac{1-\theta a s^2}{(1+\theta\mathrm{s})(1+as)}L\nabla\mathrm{f}\left(L\mathrm{y}^k+x_0\right)$$
$$\Rightarrow \mathrm{v}^{k+1}=\frac{\mathrm{v}^k-\mathrm{s}L\nabla\mathrm{f}\left(L\mathrm{y}^k+x_0\right)}{1+\theta\mathrm{s}}+\frac{as^2}{(1+as)}L\nabla\mathrm{f}\left(L\mathrm{y}^k+x_0\right).$$
3) For (32c), one has $\mathrm{x}^{k+1}=L\mathrm{z}^{k+1}+\mathrm{x}_0$ by the explicit discretization scheme.
To sum up, we get a distributed fixed inertial discrete time algorithm (DFIDA) as follows:
$$\begin{cases} \mathrm{y}^k=\mathrm{z}^k+\mathrm{s}\mathrm{v}^k, \\ \mathrm{w}^k=(1+\theta\mathrm{s})^{-1}\left(\mathrm{v}^k-\mathrm{s}L\nabla\mathrm{f}\left(L\mathrm{y}^k+x_0\right)\right), \\ \mathrm{z}^{k+1}=\mathrm{y}^k-\mathrm{s}^2L\nabla\mathrm{f}\left(L\mathrm{y}^k+x_0\right), \\ \mathrm{v}^{k+1}=\mathrm{w}^k+(1+as)^{-1}as^2L\nabla\mathrm{f}\left(L\mathrm{y}^k+x_0\right), \\ \mathrm{x}^{k+1}=Lz^{k+1}+\mathrm{x}_0, \end{cases} \tag{33}$$
where $\theta=(2-\frac{\sqrt{2}}{2})\sqrt{\mu}\lambda_2$, $\mathrm{s}=\frac{1}{\sqrt{l\varpi_{\max}(L)}}$, $a=\sqrt{\mu}\lambda_2$.

Theorem 4: Suppose Assumptions 1, 3 and $\sum_{i=1}^{m} x_{0,i} = \sum_{i=1}^{m} d_i$ hold, the proposed DFIDA (33) has the following a faster linear convergence rate

$$\begin{aligned} & f\left(\mathrm{z}^{k+1}\right) - f\left(\mathrm{z}^*\right) = \mathrm{f}\left(\mathrm{x}^{k+1}\right) - \mathrm{f}\left(\mathrm{x}^*\right) \\ & = O\left(\left(1 - \frac{(2-\sqrt{2})\sqrt{\mu}\lambda_2}{\sqrt{l}\varpi_{\max}(L)} + o\left(\frac{\sqrt{\mu}\lambda_2}{\sqrt{l}\varpi_{\max}(L)}\right)\right)^{k+1}\right), \\ & \|\mathrm{z}^{k+1} - \mathrm{z}^*\| \\ & = O\left(\left(1 - \frac{(2-\sqrt{2})\sqrt{\mu}\lambda_2}{\sqrt{l}\varpi_{\max}(L)} + o\left(\frac{\sqrt{\mu}\lambda_2}{\sqrt{l}\varpi_{\max}(L)}\right)\right)^{\frac{k+1}{2}}\right). \end{aligned} \tag{34}$$

Proof: The proof is given in Appendix D, available in the online supplemental material. □

Remark 1: It should be noted that the parameters s, θ and s of DDIDA (27), DFICA (29) and DFIDA (33) are related to the Fiedler eigenvalue λ_2, Lipschitz constant l, strongly convex parameter μ and $\varpi_{\max}(L)$. For λ_2 and $\varpi_{\max}(L)$, some valid distributed results can be utilized. For instance, distributed computational approach in [50]. Moreover, l, μ can be obtained in the following distributed manner, i.e., each agent i communicates with its neighbor to get l by using $\min\{\cdot\}$ operation of l_i, l_j, $j \in \mathcal{N}_i$, and to obtain μ by applying $\max\{\cdot\}$ operation of μ_i, μ_j, $j \in \mathcal{N}_i$.

Remark 2: For the supply-demand constraint (1b), there are two ways to achieve it in a distributed manner. First, leader election is one method. That is, the 1 token is passed in the network. If any agent i gets a token and also knows d, then it chooses $x_i = \sum_{i=1}^{n} d_i$ and stops passing the token. For those who do not get a token, they choose $x_i = 0$, $j \neq i$. Second, the 1 token is passed into the network. If any agent i gets a token and also knows $\sum_{i=1}^{n} d_i$, then they choose $0 \leqslant x_i \leqslant \sum_{i=1}^{n} d_i$ and transfer the token and the remaining resource $d - x_i$ to the next agent. In addition, if agent i has two or more outer neighbors, an outer neighbor agent is randomly selected to transfer the above information. Thus, the initial network-wide supply and demand constrained distribution is completed.

V. Numerical Simulations

Example 1: DDICA (23) and DDIDA (27) for log-sum-exp functions

In this example, consider a resource dispatch problem:

$$\min \sum_{i=1}^{n} \mathrm{f}_i(x_i); \text{s.t.} \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} d_i. \tag{35}$$

Its local function is a log-sum-exp function $\mathrm{f}_i(x_i) = \mathfrak{a} \log\left[\sum_{i=1}^{m} \exp(\frac{\gamma_{i,j}^T x_i - \beta_{i,j}}{\mathfrak{a}})\right]$, where $n = 10, m = 50, \mathfrak{a} = 20$, $d = 30$, $x_i, \gamma_{i,j}, \beta_{i,j} \in R$ are random scalars generated from a uniform distribution on the interval $[0, 1]$. In addition, its communication graph of 10 agents is presented in Fig. 5. Simulation results of DDICA (23) and DDIDA (27) are shown in Figs. 3 and 4. In Fig. 3, it displays that the variables x are globally asymptotically stable in (left) and the supply demand balance is satisfied in (middle). In addition, the results in Fig. 3 (right) shows that the DDICA (23) has a much faster performance than DDFA [23], incremental cost consensus (second-order) [26] and incremental cost consensus (first-order) algorithm [20]. In Fig. 4 (left) and (middle), the global convergence property and supply demand balance constraint of DDIDA (27) can be obtained. In addition, the results in Fig. 4 (right) shows that

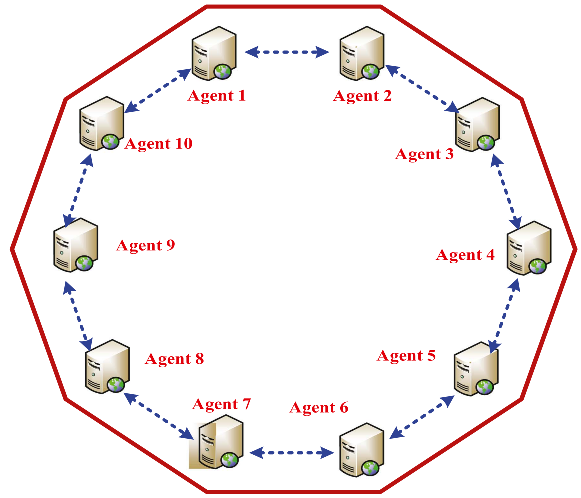

Fig. 5. The communication graph of 10 agents.

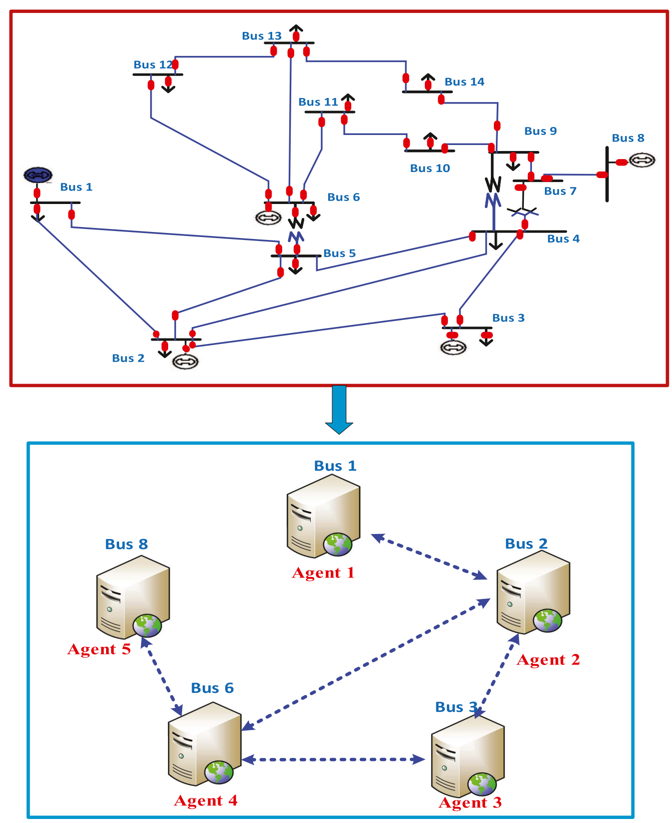

Fig. 6. IEEE-14 bus system and its communication topology.

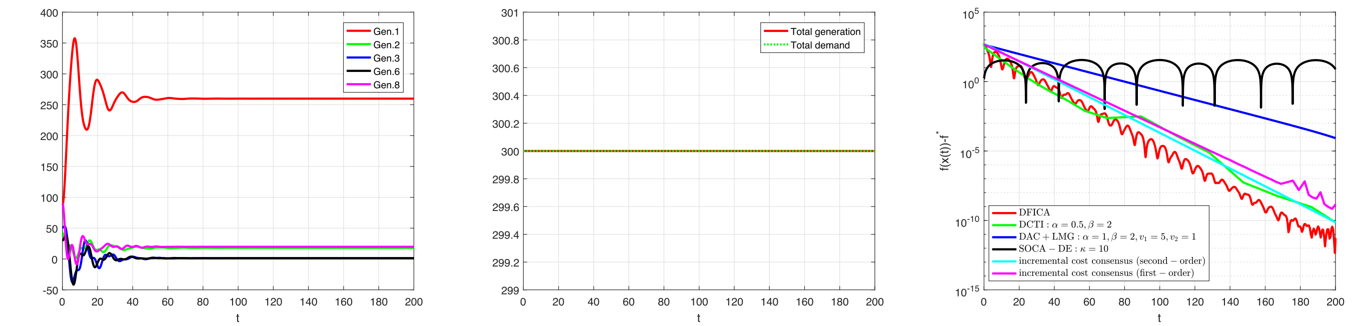

Fig. 7. Trajectories of state x of DFICA (29) (left). The trajectories of supply demand balance (middle). Comparison of convergence rates among DFICA (29), DCTI [39], DAC+LMG [46], SOCA-DE [28], incremental cost consensus (second-order) algorithm [26] and incremental cost consensus (first-order) algorithm [20] (right).

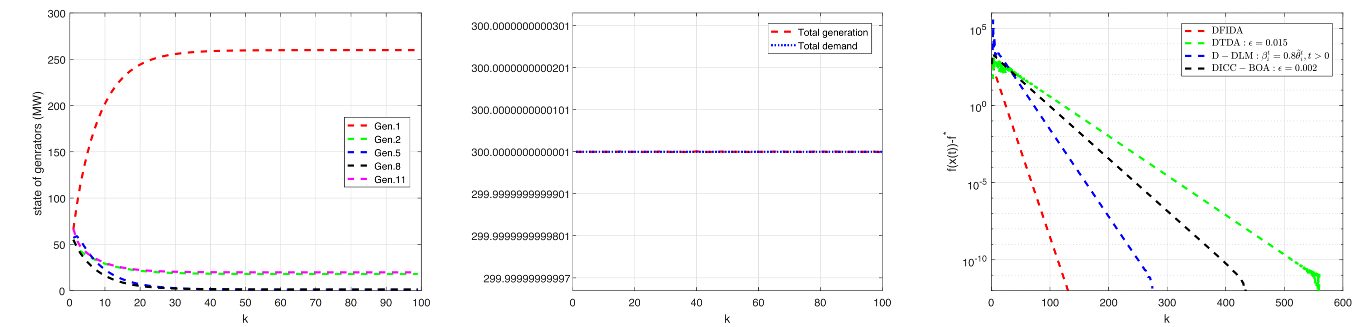

Fig. 8. Trajectories of state x of DFIDA (33) (left). The trajectories of supply demand balance (middle). Comparison of convergence rates among DFIDA (33), D-DLM [14], DTDA [39] and DICC-BOA [47] (right).

TABLE II
COST FUNCTION PARAMETERS

Bus	$a_{i,1}$ (\$/MW^{2}h)	$a_{i,2}$ (\$/MWh)
1	0.04	2.0
2	0.03	3.0
3	0.035	4.0
6	0.03	4.0
8	0.04	2.5

the DDIDA (27) has a faster convergence rate than DDTA [15], Mirror-EXTRA [13] and **Predicted** $O(\frac{1}{k^2})$ **rate.**

Example 2: DFICA (29) and DFIDA (33) for Economic Dispatch in Smart Grid

In this example, a numerical simulation on smart grid is proposed to illustrate the effectiveness and superiority of DFICA (29) and DFIDA (33). The convergence rate comparisons are also considered. The IEEE-14 bus system and its communication topology are displayed in Fig. 6. The cost function parameters of generators are given in Table II [39], in which the cost function is $\mathrm{f}_i(x_i) = a_{i,1}x_i^2 + a_{i,2}x_i$, i.e., it is a strongly convex function.

First, the continuous time case, i.e., DFICA (29) is considered. Let $P_D = 300\mathbf{MW}$ and $x_0 = [66, 55, 57, 55, 67]^T$. The trajectories of state x, supply-demand constraint, and error of $f(x(t)) - f^*$ are shown as in Fig. 7. As can be seen from the Fig. 7 (left) that the DFICA (29) is globally asymptotically stable and converges to the optimal solution $x^* = [259.85, 17.98, 1.1254, 1.313, 19.735]^T$. The supply demand balance always holds from Fig. 7 (middle). In addition, the convergence performance comparison of DFICA (29) is carried out in Fig. 7 (right), where the error $\log_{10}(\mathrm{f}(x(t)) - f^*)$ is applied as the comparison indicator. As is shown in Fig. 7 (right), the DFICA (29) has an exponential convergence rate which is consistent with the conclusion of Theorem 5, which performs slightly better than DCTI [39], DAC+LMG [46], SOCA-DE [28], incremental cost consensus (second-order) algorithm [26] and incremental cost consensus (first-order) algorithm [20].

Next, the discrete time case of DFICA (29), i.e., the DFIDA (33) is also considered. Similar to the above continuous time case. Letting $P_D = 300\mathbf{MW}$ and $x_0 = [66, 55, 57, 55, 67]^T$. The results of DFIDA (33) are shown in Fig. 8. As can be seen from the Fig. 8 (left) and (middle) that the trajectories of $Gens$ are globally asymptotically stable and the supply demand equality constraint always holds when using the initial values $x_0 = [66, 55, 57, 55, 67]^T$. Fig. 8 (right) shows the result of the comparative test on convergence rates of DFIA (8) and other distributed algorithms with error measure $\mathrm{f}(x^k) - f^*$. Fig. 8 (right) shows that DFIDA (33) has a linear convergence rate, and it improves the convergence rate in comparison with the distributed algorithms with inertial term D-DLM [14], DTDA [39] and DICC-BOA [47] without inertial accelerated item.

VI. Conclusion

In this article, we have proposed two accelerated distributed (damped and fixed) inertial continuous time and distributed discrete time algorithms for solving RAP under an undirected graph. For solving RAP with a convex function, a distributed damped continuous time algorithm with convergence rate $O(\frac{1}{t^2})$ has been proposed. Then, by the special Nesterov's discretization scheme, we have proposed a novel rate-matching distributed damped discrete time algorithm with a fixed step-size, which has a convergence rate of $O(\frac{1}{t^2})$. To deal with RAP with a strongly convex objective function, we have found that the transformed distributed unconstrained problem (4) satisfies the **PL** and **QG** conditions although the original objective function is strongly convex. Further, we have proposed a distributed fixed inertial continuous time algorithm with an explicit and accelerated exponential convergence rate. Later, based on explicit Euler, semi-implicit and sufficient decrease update schemes, we have also proposed a rate-matching novel distributed fixed inertial discrete time algorithm with a fixed step-size, which has a linear convergence rate (like $O(q^k)$, $q \approx 1 - C\sqrt{\frac{\mu}{L}}$). Comparison experiments with existing state-of-the-art distributed continuous and discrete time algorithms further illustrate the effectiveness and superiority of our proposed distributed inertial algorithms. Considering that many resource allocation problems in practical applications are nonsmooth, we further study accelerated inertial algorithms to solve nonsmooth resource scheduling problems in future work; in addition, only the coupling equality constraint is considered in this article, we will further study resource allocation problems with general constraints and design the corresponding inertial algorithms.

References

[1] C. Li, X. Yu, W. Yu, T. Huang, and Z. Liu, "Distributed event-triggered scheme for economic dispatch in smart grids," *IEEE Trans. Ind. Informat.*, vol. 12, no. 5, pp. 1775–1785, Oct. 2016.

[2] H. Peng and X. Shen, "Deep reinforcement learning based resource management for multi-access edge computing in vehicular networks," *IEEE Trans. Netw. Sci Eng.*, vol. 7, no. 4, pp. 2416–2428, Oct.–Dec. 2020.

[3] K. Peng, H. Huang, B. Zhao, A. Jolfaei, X. Xu, and M. Bilal, "Intelligent computation offloading and resource allocation in IIoT with end-edge-cloud computing using NSGA-III," *IEEE Trans. Netw. Sci Eng.*, vol. 10, no. 5, pp. 3032–3046, Sep.-Oct. 2023.

[4] S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, pp. 1–122, 2011.

[5] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Found. Trends Netw*., vol. 1, pp. 1–144, 2006.

[6] A. Nedić, A. Olshevsky, and W. Shi, "Decentralized consensus optimization and resource allocation," in *Large-Scale and Distributed Optimization*. Cham, Switzerland: Springer, 2018. pp. 247–287.

[7] C. E. Lin, S. T. Chen, and C. L. Huang, "A direct Newton-Raphson economic dispatch," *IEEE Trans. Power Syst.*, vol. 7, no. 3, pp. 1149–1154, Aug. 1992.

[8] N. Amjady and H. Nasiri-Rad, "Economic dispatch using an efficient real-coded genetic algorithm," *IET Gener. Transmiss. Distrib.*, vol. 3, no. 3, pp. 266–278, 2009.

[9] S. Yang, S. Tan, and J. X. Xu, "Consensus based approach for economic dispatch problem in a smart grid," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4416–4426, Nov. 2013.

[10] I. Necoara, "Random coordinate descent algorithms for multi-agent convex optimization over networks," *IEEE Trans. Autom. Control.*, vol. 58, no. 8, pp. 2001–2012, Aug. 2013.

[11] T. T. Doan and A. Olshevsky, "Distributed resource allocation on dynamic networks in quadratic time," *Syst. Control Lett.*, vol. 99, pp. 57–63, 2017.

[12] N. S. Aybat and E. Y. Hamedani, "A distributed ADMM-like method for resource sharing under conic constraints over time-varying networks," *SIAM J. Optim.*, vol. 29, no. 4, pp. 3036–3068, 2019.

[13] A. Nedić, A. Olshevsky, and W. Shi, "Improved convergence rates for distributed resource allocation," in *Proc. IEEE Conf. Decis. Control*, 2018, pp. 172–177.

[14] Q. Lü, X. Liao, H. Li, and T. Huang, "Achieving acceleration for distributed economic dispatch in smart grids over directed networks," *IEEE Trans. Netw. Sci Eng.*, vol. 7, no. 3, pp. 1988–1999, Jul.–Sep. 2020.

[15] L. Liu and G. Yang, " Distributed optimal economic environmental dispatch for microgrids over time-varying directed communication graph," *IEEE Trans. Netw. Sci Eng.*, vol. 8, no. 2, pp. 1913–1924, Apr.–Jun. 2021.

[16] B. Wang, C. Zhang, C. Li, G. Yang, and Z. Dong, "Transactive energy sharing in a microgrid via an enhanced distributed adaptive robust optimization approach," *IEEE Trans. Smart Grid.*, vol. 13, no. 3, pp. 2279–2293, May 2022.

[17] X. Li, C. Li, G. Chen, and Z. Dong, "A risk-averse energy sharing market game for renewable energy microgrid aggregators," *IEEE Trans. Power Syst.*, vol. 37, no. 5, pp. 3528–3539, Sep. 2022.

[18] A. Cherukuri and J. Cortés, "Distributed generator coordination for initialization and anytime optimization in economic dispatch," *IEEE Trans. Control Netw. Syst.*, vol. 2, no. 3, pp. 226–237, Sep. 2015.

[19] A. Cherukuri and J. Cortes, "Initialization-free distributed coordination-for economic dispatch under varying loads and generator commitment," *Automatica*, vol. 74, pp. 183–193, Dec. 2016.

[20] W. Yu, C. Li, X. Yu, G. Wen, and J. Lü, "Economic power dispatch in smart grids: A framework for distributed optimization and consensus dynamics," *Sci. China Inf. Sci.*, vol. 61, no. 1, pp. 1–16, 2018.

[21] C. Li, X. Yu, T. Huang, and X. He, "Distributed optimal consensus over resource allocation network and its application to dynamical economic dispatch," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2407–2418, Jun. 2018.

[22] P. Yi, Y. Hong, and F. Liu, "Initialization-free distributed algorithms for optimal resource allocation with feasibility constraints and application to economic dispatch of power systems," *Automatica*, vol. 74, pp. 259–269, 2016.

[23] X. Zeng, P. Yi, Y. Hong, and L. Xie, "Distributed continuous-time algorithms for nonsmooth extended monotropic optimization problems," *SIAM. J. Control Optim.*, vol. 56, no. 6, pp. 3973–3993, 2018.

[24] L. Bai, M. Ye, C. Sun, and G. Hu, "Distributed economic dispatch control via saddle point dynamics and consensus algorithms," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 2, pp. 898–905, Mar. 2019.

[25] X. Wang, S. Yang, Z. Guo, M. Lian, and T. Huang, "A distributed dynamical system for optimal resource allocation over state-dependent networks," *IEEE Trans. Netw. Sci Eng.*, vol. 9, no. 4, pp. 2940–2951, Jul./Aug. 2022.

[26] X. He, D. W. C. Ho, T. Huang, J. Yu, H. Abu-Rub, and C. Li, "Second-order continuous-time algorithms for economic power dispatch in smart grids," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 9, pp. 1482–1492, Sep. 2018.

[27] Z. Deng, S. Liang, and W. Yu, "Distributed optimal resource allocation of second-order multiagent systems," *Int. J. Robust Nonlinear Control* vol. 28, no. 14, pp. 4246–4260, 2018.

[28] D. Wang, Z. Wang, C. Wen, and W. Wang, "Second-order continuous-time algorithm for optimal resource allocation in power systems," *IEEE Trans. Ind. Informat.*, vol. 15, no. 2, pp. 626–637, Feb. 2019.

[29] Y. Zhu, G. Wen, W. Yu, and X. Yu, "Continuous-time distributed proximal gradient algorithms for nonsmooth resource allocation over general digraphs," *IEEE Trans. Netw. Sci Eng.*, vol. 8, no. 2, pp. 1733–1744, Apr.–Jun. 2021.

[30] Z. Guo, M. Lian, S. Wen, and T. Huang, "An adaptive multi-agent system with duplex control laws for distributed resource allocation," *IEEE Trans. Netw. Sci Eng.*, vol. 9, no. 2, pp. 389–400, Mar./Apr. 2022.

[31] X. Zeng, J. Lei, and J. Chen, "Dynamical primal-dual accelerated method with applications to network optimization," *IEEE Trans. Autom. Control*, vol. 68, no. 3, pp. 1760–1767, Mar. 2023.

[32] W. Krichene, A. Bayen, and P. Bartlett, "Accelerated mirror descent in continuous and discrete time," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 2845–2853.

[33] W. Su, S. Boyd, and E. J. Candes, "A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 5312–5354, 2016.

[34] H. Attouch, Z. Chbani, J. Peypouquet, and P. Redont, "Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity," *Math Prog.*., vol. 168, no. 1, pp. 123–175, 2018.

[35] H. Attouch and A. Cabot, "Convergence of a relaxed inertial proximal algorithm for maximally monotone operators," *Math. Prog.*, vol. 184, no. 1, pp. 243–287, 2020.
[36] X. He, R. Hu, and Y. P. Fang, "Convergence rates of inertial primal-dual dynamical methods for separable convex optimization problems," *SIAM J. Control Optim.*, vol. 59, no. 5, pp. 3278–3301, 2021.
[37] H. Luo and L. Chen, "From differential equation solvers to accelerated first-order methods for convex optimization," *Math. Prog.*, vol. 195, pp. 735–781, 2022, doi: 10.1007/s10107-021-01713-3.
[38] R. Chen, T. Yang, and T. Chai, "Distributed accelerated optimization algorithms: Insights from an ODE," *Sci. China. Technol. Sci.*, vol. 63, no. 9, pp. 1647–1655, 2020.
[39] X. Shi, R. Zheng, Z. Lin, T. Yang, and G. Yan, "An exponentially convergent distributed algorithm for resource allocation problem," *Asian J. Control.*, vol. 23, no. 2, pp. 1072–1082, 2021.
[40] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, Berlin, Germany: Springer, 2003.
[41] B. T. Polyak, "Some methods of speeding up the convergence of iteration methods," *USSR Comput. Math. Math. Phys.*, vol. 4, no. 5, pp. 1–17, 1964.
[42] O. Sebbouh, C. Dossal, and A. Rondepierre, "Nesterov's acceleration and Polyak's heavy ball method in continuous time: Convergence rate analysis under geometric conditions and perturbations," Tech. Rep. 02173978, Jul. 2019.
[43] J. W. Siegel, "Accelerated first-order methods: Differential equations and Lyapunov functions," 2019, *arXiv:1903.05671*.
[44] J. F. Aujol, C. Dossal, and A. Rondepierre, "Convergence rates of the heavy-ball method with Lojasiewicz property," Ph.D. dissertation, IMB-Institut de Mathmatiques de Bordeaux; INSA Toulouse; UPS Toulouse, 2020.
[45] W. T. Lin, Y. W. Wang, C. Li, and X. Yu, "Predefined-time optimization for distributed resource allocation," *J. Franklin Inst.*, vol. 357, no. 16, pp. 11323–11348, 2020.
[46] A. Cherukuri and J. Cortés, "Initialization-free distributed coordination for economic dispatch under varying loads and generator commitment," *Automatica*, vol. 74, pp. 183–193, 2016.
[47] A. Wang and W. Liu, "Distributed incremental cost consensus-based optimization algorithms for economic dispatch in a microgrid," *IEEE Access*, vol. 8, pp. 12933–12941, 2020.
[48] Q. Liu, X. Le, and K. Li, "A distributed optimization algorithm based on multiagent network for economic dispatch with region partitioning," *IEEE Trans. Cybern.*, vol. 51, no. 5, pp. 2466–2475, May 2021.
[49] H. Karimi, J. Nutini, and M. Schmidt, "Linear convergence of gradient and proximal-gradient methods under the Polyak-Lojasiewicz condition," In *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discov. Databases*, 2016, pp. 795–811.
[50] A. Bertrand and M. Moonen, "Distributed computation of the Fiedler vector with application to topology inference in ad hoc networks," *Signal Process.*, vol. 93, no. 3, pp. 1106–1117, 2013.

You Zhao received the M.S. degree in signal and information processing from the School of Electronics and Information Engineering, Southwest University, Chongqing, China, in 2018. He is currently working toward the Ph.D. degree with the College of Computer Science, Chongqing University, Chongqing, China. His research interests include neurodynamic optimization, distributed optimization, compressed sensing, and smart grid.

Xiaofeng Liao (Fellow, IEEE) received the B.S. and M.S. degrees in mathematics from Sichuan University, Chengdu, China, in 1986 and 1992, respectively, and the Ph.D. degree in circuits and systems from the University of Electronic Science and Technology of China, Chengdu, China, in 1997. From 1999 to 2012, he was a Professor with Chongqing University, Chongqing, China. From January 2013 to November 2018, he was a Professor and the Dean with the College of Electronic and Information Engineering, Southwest University, Chongqing, China. He is currently a Professor and the Dean with the College of Computer Science, Chongqing University. He is also a Yangtze River Scholar with the Ministry of Education of China, Beijing, China. From November 1997 to April 1998, he was a Research Associate with the Chinese University of Hong Kong, Hong Kong. From October 1999 to October 2000, he was a Research Associate with the City University of Hong Kong, Hong Kong. From March 2001 to June 2001 and March 2002 to June 2002, he was a Senior Research Associate with the City University of Hong Kong. From March 2006 to April 2007, he was a Research Fellow at the City University of Hong Kong. He holds five patents, and has authored ore coauthored four books and more than 300 international journal and conference papers. His research interests include neural networks, nonlinear dynamical systems, bifurcation and chaos, and cryptography. Dr. Liao is also an Associate Editor for the IEEE TRANSACTIONS ON CYBERNETICS and IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.

Xing He (Member, IEEE) received the B.S. degree in mathematics and applied mathematics from the Department of Mathematics, Guizhou University, Guiyang, China, in 2009, and the Ph.D. degree in computer science and technology from Chongqing University, Chongqing, China, in 2013. From November 2012 to October 2013, he was a Research Assistant with the Texas A & M University at Qatar, Doha, Qatar. From December 2015 to February 2016, he was a Senior Research Associate with the City University of Hong Kong, Hong Kong. He is currently a Professor with the School of Electronics and Information Engineering, Southwest University, Chongqing, China. His research interests include neural networks, bifurcation theory, optimization method, smart grid, and nonlinear dynamical system.