Distributed Continuous-Time Algorithms for Optimal Resource Allocation With Time-Varying Quadratic Cost Functions

Bo Wang, Shan Sun, and Wei Ren, *Fellow, IEEE*

***Abstract*—In this article, we propose distributed continuous-time algorithms to solve the optimal resource allocation problem with certain time-varying quadratic cost functions for multiagent systems. The objective is to allocate a quantity of resources while optimizing the sum of all the local time-varying cost functions. Here, the optimal solutions are trajectories rather than some fixed points. We consider a large number of agents that are connected through a network, and our algorithms can be implemented using only local information. By making use of the prediction–correction method and the nonsmooth consensus idea, we first design two distributed algorithms to deal with the case when the time-varying cost functions have identical Hessians. We further propose an estimator-based algorithm which uses distributed average tracking theory to estimate certain global information. With the help of the estimated global information, the case of nonidentical constant Hessians is addressed. In each case, it is proved that the solutions of the proposed dynamical systems with certain initial conditions asymptotically converge to the optimal trajectories. We illustrate the effectiveness of the proposed distributed continuous-time optimal resource allocation algorithms through simulations.**

***Index Terms*—Distributed algorithms, optimization, resource allocation, time-varying cost functions.**

I. Introduction

RECENT years have witnessed a surge of interest in the distributed optimization problem [1]–[12]. Many practical situations have led to a strong demand for distributed optimal control in multiagent systems, where all the agents cooperatively fulfill a complex task in distributed ways. The optimal resource allocation or economic dispatch problem, which requires the agents to satisfy a common demand with their own constraints, is one of the fundamental aspects of distributed optimization and has been studied in power systems and other fields for many years using discrete-time algorithms [1]–[5]. A distributed algorithm for optimal economic dispatch using frequency control in discrete time is presented in [1] for an electric grid system. Considering transmission losses and generator constraints in a smart grid system, a consensus-based distributed running algorithm is proposed in [2] for the optimal economic power dispatch problem. Two cases are studied in [3] using distributed primal–dual subgradient algorithms incorporating the global constraint that is the intersection of all the local constraints. The Laplacian-gradient dynamics, where each agent is required to exchange the gradient information over the network, combined with nonsmooth exact penalty functions, is introduced in [4] to dynamically converge to the fixed optimal solution for the economic dispatch problem. To tackle the time-varying loads, the method in [4] is further developed in [5] based on dynamic average consensus to estimate the mismatch in load satisfaction. This method guarantees the convergence of the economic dispatch problem while the initial allocation is arbitrary.

Contrary to the above conventional discrete-time algorithms, some researchers have been devoted to designing distributed continuous-time algorithms for the optimal resource allocation problem in recent years [6]–[10]. The initialization-free distributed algorithms are applied in [6] to solve the optimal resource allocation problem with general local constraints. A distributed continuous-time algorithm based on average consensus theory is proposed in [7] for the optimal resource allocation problem, where all the agents need to satisfy multiple weighted demands. Compared with [5], all of the agents in [7] communicate the Lagrange multiplier information rather than the gradient with neighbors. By using nonsmooth penalty functions and differential inclusions, a distributed continuous-time method is presented in [8] for the resource allocation in power systems with second-order dynamics. This approach, which uses the exact approximate function and average consensus theory, is a combination of [5] and [7]. Distributed continuous-time subgradient algorithms are proposed in [9] to handle the nonsmooth optimization problem with general local constraints. The saddle point dynamics and consensus protocols are employed in [10] to address the optimal resource allocation problem in a power grid system with distributed thermal generators that have time-invariant quadratic cost functions.

All the above algorithms assume that the local cost functions or resources are time invariant. In many practical situations,

Manuscript received December 11, 2019; revised December 14, 2019, March 25, 2020, and June 24, 2020; accepted August 15, 2020. Date of publication September 1, 2020; date of current version December 16, 2020. This work was supported by the National Science Foundation under Grant ECCS-1920798. Recommended by Associate Editor P. Barooah. *(Corresponding author: Wei Ren.)*

Bo Wang is with the School of Automation, Beijing Institute of Technology, Beijing 100081, China (e-mail: 3120160440@bit.edu.cn).

Shan Sun and Wei Ren are with the Department of Electrical and Computer Engineering, University of California at Riverside, Riverside, CA 92521 USA (e-mail: ssun029@ucr.edu; ren@ee.ucr.edu).

Digital Object Identifier 10.1109/TCNS.2020.3020972

however, the cost functions and resources depend explicitly on time. Examples include the utility maximization problem for wireless sensor networks [11] and tracking of moving targets [13], [14], and the optimal resource allocation problem [15]. As a result, the optimal solution is no longer a fixed point but a time-varying trajectory that the agents need to not only collectively find but also track asymptotically. The decentralized discrete-time prediction–correction method is proposed in [11] to solve the constrained optimization problem with time-varying smooth strongly convex cost functions while involving bounded tracking errors. An extension of the above prediction–correction method in the dual space is shown in [12]. The aforementioned discrete-time algorithms deal with the time-varying resource allocation problem by sampling the time-varying cost functions at each discrete time instant, which results in bounded tracking errors that depend on the sampling period, the computation time at each step, and the number of iterations. Continuous-time methods are proposed in [13] and [15] to address the time-varying optimal resource allocation problem. The prediction–correction interior-point method in [13] is a centralized algorithm. The method in [15] is distributed and deals with time-varying resources, but the quadratic cost functions are required to be time invariant. In all the above-mentioned articles, the distributed continuous-time optimal resource allocation problem with time-varying cost functions is not addressed.

Motivated by the above discussion, this article proposes novel distributed continuous-time algorithms to address the optimal resource allocation problem with certain time-varying quadratic cost functions. The algorithms guarantee that the state of each agent will asymptotically track the time-varying optimal trajectory while satisfying the common demand. We first deal with the case of cost functions with identical Hessians by combining the prediction–correction method and the nonsmooth consensus idea. With the help of distributed estimators employed to estimate certain global information shared by all agents, we then deal with the case of cost functions with nonidentical constant Hessians. In each case, it is proved that the states of the proposed dynamical systems can globally asymptotically converge to the corresponding optimal trajectories. In addition, unlike some of the results in the existing literature, it is not necessary for each agent to obtain the gradient information of its neighbors in all the proposed algorithms.

Comparison with the literature: To address the optimal resource allocation problem, [1]–[5] propose distributed discrete-time algorithms for time-invariant cost functions, while [11] and [12] propose distributed discrete-time algorithms for time-varying cost functions but subject to bounded tracking errors. Also, [6]–[10] and [15] propose distributed continuous-time algorithms for time-invariant cost functions. In contrast, the current article proposes distributed continuous-time algorithms for time-varying cost functions with zero tracking errors. In addition, with our proposed algorithms, there is no need to exchange the gradient information among the agents as in [4] and [5] or estimate the Hessian inverse of the global cost function by truncating its Taylor expansion as in [11], which leads to increasing computation cost and decreasing convergence accuracy. While the current article gains some insight from [7], [8], and [10], the results therein are limited to time-invariant cost functions. In contrast, the current article deals with time-varying cost functions. The most relevant results to the current article are provided in [13] and [14]. While the equality constraint is taken into account in [13] for the time-varying optimization problem, the algorithms are centralized and each agent needs to obtain certain common global information. Distributed continuous-time algorithms including an estimator-based algorithm to handle nonidentical Hessians are studied in [14] for the time-varying unconstrained optimization problem, but the results cannot be directly applied to the time-varying optimal resource allocation problem in the presence of a common equality constraint.

II. Preliminaries

A. Notation and Graph Theory

Let $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{R}_{++}$ denote the set of real, non-negative, and positive numbers, respectively. Let $\mathbb{Z}_{++}$ denote the set of positive integers. Let $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$, where $n, m \in \mathbb{Z}_{++}$, denote the set of real-valued vectors with n entries and all $m \times n$ real matrices, respectively. Let $|\cdot|$ be the absolute value of a real number and $||\cdot||_1, ||\cdot||_2, ||\cdot||_\infty$ denote the 1-norm, 2-norm, and $\infty-$norm of a vector or a matrix, respectively. Let $\mathbf{1}_n$ denote the column vector with n ones, $\mathbf{0}_n$ denote the column vector with all entries being zero, and $\mathbf{I}_n$ denote the $n \times n$ identity matrix. For a number $x \in \mathbb{R}$, the standard sign function is denoted by $\text{sign}(x)$, i.e., $\text{sign}(x) = 1$ when $x > 0$, $\text{sign}(x) = -1$ when $x < 0$, and $\text{sign}(x) = 0$ otherwise. For a vector $x = [x_1, ..., x_n]^T \in \mathbb{R}^n$, let $\text{sign}(x) \triangleq [\text{sign}(x_1), ..., \text{sign}(x_n)]^T$. Let $\nabla_x f(x,t) \in \mathbb{R}^n$ and $H(x, t, f(\cdot)) \in \mathbb{R}^{n\times n}$ represent the gradient and Hessian of the function $f(x,t) : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ with respect to $x \in \mathbb{R}^n$, and $\nabla_{xt} f(x,t) \in \mathbb{R}^n$ denote the partial derivative of $\nabla_x f(x,t)$ with respect to $t \in \mathbb{R}_+$. A symmetric matrix $P \in \mathbb{R}^{n\times n}$ is said to be positive semidefinite, which is denoted by $P \succeq 0$, if $x^T Px \geq 0$ for all $x \in \mathbb{R}^n$. For two symmetric matrices $P \in \mathbb{R}^{n\times n}$ and $Q \in \mathbb{R}^{n\times n}$, $P \succeq Q$ means that $P - Q$ is positive semidefinite.

For a group of $N \in \mathbb{Z}_{++}$ agents, each agent is considered a node and the communication topology among them is denoted by an undirected graph $G(\mathcal{A}) = \{\mathcal{I}, \mathcal{E}, \mathcal{A}\}$. Here, the node index set is denoted by $\mathcal{I} = \{1, 2, ..., N\}$, $\mathcal{E} \subseteq \mathcal{I} \times \mathcal{I}$ is the set of edges (i,j), $i, j \in \mathcal{I}$, which means that agents i and j are able to get information from each other, and $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{N\times N}$ is the adjacency matrix associated with the undirected graph $G(\mathcal{A})$, where $a_{ij} = a_{ji} = 1$ if edge $(j,i) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise, and $a_{ii} = 0$ for all $i \in \mathcal{I}$. An undirected path in an undirected graph is defined by a sequence of edges of the form $(i,j), (j,s), ...$, where $i, j, s \in \mathcal{I}$. The undirected graph is called connected if there exists an undirected path between any two distinct nodes in $\mathcal{I}$. The set of neighbors of node i is denoted by $N_i = \{j : (j,i) \in \mathcal{E}\}$. Let the matrix $L = [l_{ij}] \in \mathbb{R}^{N\times N}$, in which $l_{ii} = \sum_{j=1, j\neq i}^{N} a_{ij}$ and $l_{ij} = -a_{ij}$ for $i \neq j$, denote the Laplacian matrix associated with the undirected graph $G(\mathcal{A})$. By arbitrarily assigning directions for the edges, let $B = [b_{ik}] \in \mathbb{R}^{N\times \ell}$ denote the incidence matrix associated with the undirected graph $G(\mathcal{A})$, where $k = 1, 2, ..., \ell$, and $\ell \in \mathbb{Z}_{++}$ is the cardinality of the edge set $\mathcal{E}$. The incidence matrix describes the relationship among these agents using vertices and edges. That is, for the edge e_k toward the node i, $b_{ik} = 1$, for the edge e_k left from node i, $b_{ik} = -1$, and $b_{ik} = 0$ otherwise. Note that the Laplacian matrix L of the undirected graph $G(\mathcal{A})$ is symmetric positive semidefinite and $L = BB^T$. The eigenvalues of L are ordered as $\lambda_1(L) = 0 < \lambda_2(L) \leq \cdots \leq \lambda_N(L)$ when the undirected graph $G(\mathcal{A})$ is connected, and $\lambda_2(L)$ denotes the second smallest eigenvalue of the Laplacian matrix L, which has the following properties [16]:

$$x^T L x = \frac{1}{2}\sum_{i=1}^{N}\sum_{j\in N_i}(x_j - x_i)^2, \tag{1}$$

$$x^T L x \geq \lambda_2(L) x^T x \quad \text{when } \mathbf{1}_N^T x = 0 \text{ and } x \neq 0 \tag{2}$$

where $x = [x_1, x_2, ..., x_N]^T \in \mathbb{R}^N$. Defining $\Pi \triangleq (\mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^T)$, we then have the following property [17]:

$$\Pi = L(L)^+ = BB^T(BB^T)^+ = B(B^TB)^+B^T \tag{3}$$

for the incidence matrix B and the Laplacian matrix L associated with any connected undirected network $G(\mathcal{A})$. Here, $(\cdot)^+$ is the generalized inverse.

B. Nonsmooth Consensus Algorithms

The nonsmooth consensus idea plays an important role in our distributed optimization algorithm development. We introduce two nonsmooth consensus algorithms that will be exploited in our main results. Especially, in the first algorithm in Section III-A, we need to exploit the nonsmooth consensus idea for systems involving a damping term and a time-varying disturbance term. Consider N agents with dynamics

$$\dot{r}_i(t) = -\nu r_i(t) + d_i(t) + u_i(t), \quad i \in \mathcal{I} \tag{4}$$

where $r_i(t) \in \mathbb{R}$ is the variable associated with the ith agent, $\nu \in \mathbb{R}_{++}$, $u_i(t) \in \mathbb{R}$ is the control input for the ith agent, and $d_i(t) \in \mathbb{R}$ is the time-varying function satisfying $\sup_{t\geq 0}|d_i(t)| < D \in \mathbb{R}_{++}$ for the ith agent. We have the following lemma.

Lemma 1: For the agents with dynamics (4) over a fixed connected undirected graph topology $G(\mathcal{A})$, the variables r_i for $i \in \mathcal{I}$ will reach consensus using the nonsmooth controller

$$u_i(t) = -\eta \sum_{j\in N_i} \text{sign}\left(r_i(t) - r_j(t)\right) \tag{5}$$

where $\eta \in \mathbb{R}_{++}$ satisfies $\eta \geq D\sqrt{\frac{2N}{\lambda_2(L)}}$. Here, $\lambda_2(L)$ is the second smallest eigenvalue of the Laplacian matrix L defined in Section II-A.

Proof: See Appendix A. ∎

Moreover, in the second algorithm in Section III-A, we need to exploit the nonsmooth consensus idea to guarantee the output consensus for systems involving a time-varying damping term and a nonuniform time-varying signal in their outputs. Consider N agents with dynamics

$$\dot{p}_i(t) = \sigma(t)p_i(t) + u_i(t), \tag{6a}$$

$$q_i(t) = \phi_i(t) + p_i(t), \quad i \in \mathcal{I} \tag{6b}$$

where $p_i(t) \in \mathbb{R}$ is the variable associated with the ith agent, $q_i(t) \in \mathbb{R}$ is the output, $|\sigma(t)| \leq \sigma_{\max} \in \mathbb{R}_{++}$, and $\phi_i(t) \in \mathbb{R}$ is the time-varying signal for the ith agent. Here, the time-varying signals $\phi_i(t)$ and $\phi_j(t)$ for each $(i, j) \in \mathcal{E}$ satisfy

$$\sup_{t\geq 0} |\phi_i(t) - \phi_j(t)| \leq \phi_{\max}$$
$$\sup_{t\geq 0} |\dot{\phi}_i(t) - \dot{\phi}_j(t)| \leq \varphi_{\max}. \tag{7}$$

We then have the following lemma.

Lemma 2: For the agents with dynamics (6) over a fixed connected undirected graph topology $G(\mathcal{A})$, the outputs $q_i(t)$ for $i \in \mathcal{I}$ will reach consensus eventually using the controller

$$u_i(t) = -\zeta \sum_{j\in N_i}\left(q_i(t) - q_j(t)\right) - \mu \sum_{j\in N_i} \text{sign}\left(q_i(t) - q_j(t)\right) \tag{8}$$

where $\zeta \in \mathbb{R}_{++}$ satisfies $\zeta > \sigma_{\max} \cdot \lambda_{\max}\{(B^TB)^+\}$, $\mu \in \mathbb{R}_{++}$ satisfies $\mu > 1 + (\varphi_{\max} + \sigma_{\max}\phi_{\max})\left|\left|(B^TB)^+\right|\right|_\infty$, B is the incidence matrix defined in Section II-A, and $\lambda_{\max}\{(B^TB)^+\}$ is the largest eigenvalue of $(B^TB)^+$.

Proof: See Appendix B. ∎

Remark 1: The proposed system (6) with the controller (8) can guarantee the consensus of the outputs when $\sigma(t)$ is time varying and bounded while the methods in [17] deal with the case that $\sigma(t)$ is a negative constant.

C. Prediction–Correction Method and Problem Formulation

Because the proposed distributed algorithms gain some insight from the centralized prediction–correction method in [13], we first review some results from [13]. Consider the time-varying optimization problem

$$x^*(t) = \underset{x(t)\in\mathbb{R}^n}{\text{argmin}}\ f_0\left(x(t), t\right) \quad \text{s.t.} \quad Ax(t) = b \tag{9}$$

where $x^*(t)$ is the minimizer at any given time $t \in \mathbb{R}_+$, the matrix $A \in \mathbb{R}^{q\times n}$ includes the constraint parameters, $q \in \mathbb{Z}_{++}$ satisfies $q < n$, and $b \in \mathbb{R}^q$ is the demand that should be collectively satisfied by all states of $x(t)$. Before giving the lemmas, we need the following assumptions.

Assumption 1: The cost function $f_0(x(t), t)$ is twice continuously differentiable and uniformly strongly convex with respect to $x(t)$ for all t, i.e., $H(x(t), t, f_0(\cdot)) \succeq mI_n$ for some $m \in \mathbb{R}_{++}$, as well as continuously differentiable with respect to t. Here, $H(x(t), t, f_0(\cdot))$ denotes the Hessian of $f_0(x(t), t)$ with respect to $x(t)$.

Assumption 2: The Slater's condition always holds. There exists at least one $x(t) \in \mathbb{R}^n$ such that $Ax(t) = b$ for each t. That is, the optimization problem is feasible at all times.

The uniform strong convexity of $f_0(x(t), t)$ in Assumption 1 ensures the uniqueness of the optimal solution at any given time t. Assumption 2 ensures that the optimal solution $x^*(t)$ for all t can be characterized using the Karush–Kuhn–Tucker conditions.

Define the Lagrange function associated with the optimization problem (9) as

$$\mathcal{L}\left(x(t),\lambda(t),t\right) = f_0\left(x(t),t\right) + \lambda(t)^T\left(Ax(t)-b\right) \quad (10)$$

where $\lambda(t) \in \mathbb{R}^q$ is the Lagrange multiplier. Note that the Lagrange function (9) is strongly convex in $x(t)$ and concave in $\lambda(t)$ under Assumption 1. Define the dual function $\mathcal{F}(\lambda(t),t) = \min_{x(t)\in\mathbb{R}^n} \mathcal{L}(x(t),\lambda(t),t)$ and the corresponding dual optimizer $\lambda^*(t) = \text{argmax}_{\lambda(t)\in\mathbb{R}^q} \mathcal{F}(\lambda(t),t)$. We also define the variable $z(t) = [x(t)^T, \lambda(t)^T]^T \in \mathbb{R}^{n+q}$ and the optimal solution $z^*(t) = [x^*(t)^T, \lambda^*(t)^T]^T \in \mathbb{R}^{n+q}$. The following lemma is then obtained.

Lemma 3 ([13]): If Assumptions 1 and 2 hold, the variable $z(t)$ satisfying

$$H\left(z(t),t,\mathcal{L}(\cdot)\right)\dot{z}(t) = -\alpha\nabla_z\mathcal{L}\left(z(t),t\right) - \nabla_{zt}\mathcal{L}\left(z(t),t\right) \quad (11)$$

globally exponentially converges to the optimal solution $z^*(t)$ with some $\alpha \in \mathbb{R}_{++}$.

Note that for all $z(t) \in \mathbb{R}^{n+q}$, the Hessian of the Lagrange function (10) satisfies $\left|\left|H^{-1}(z(t),t,\mathcal{L}(\cdot))\right|\right|_2 \leq M$ for some $M \in \mathbb{R}_{++}$ when Assumptions 1 and 2 hold. We then have the following lemma.

Lemma 4 ([13]): Suppose that Assumptions 1 and 2 hold. If the time evolution of $\nabla_z\mathcal{L}(z(t),t)$ satisfies

$$\dot{\nabla}_z\mathcal{L}\left(z(t),t\right) = -\alpha\nabla_z\mathcal{L}\left(z(t),t\right)$$

where $\alpha \in \mathbb{R}_{++}$, then the inequality

$$||z(t) - z^*(t)||_2 \leq M\left|\left|\nabla_z\mathcal{L}\left(z(t),t\right)\right|\right|_2 e^{-\alpha t}$$

is obtained, which means that the variable $z(t)$ exponentially converges to the optimal solution $z^*(t)$.

To solve the distributed optimal resource allocation problem for a multiagent system, we aim to propose distributed dynamical systems whose solutions converge to the optimal trajectory

$$x^*(t) = \underset{x(t)\in\mathbb{R}^N}{\text{argmin}} \sum_{i=1}^N f_i\left(x_i(t),t\right) \quad \text{s.t.} \quad \sum_{i=1}^N x_i(t) = b \quad (12)$$

where $x_i(t) \in \mathbb{R}$ is the state of the ith agent, $x(t) = [x_1(t),\ldots,x_N(t)]^T$, $b \in \mathbb{R}$ is the desired constant demand that should be satisfied by all agents collectively, and $f_i(x_i(t),t)$ is the time-varying convex and differentiable local cost function for the ith agent. To gain insight, we first derive the centralized solution to (12) by letting $\sum_{i=1}^N f_i(x_i(t),t)$ play the role of $f_0(x(t),t)$ in (9). In the following, for simplicity, we remove the time-dependence associated with the variables (e.g., x_i instead of $x_i(t)$) when appropriate.

Suppose that the global cost function $\sum_{i=1}^N f_i(x_i,t)$ satisfies Assumption 1. Define the Lagrange function associated with the optimization problem (12) as

$$\begin{aligned}\mathcal{L}\left(x,\lambda,t\right) &= \sum_{i=1}^N f_i(x_i,t) + \lambda\left(\sum_{i=1}^N x_i - b\right)\\ &= \sum_{i=1}^N f_i(x_i,t) + \lambda\left(\sum_{i=1}^N (x_i - b_i)\right)\end{aligned}$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier, and $b_i \in \mathbb{R}$ for $i \in \mathcal{I}$ is designated such that $\sum_{i=1}^N b_i = b$. Also define $z = [x^T,\lambda]^T$. Note that there is only one equality constraint for the optimization problem (12), and the optimal solution always exists for $t \in \mathbb{R}_+$ under Assumption 2. Therefore, by invoking the prediction–correction method described in Lemma 3, the variable z satisfying (11) globally exponentially converges to the optimal solution z^*. Here, $\nabla_z\mathcal{L}(z,t) = [\nabla_{x_1}f_1(x_1,t) + \lambda,\ldots,\nabla_{x_N}f_N(x_N,t)+\lambda, \sum_{i=1}^N x_i - b]^T$, $\nabla_{zt}\mathcal{L}(z,t) = [\nabla_{x_1t}f_1(x_1,t),\ldots,\nabla_{x_Nt}f_N(x_N,t),0]^T$, and $H(z,t,\mathcal{L}(\cdot))$ is given by

$$\begin{bmatrix} H_1\left(x_1,t,f_1(\cdot)\right) & \cdots & 0 & 1\\ \vdots & \ddots & \vdots & \vdots\\ 0 & \cdots & H_N\left(x_N,t,f_N(\cdot)\right) & 1\\ 1 & \cdots & 1 & 0\end{bmatrix}$$

where $H_i(x_i(t),t,f_i(\cdot))$ denotes the Hessian of the local cost function for the ith agent. Substituting the above $\nabla_z\mathcal{L}(z,t)$, $\nabla_{zt}\mathcal{L}(z,t)$, and $H(z,t,\mathcal{L}(\cdot))$ to (11) gives

$$H_i\left(x_i,t,f_i(\cdot)\right)\dot{x}_i + \dot{\lambda} = -\alpha\nabla_{x_i}f_i(x_i,t) - \alpha\lambda - \nabla_{x_it}f_i(x_i,t), \quad (13a)$$

$$\sum_{i=1}^N \dot{x}_i = -\alpha\left(\sum_{i=1}^N x_i - b\right). \quad (13b)$$

Note that (13) is a centralized algorithm as it requires global information from all agents.

When the time-varying cost functions $f_i(x_i,t)$ for $i \in \mathcal{I}$ have identical Hessians, we let $H(t) = H_i(x_i,t,f_i(\cdot))$. Summing up (13a) for all agents, replacing $H_i(x_i,t,f_i(\cdot))$ with $H(t)$ for $i \in \mathcal{I}$, and multiplying both sides of (13b) by $H(t)$, (13) becomes

$$H(t)\sum_{i=1}^N \dot{x}_i + N\dot{\lambda} = -\alpha\sum_{i=1}^N \nabla_{x_i}f_i(x_i,t) - \alpha N\lambda - \sum_{i=1}^N \nabla_{x_it}f_i(x_i,t), \quad (14a)$$

$$H(t)\sum_{i=1}^N \dot{x}_i = -H(t)\alpha\left(\sum_{i=1}^N x_i - b\right). \quad (14b)$$

Manipulating (13a) after replacing $H_i(x_i,t,f_i(\cdot))$ with $H(t)$ and substituting (14b) to (14a), we then obtain that

$$\dot{x}_i = -H^{-1}(t) \times \left(\alpha\nabla_{x_i}f_i(x_i,t) + \alpha\lambda + \nabla_{x_it}f_i(x_i,t) + \dot{\lambda}\right) \quad (15a)$$

$$\dot{\lambda} = -\alpha\lambda - \frac{\alpha}{N}\sum_{i=1}^N \nabla_{x_i}f_i(x_i,t) - \frac{1}{N}\sum_{i=1}^N \nabla_{x_it}f_i(x_i,t) + \frac{\alpha H(t)}{N}\left(\sum_{i=1}^N x_i - b\right). \quad (15b)$$

Note from (15b) that the Lagrange multiplier λ is updated using global information from all agents and (15a) relies on λ, which implies that (15) is a centralized algorithm. The structures of (15) and (13) will be exploited in the next section to derive distributed solutions.

III. Main Results

In this section, we focus on certain time-varying quadratic cost functions for the problem (12). That is, $f_i(x_i, t)$ for $i \in \mathcal{I}$ is given by

$$f_i\left(x_i(t), t\right) = \frac{1}{2} a_i(t) x_i^2(t) + c_i(t) x_i(t) + g_i(t) \tag{16}$$

where $a_i(t) \in \mathbb{R}_{++}$, $c_i(t) \in \mathbb{R}$, and $g_i(t) \in \mathbb{R}$. Note that $\nabla_{x_i} f_i(x_i, t) = a_i(t) x_i + c_i(t)$, $\nabla_{x_i t} f_i(x_i, t) = \dot{a}_i(t) x_i + \dot{c}_i(t)$, and $H_i(x_i, t, f_i(\cdot)) = a_i(t)$. We consider the cases that the quadratic cost functions defined by (16) have identical Hessians and nonidentical constant Hessians. To distinguish the different cases, we keep the time dependency for $a_i(t)$, $c_i(t)$, and $g_i(t)$. We aim to design distributed algorithms for each agent using local information with inspiration from the centralized algorithms (13) and (15).

A. Distributed Time-Varying Optimal Resource Allocation with Identical Hessians

This section considers the case of identical Hessians for the optimization problem defined by (12) and (16).

With identical Hessians in (16), we let $a_i(t) = a(t)$ for $i \in \mathcal{I}$. As a result, (15) becomes

$$\begin{aligned} \dot{x}_i = & - a^{-1}(t) \left(\alpha a(t) x_i + \alpha c_i(t) + \alpha \lambda + \dot{a}(t) x_i \right. \\ & \left. + \dot{c}_i(t) + \dot{\lambda}\right), \end{aligned} \tag{17a}$$

$$\begin{aligned} \dot{\lambda} = & - \alpha\lambda - \frac{\alpha}{N} \sum_{i=1}^{N} c_i(t) - \frac{1}{N} \sum_{i=1}^{N} \left(\dot{a}(t) x_i + \dot{c}_i(t)\right) \\ & - \frac{\alpha b a(t)}{N}. \end{aligned} \tag{17b}$$

Define $w_i = a(t) x_i + \lambda$ for the ith agent. Note that $\dot{w}_i = \dot{a}(t) x_i + a(t) \dot{x}_i + \dot{\lambda}$ for $i \in \mathcal{I}$. After some manipulation, it follows from (17a) that

$$\dot{w}_i = -\alpha w_i - \alpha c_i(t) - \dot{c}_i(t). \tag{18}$$

Now if distributed algorithms can be designed such that agent i's state x_i tracks that of (17a) for $i \in \mathcal{I}$, then the optimization problem defined by (12) and (16) with identical Hessians can be solved in a distributed manner. Considering the definition of w_i for $i \in \mathcal{I}$, the problem is then equivalent to tracking w_i and λ simultaneously for the ith agent. Based on the above idea, we next consider two distributed continuous-time algorithms, each of which has its own merits, to deal with the optimization problem with identical Hessians. In the following, to distinguish between the centralized system (17) and our proposed distributed algorithm, we use $\hat{x}_i$ to denote agent i's state and $\hat{\lambda}_i$ to denote agent i's Lagrange multiplier in our distributed algorithm. Before moving on, an assumption on (16) is needed.

Assumption 3: In (16), $a_i(t) = a(t) \in \mathbb{R}_{++}$, and $|a(t)|$, $|\dot{a}(t)|$, $|c_i(t)|$, and $|\dot{c}_i(t)|$ are upper bounded for $i \in \mathcal{I}$.

1) Algorithm 1 of Identical Hessians Case:

For each agent, we propose the distributed algorithm

$$\dot{\hat{x}}_i = - \alpha(\hat{x}_i - b_i) + \tau a(t)^{-1} \sum_{j \in N_i} \operatorname{sign}(\hat{\lambda}_i - \hat{\lambda}_j), \tag{19a}$$

$$\begin{aligned} \dot{\hat{\lambda}}_i = & - \alpha\hat{\lambda}_i - \alpha c_i(t) - \dot{c}_i(t) - \alpha b_i a(t) - \dot{a}(t)\hat{x}_i \\ & - \tau \sum_{j \in N_i} \operatorname{sign}(\hat{\lambda}_i - \hat{\lambda}_j) \end{aligned} \tag{19b}$$

where $\tau \in \mathbb{R}_{++}$ is the control gain, $\alpha > \sup_{t \geq 0} |\dot{a}(t) a^{-1}(t)|$ can be adjusted to improve the convergence rate, and $b_i \in \mathbb{R}$ for $i \in \mathcal{I}$ is arbitrarily designated such that their sum equals the total demand b. In particular, τ satisfies $\tau \geq D_\tau \sqrt{\frac{2N}{\lambda_2(L)}}$ with $D_\tau > \sup_{t \geq 0} |\alpha c_i(t) + \dot{c}_i(t) + \alpha b_i a(t)| + \sup_{t \geq 0} \left|\dot{a}(t) a^{-1}(t) (\sup_{t \geq 0} |c_i(t) + \alpha^{-1} \dot{c}_i(t)| + |a(0) \hat{x}_i(0) + \hat{\lambda}_i(0)|)\right|$ for all $i \in \mathcal{I}$, $\lambda_2(L)$ is defined in Section II-A, and $b_i = b/N$ for $i \in \mathcal{I}$ is a possible selection, which implies that every agent knows the common demand and the size of the group. Another special selection is that only one agent knows the common demand, for example, $b_1 = b$, and $b_i = 0$ for $i \in \{2, 3, ..., N\}$. Note that due to Assumption 3, α and D_τ are well defined. The following theorem is then obtained.

Theorem 1: Suppose that the fixed graph $G(\mathcal{A})$ is undirected and connected. If Assumptions 2 and 3 hold, then the state $\hat{x}_i$ and Lagrange multiplier $\hat{\lambda}_i$ of the ith agent with dynamics (19) will converge to the corresponding optimal state and Lagrange multiplier as $t \to \infty$ for the optimization problem defined by (12) and (16), respectively.

Proof: Note that Assumption 3 guarantees that the global cost function $\sum_{i=1}^{N} f_i(x_i, t)$ satisfies Assumption 1. Then, it follows from the derivation of (15) that the solution of the dynamical system (17) will converge to the corresponding optimal trajectory for the optimization problem defined by (12) and (16). Next, we show that the state $\hat{x}_i$ and Lagrange multiplier $\hat{\lambda}_i$ of the ith agent with dynamics (19) can track the corresponding part in (17) in a distributed manner.

Define $\hat{w}_i = a(t)\hat{x}_i + \hat{\lambda}_i$ for $i \in \mathcal{I}$. After some manipulation, it follows from (19) that $\dot{\hat{w}}_i = \dot{a}(t)\hat{x}_i + a(t)\dot{\hat{x}}_i + \dot{\hat{\lambda}}_i = -a(t)\alpha\hat{x}_i - \alpha\hat{\lambda}_i - \alpha c_i(t) - \dot{c}_i(t)$ for $i \in \mathcal{I}$, which can be written as

$$\dot{\hat{w}}_i = -\alpha\hat{w}_i - \alpha c_i(t) - \dot{c}_i(t). \tag{20}$$

Define $e_{wi} = \hat{w}_i - w_i$ for $i \in \mathcal{I}$. It follows from (18) and (20) that $\dot{e}_{wi} = -\alpha e_{wi}$, which implies that $\lim_{t \to \infty} e_{wi} = 0$ for $i \in \mathcal{I}$.

Because the fixed graph $G(\mathcal{A})$ is undirected, we have $\sum_{i=1}^{N} \sum_{j \in N_i} \operatorname{sign}(\hat{\lambda}_i - \hat{\lambda}_j) = 0$. Note that $\sum_{i=1}^{N} b_i = b$. Summing up (19a) for all agents, we have $\sum_{i=1}^{N} \dot{\hat{x}}_i = -\alpha(\sum_{i=1}^{N} \hat{x}_i - b)$. Substituting (17b) to (17a) and summing up (17a) for all agents, we obtain that $\sum_{i=1}^{N} \dot{x}_i = -\alpha(\sum_{i=1}^{N} x_i - b)$. Define $e_x = \sum_{i=1}^{N} \hat{x}_i - \sum_{i=1}^{N} x_i$. It follows that $\dot{e}_x = -\alpha e_x$, which implies that $\lim_{t \to \infty} e_x = 0$.

Summing up (17b) for all agents, we obtain that

$$\sum_{i=1}^{N}\dot{\lambda} = -\alpha N\lambda - \sum_{i=1}^{N}\dot{a}(t)x_i - \alpha\sum_{i=1}^{N}c_i(t) - \sum_{i=1}^{N}\dot{c}_i(t) - \alpha b a(t). \quad (21)$$

Because the fixed graph $G(\mathcal{A})$ is undirected and $\sum_{i=1}^{N} b_i = b$, we obtain from (19b) that

$$\sum_{i=1}^{N}\dot{\hat{\lambda}}_i = -\alpha\sum_{i=1}^{N}\hat{\lambda}_i - \sum_{i=1}^{N}\dot{a}(t)\hat{x}_i - \alpha\sum_{i=1}^{N}c_i(t) - \sum_{i=1}^{N}\dot{c}_i(t) - \alpha b a(t). \quad (22)$$

Define $e_\lambda = \sum_{i=1}^{N}\hat{\lambda}_i - N\lambda$. It follows from (21) and (22) that $\dot{e}_\lambda = -\alpha e_\lambda - \dot{a}(t)e_x$. Because $\lim_{t\to\infty} e_x = 0$ and $|\dot{a}(t)|$ is upper bounded due to Assumption 3, we then have $\lim_{t\to\infty} e_\lambda = 0$.

Because $|c_i(t)|$ and $|\dot{c}_i(t)|$ are upper bounded due to Assumption 3, it follows from (20) that $\hat{w}_i(t)$ is bounded and satisfies $|\hat{w}_i(t)| \le \sup_{t\ge 0}|c_i(t) + \alpha^{-1}\dot{c}_i(t)| + |\hat{w}_i(0)|$. Replacing $\hat{w}_i(0)$ with $a(0)\hat{x}_i(0) + \hat{\lambda}_i(0)$, we have $|\hat{w}_i(t)| \le \sup_{t\ge 0}|c_i(t) + \alpha^{-1}\dot{c}_i(t)| + |a(0)\hat{x}_i(0) + \hat{\lambda}_i(0)|$. According to the definition of $\hat{w}_i$, we have $\hat{x}_i = a^{-1}(t)(\hat{w}_i - \hat{\lambda}_i)$. Substituting it to the right-hand side of (19b), we have

$$\dot{\hat{\lambda}}_i = -\left(\alpha - \dot{a}(t)a^{-1}(t)\right)\hat{\lambda}_i + s_i(t) - \tau\sum_{j\in N_i}\operatorname{sign}(\hat{\lambda}_i - \hat{\lambda}_j)$$

where $s_i(t) = -\alpha c_i(t) - \dot{c}_i(t) - \alpha b_i a(t) - \dot{a}(t)a^{-1}(t)\hat{w}_i$. Based on the definition of D_τ, it can be verified that $D_\tau > \sup_{t\ge 0}|s_i(t)|$. Because $\alpha > \sup_{t\ge 0}|\dot{a}(t)a^{-1}(t)|$ and $\tau \ge D_\tau\sqrt{\frac{2N}{\lambda_2(L)}}$, it follows from Lemma 1 that $\hat{\lambda}_i$ for $i \in \mathcal{I}$ will reach consensus, i.e., $\hat{\lambda}_i \to \hat{\lambda}_j$, $\forall i, j \in \mathcal{I}$, as $t \to \infty$.

Combining with the fact that $\lim_{t\to\infty} e_\lambda = 0$, it follows that $\lim_{t\to\infty}|\hat{\lambda}_i - \lambda| = 0$ for $i \in \mathcal{I}$. By further invoking the definitions of $\hat{w}_i$ and w_i and the fact that $\lim_{t\to\infty} e_{wi} = 0$, we obtain that $\lim_{t\to\infty}|\hat{x}_i - x_i| = 0$ for $i \in \mathcal{I}$.

Since the solution of (17) will converge to the optimal trajectories x^* and λ^* for the optimization problem defined by (12) and (16), the state $\hat{x}_i$ and Lagrange multiplier $\hat{\lambda}_i$ of the ith agent with dynamics (19) will converge to the corresponding optimal state and Lagrange multiplier, respectively. ■

Remark 2: For the optimal resource allocation problem with time-varying resources as in [15], our proposed algorithm (19) can be slightly modified to handle the same situation when the cost functions have identical Hessians. Suppose that the resource $b_i(t) \in \mathbb{R}$ assigned to the ith agent is bounded as well as its time derivative $\dot{b}_i(t) \in \mathbb{R}$ as in [15] for $i \in \mathcal{I}$. For each agent, we propose the distributed algorithm

$$\dot{\hat{x}}_i = -\alpha(\hat{x}_i - b_i) + \dot{b}_i + \tau a(t)^{-1}\sum_{j\in N_i}\operatorname{sign}(\hat{\lambda}_i - \hat{\lambda}_j), \quad (23a)$$

$$\dot{\hat{\lambda}}_i = -\alpha\hat{\lambda}_i - \alpha c_i(t) - \dot{c}_i(t) - \alpha b_i a(t) - a(t)\dot{b}_i - \dot{a}(t)\hat{x}_i - \tau\sum_{j\in N_i}\operatorname{sign}(\hat{\lambda}_i - \hat{\lambda}_j) \quad (23b)$$

where $\alpha > \sup_{t\ge 0}|\dot{a}(t)a^{-1}(t)|$, and τ satisfies $\tau \ge D_\tau\sqrt{\frac{2N}{\lambda_2(L)}}$ with $D_\tau > \sup_{t\ge 0}|\alpha c_i(t) + \dot{c}_i(t) + \alpha b_i a(t) + a(t)\dot{b}_i| + \sup_{t\ge 0}\left|\dot{a}(t)a^{-1}(t)(\sup_{t\ge 0}|c_i(t) + \alpha^{-1}\dot{c}_i(t)| + |a(0)\hat{x}_i(0) + \hat{\lambda}_i(0)|)\right|$ for all $i \in \mathcal{I}$. The proof is similar to Theorem 1 and hence is omitted here. Note that while [15] allows for nonidentical Hessians, it requires $a_i(t)$ and $c_i(t)$ in (16) to be constant. In contrast, with identical Hessians, the algorithm (23) is able to deal with time-varying $a_i(t)$ and $c_i(t)$ in (16).

Note that the selection of τ is affected by D_τ, which depends on all agents' initial state variables $\hat{x}_i(0)$ and $\hat{\lambda}_i(0)$. To relax such a requirement, we develop the following algorithm where the control gains are independent on the initial state variables.

2) Algorithm 2 of Identical Hessians Case: For each agent, we propose the following distributed algorithm:

$$\dot{b}_i = a^{-1}(t) \times \left(\zeta_o\sum_{j\in N_i}(\delta_i - \delta_j) + \mu_o\sum_{j\in N_i}\operatorname{sign}(\delta_i - \delta_j)\right), \quad (24a)$$

$$\dot{\hat{x}}_i = -\alpha(\hat{x}_i - b_i) + \dot{b}_i, \quad (24b)$$

$$\dot{\hat{\lambda}}_i = -\alpha\hat{\lambda}_i - \alpha c_i(t) - \dot{c}_i(t) - a(t)\alpha b_i - a(t)\dot{b}_i - \dot{a}(t)\hat{x}_i, \quad (24c)$$

$$\delta_i = -c_i(t) - a(t)b_i \quad (24d)$$

where $\zeta_o \in \mathbb{R}_{++}$ satisfies $\zeta_o > D_\zeta\lambda_{\max}\{(B^TB)^+\}$, $D_\zeta \in \mathbb{R}_{++}$ is the upper bound of $|\dot{a}(t)a^{-1}(t)|$, $\lambda_{\max}\{(B^TB)^+\}$ is the largest eigenvalue of $(B^TB)^+$, $\mu_o \in \mathbb{R}_{++}$ satisfies $\mu_o > 1 + \left|\left|(B^TB)^+\right|\right|_\infty\left|\left|B^T(-\dot{C}(t) + \dot{a}(t)a^{-1}(t)C(t))\right|\right|_\infty$, $C(t)$ denotes $[c_1(t), ..., c_2(t)]^T$, and $b_i \in \mathbb{R}$ for $i \in \mathcal{I}$ is initially designated such that $\sum_{i=1}^{N} b_i(0) = b$. The following theorem is then obtained.

Theorem 2: Suppose that the fixed graph $G(\mathcal{A})$ is undirected and connected. If Assumptions 2 and 3 hold, then the state $\hat{x}_i$ and Lagrange multiplier $\hat{\lambda}_i$ of the ith agent with dynamics (24) will converge to the corresponding optimal state and Lagrange multiplier as $t \to \infty$ for the optimization problem defined by (12) and (16), respectively.

Proof: Note that Assumption 3 guarantees that the global cost function $\sum_{i=1}^{N} f_i(x_i, t)$ satisfies Assumption 1. Next, we show that the state $\hat{x}_i$ and Lagrange multiplier $\hat{\lambda}_i$ of the ith agent with dynamics (24) can track the corresponding part in (17) in a distributed manner.

Define $\hat{w}_i = a(t)\hat{x}_i + \hat{\lambda}_i$ for $i \in \mathcal{I}$. Note that $a(t)$ is time varying according to Assumption 3. After some manipulation, it follows from (24b) and (24c) that $\dot{\hat{w}}_i = \dot{a}(t)\hat{x}_i + a(t)\dot{\hat{x}}_i + \dot{\hat{\lambda}}_i = -a(t)\alpha\hat{x}_i - \alpha\hat{\lambda}_i - \alpha c_i(t) - \dot{c}_i(t)$ for $i \in \mathcal{I}$, which can also be written as (20). Define $e_{wi} = \hat{w}_i - w_i$ for $i \in \mathcal{I}$. It follows from (18) and (20) that $\dot{e}_{wi} = -\alpha e_{wi}$, which implies that $\lim_{t\to\infty} e_{wi} = 0$ for $i \in \mathcal{I}$.

Substituting (17b) to (17a) and summing up (17a) for all agents, we obtain that $\sum_{i=1}^N \dot{x}_i = -\alpha(\sum_{i=1}^N x_i - b)$. Because the fixed graph $G(\mathcal{A})$ is undirected, it follows from (24a) that $\sum_{i=1}^N \dot{b}_i = 0$. We then obtain that $\sum_{i=1}^N \dot{\hat{x}}_i = -\alpha(\sum_{i=1}^N \hat{x}_i - b)$. Define $e_x = \sum_{i=1}^N \hat{x}_i - \sum_{i=1}^N x_i$. It follows that $\dot{e}_x = -\alpha e_x$, which implies that $\lim_{t\to\infty} e_x = 0$.

Note that (24c) is equivalent to

$$\dot{\hat{\lambda}}_i = -\alpha\hat{\lambda}_i - \alpha c_i(t) - a(t)\alpha b_i - \dot{c}_i(t) - a(t)\dot{b}_i - \dot{a}(t)b_i - \dot{a}(t)(\hat{x}_i - b_i). \tag{25}$$

It follows from (24d) that $\dot{\delta}_i = -\dot{c}_i(t) - \dot{a}(t)b_i - a(t)\dot{b}_i$ for $i \in \mathcal{I}$. Substituting (24d) and $\dot{\delta}_i$ to (25), we obtain that (24b)–(24d) can be written as

$$(\dot{\hat{x}}_i - \dot{b}_i) = -\alpha(\hat{x}_i - b_i),$$
$$(\dot{\hat{\lambda}}_i - \dot{\delta}_i) = -\alpha(\hat{\lambda}_i - \delta_i) - \dot{a}(t)(\hat{x}_i - b_i)$$

which implies that $\hat{\lambda}_i \to \delta_i$ for $i \in \mathcal{I}$ as $t \to \infty$ under Assumption 3 that $|\dot{a}(t)|$ is bounded. Define $v_i = -a(t)b_i$ such that $\dot{v}_i = -\dot{a}(t)b_i - a(t)\dot{b}_i$ for $i \in \mathcal{I}$. Substituting v_i to (24d) and using (24a), we obtain that

$$\delta_i = -c_i(t) + v_i,$$
$$\dot{v}_i = \dot{a}(t)a^{-1}(t)v_i - \zeta_o \sum_{j\in N_i} (\delta_i - \delta_j) - \mu_o \sum_{j\in N_i} \text{sign}(\delta_i - \delta_j).$$

Because the fixed undirected graph $G(\mathcal{A})$ is connected, $|a(t)|$ and $|\dot{a}(t)|$ are bounded, and $|c_i(t)|$ and $|\dot{c}_i(t)|$ are bounded for $i \in \mathcal{I}$ under Assumption 3, it follows from Lemma 2 that δ_i will reach consensus as $t \to \infty$ for $i \in \mathcal{I}$. Therefore, we can obtain that $\hat{\lambda}_i \to \hat{\lambda}_j$, $\forall i, j \in \mathcal{I}$, as $t \to \infty$. Combining with the fact that $\lim_{t\to\infty} \sum_{i=1}^N e_{wi} = 0$ and $\lim_{t\to\infty} e_x = 0$, it follows that $\sum_{i=1}^N \hat{\lambda}_i \to N\lambda$ as $t \to \infty$, which results in $\lim_{t\to\infty} |\hat{\lambda}_i - \lambda| = 0$ for $i \in \mathcal{I}$. By further invoking the definitions of $\hat{w}_i$ and w_i, we obtain that $\lim_{t\to\infty} |\hat{x}_i - x_i| = 0$ for $i \in \mathcal{I}$.

Since the solution of (17) will converge to the optimal trajectories x^* and λ^* for the optimization problem defined by (12) and (16), the state $\hat{x}_i$ and Lagrange multiplier $\hat{\lambda}_i$ of the ith agent with dynamics (24) will converge to the corresponding optimal state and Lagrange multiplier, respectively. ■

Remark 3: The algorithm (24) can be slightly modified to relax the assumption that the upper bounds of $|c_i(t)|$ and $|\dot{c}_i(t)|$ for $i \in \mathcal{I}$ are known in advance by replacing the parameter μ_o in (24a) with some adaptive gains as in [17]. But it is still necessary to know the bounds of the identical time-varying Hessian and its time derivative.

Remark 4: Each of (19) and (24) has its own merits. It is more efficient to implement (19), which has fewer variables and less computational cost than (24). The control gains of (24), however, are independent on the initial state variables. It is also worth mentioning that in the special case that $a(t)$ is a constant, the control gain τ in (19) would not rely on the initial state variables but instead satisfy $\tau \geq D_\tau \sqrt{\frac{2N}{\lambda_2(L)}}$ with $D_\tau > \sup_{t\geq 0} |\alpha c_i(t) + \dot{c}_i(t) + \alpha b_i a(t)|$ for $i \in \mathcal{I}$. In addition, (19) can be modified to handle the situation in [15] where the resource assigned to each agent is time varying and bounded (see Remark 2) but not for (24).

B. Estimator-Based Distributed Time-Varying Optimal Resource Allocation With Nonidentical Constant Hessians

This section considers the case of nonidentical constant Hessians for the optimization problem defined by (12) and (16).

With nonidentical constant Hessians in (16), we let $a_i(t) = a_i$ for $i \in \mathcal{I}$. As a result, (13) becomes

$$a_i\dot{x}_i + \dot{\lambda} = -a_i\alpha x_i - \alpha c_i(t) - \alpha\lambda - \dot{c}_i(t), \tag{26a}$$

$$\sum_{i=1}^N \dot{x}_i = -\alpha\left(\sum_{i=1}^N x_i - b\right). \tag{26b}$$

By manipulating (26a), we then obtain that

$$\dot{x}_i = -a_i^{-1}\left(a_i\alpha x_i + \alpha c_i(t) + \alpha\lambda + \dot{c}_i(t) + \dot{\lambda}\right). \tag{27}$$

Summing up (27) for all agents results in

$$\sum_{i=1}^N \dot{x}_i = -\alpha\sum_{i=1}^N x_i - \sum_{i=1}^N a_i^{-1}\left(\alpha c_i(t) + \dot{c}_i(t)\right) - (\alpha\lambda + \dot{\lambda})\sum_{i=1}^N a_i^{-1}. \tag{28}$$

Letting the right sides of (26b) and (28) be equal, we obtain that

$$\dot{\lambda} = -\alpha\lambda - \left(\sum_{i=1}^N \left(a_i^{-1}\left(\alpha c_i(t) + \dot{c}_i(t)\right) + \alpha b_i\right)\right) \times \left(\sum_{i=1}^N a_i^{-1}\right)^{-1}. \tag{29}$$

Note from (29) that the Lagrange multiplier λ is updated using global information from all agents and (27) relies on λ, which implies that (27) with (29) is a centralized algorithm.

In order to relax the condition that every agent knows the common global information $\sum_{i=1}^N (a_i^{-1}\alpha c_i(t) + a_i^{-1}\dot{c}_i(t) + \alpha b_i)$ and $\sum_{i=1}^N a_i^{-1}$ in (29), we propose a distributed estimator, which uses only local information, for each agent. Before moving on, we need the following assumption.

Assumption 4: In (16), $a_i(t) = a_i \in \mathbb{R}_{++}$ is a constant, and $|c_i(t)|$, $|\dot{c}_i(t)|$, and $|\ddot{c}_i(t)|$ are upper bounded for $i \in \mathcal{I}$.

For each agent, we design the estimator

$$\dot{\xi}_i(t) = \gamma \sum_{j\in N_i} \text{sign}\left(\omega_j(t) - \omega_i(t)\right), \tag{30a}$$

$$\omega_i(t) = \xi_i(t) + a_i^{-1}\left(\alpha c_i(t) + \dot{c}_i(t)\right) + \alpha b_i, \tag{30b}$$

$$\dot{\psi}_i(t) = \beta \sum_{j\in N_i} \text{sign}\left(\theta_j(t) - \theta_i(t)\right), \tag{30c}$$

$$\theta_i(t) = \psi_i(t) + a_i^{-1} \tag{30d}$$

where $\beta \in \mathbb{R}_{++}$, and $\gamma \in \mathbb{R}_{++}$ satisfies $\gamma > \sup_{t\geq 0} \left|\left|a_i^{-1}\alpha\dot{c}_i(t) + a_i^{-1}\ddot{c}_i(t)\right|\right|_\infty$ for $i \in \mathcal{I}$. In the following, to distinguish between the centralized system given by (27) and (29) and our proposed distributed algorithm, we use $\hat{x}_i$ to denote agent i's state and $\hat{\lambda}_i$ to denote agent i's Lagrange multiplier in our distributed algorithm. By leveraging the estimator (30), we propose the distributed algorithm

$$\dot{\hat{x}}_i = -a_i^{-1}\left(a_i\alpha\hat{x}_i + \alpha c_i(t) + \alpha\hat{\lambda}_i + \dot{c}_i(t) + \dot{\hat{\lambda}}_i\right), \tag{31a}$$

$$\dot{\hat{\lambda}}_i = -\alpha\hat{\lambda}_i - \theta_i^{-1}(t)\omega_i(t) - \eta_o \sum_{j\in N_i} \operatorname{sign}(\hat{\lambda}_i - \hat{\lambda}_j) \tag{31b}$$

where $\eta_o \in \mathbb{R}_{++}$ is the gain determined later, $\alpha \in \mathbb{R}_{++}$ can be adjusted to improve the convergence rate, and $b_i \in \mathbb{R}$ for $i \in \mathcal{I}$ is designated such that $\sum_{i=1}^N b_i = b$. The following theorem is then obtained.

Theorem 3: Suppose that the fixed graph $G(\mathcal{A})$ is undirected and connected. If Assumptions 2 and 4 hold, then the state $\hat{x}_i$ and Lagrange multiplier $\hat{\lambda}_i$ of the ith agent with dynamics (31) using the estimator (30) will converge to the corresponding optimal state and Lagrange multiplier as $t \to \infty$ for the optimization problem defined by (12) and (16), respectively.

Proof: Note that Assumption 4 guarantees that the global cost function $\sum_{i=1}^N f_i(x_i, t)$ satisfies Assumption 1. Next, we show that the state $\hat{x}_i$ and Lagrange multiplier $\hat{\lambda}_i$ of the ith agent with dynamics (31) will converge to the corresponding optimal state and Lagrange multiplier in a distributed manner.

Because the fixed graph $G(\mathcal{A})$ is undirected and connected, $|\dot{c}_i(t)|$ and $|\ddot{c}_i(t)|$ are upper bounded for $i \in \mathcal{I}$, and the Hessian a_i is a constant for $i \in \mathcal{I}$ under Assumption 4, it follows from Theorem 1 in [18] that there exists a $T \in \mathbb{R}_{++}$ such that $\left|\left|\omega_i(t) - \frac{1}{N}\sum_{i=1}^N(a_i^{-1}\alpha c_i(t) + a_i^{-1}\dot{c}_i(t) + \alpha b_i)\right|\right|_2 = 0$, and $\left|\left|\theta_i(t) - \frac{1}{N}\sum_{i=1}^N a_i^{-1}\right|\right|_2 = 0$ for all $t \geq T$. Define $\overline{\omega}(t) = 1/N\sum_{i=1}^N(a_i^{-1}\alpha c_i(t) + a_i^{-1}\dot{c}_i(t) + \alpha b_i)$ and $\overline{\theta}(t) = 1/N\sum_{i=1}^N a_i^{-1}$. It follows that $\omega_i(t) = \omega_j(t) = \overline{\omega}(t)$, $\theta_i(t) = \theta_j(t) = \overline{\theta}(t)$, $\forall i, j \in \mathcal{I}$, for all $t \geq T$.

Define $S_i(t) = -\theta_i^{-1}(t)\omega_i(t)$ for $i \in \mathcal{I}$. It follows from (31b) that $\dot{\hat{\lambda}}_i = -\alpha\hat{\lambda}_i + S_i(t) - \eta_o\sum_{j\in N_i}\operatorname{sign}(\hat{\lambda}_i - \hat{\lambda}_j)$ for $i \in \mathcal{I}$. For $t \geq T$, define $\overline{S}(t) = -\overline{\theta}^{-1}(t)\overline{\omega}(t)$, and it follows that $S_i(t) = S_j(t) = \overline{S}(t)$, $\forall i, j \in \mathcal{I}$. Denoting $\Psi_i(\hat{\lambda}_i, t) = -\alpha\hat{\lambda}_i + S_i(t)$ for $i \in \mathcal{I}$, we obtain that $|\Psi_i(\hat{\lambda}_i, t) - \Psi_j(\hat{\lambda}_j, t)| \leq \alpha|\hat{\lambda}_i - \hat{\lambda}_j|$, $\forall i, j \in \mathcal{I}$, for $t \geq T$. For the fixed connected undirected graph $G(\mathcal{A})$, there is a finite time $T^{'} \in \mathbb{R}_{++}$ and $\hat{\lambda} \in \mathbb{R}$ such that $\hat{\lambda}_i = \hat{\lambda}_j = \hat{\lambda}$, $\forall i, j \in \mathcal{I}$, when $t > T + T^{'}$, if we choose an appropriate parameter η_o according to Theorem 3.1 in [19]. Specifically, η_o is set such that $\eta_o > \frac{2(N-1)\alpha\cdot\max_i|\hat{\lambda}_i(T) - \overline{\hat{\lambda}}(T)|}{N\cdot a_{\min}}$, where $a_{\min}$ is the minimum positive entry of the adjacency matrix $\mathcal{A}$, $\overline{\hat{\lambda}}(t)$ denotes $1/N\sum_{j=1}^N\hat{\lambda}_j(t)$, and $\max_i|\hat{\lambda}_i(T) - \overline{\hat{\lambda}}(T)|$ is the maximum value of $|\hat{\lambda}_i(T) - \overline{\hat{\lambda}}(T)|$ for $i \in \mathcal{I}$. It follows from (31b) that

$$\begin{aligned}\dot{\hat{\lambda}} + \alpha\hat{\lambda} &= \overline{S}(t) = -\overline{\theta}^{-1}(t)\overline{\omega}(t)\\ &= -\left(\sum_{i=1}^N a_i^{-1}\right)^{-1}\left(\sum_{i=1}^N a_i^{-1}\left(\alpha c_i(t) + \dot{c}_i(t)\right) + \alpha b\right)\end{aligned} \tag{32}$$

where $\overline{\theta}(t)$ and $\overline{\omega}(t)$ are the common global information for $t > T + T^{'}$. Summing up (31a) for all agents and replacing $\dot{\hat{\lambda}}_i + \alpha\hat{\lambda}_i$ with (32) for $i \in \mathcal{I}$, we obtain that

$$\begin{aligned}\sum_{i=1}^N \dot{\hat{x}}_i &= -\sum_{i=1}^N a_i^{-1}\left(a_i\alpha\hat{x}_i + \alpha c_i(t) + \dot{c}_i(t)\right)\\ &\quad - (\dot{\hat{\lambda}} + \alpha\hat{\lambda})\cdot\sum_{i=1}^N a_i^{-1} = -\alpha\left(\sum_{i=1}^N \hat{x}_i - b\right).\end{aligned} \tag{33}$$

For $t > T + T^{'}$, define the Lyapunov function candidate $V = \frac{1}{2}(\nabla_{\hat{z}}\mathcal{L}(\hat{z}, t))^T(\nabla_{\hat{z}}\mathcal{L}(\hat{z}, t))$, where $\hat{z} = [\hat{x}_1, ..., \hat{x}_N, \hat{\lambda}]^T$ and $\mathcal{L}(\hat{z}, t) = \sum_{i=1}^N f_i(\hat{x}_i, t) + \hat{\lambda}(\sum_{i=1}^N \hat{x}_i - b)$. We then obtain the time derivative of V along (31) as

$$\begin{aligned}\dot{V} &= (\nabla_{\hat{z}}\mathcal{L}(\hat{z}, t))^T\cdot\left(H\left(\hat{z}, t, \mathcal{L}(\cdot)\right)\cdot\dot{\hat{z}} + \nabla_{\hat{z}t}\mathcal{L}(\hat{z}, t)\right)\\ &= (\nabla_{\hat{z}}\mathcal{L}(\hat{z}, t))^T\cdot\begin{bmatrix}[-a_i\alpha\hat{x}_i - \alpha c_i(t) - \alpha\hat{\lambda}]_N\\ \sum_{i=1}^N\dot{\hat{x}}_i\end{bmatrix}\end{aligned} \tag{34}$$

where $[-a_i\alpha\hat{x}_i - \alpha c_i(t) - \alpha\hat{\lambda}]_N$ denotes $[-a_1\alpha\hat{x}_1 - \alpha c_1(t) - \alpha\hat{\lambda}, ..., -a_N\alpha\hat{x}_N - \alpha c_N(t) - \alpha\hat{\lambda}]^T$, and $H(\hat{z}, t, \mathcal{L}(\cdot))$ is the Hessian of $\mathcal{L}(\hat{z}, t)$ with respect to $\hat{z}$. Note that $\nabla_{\hat{x}_i}f_i(\hat{x}_i, t) = a_i\hat{x}_i + c_i(t)$ for $i \in \mathcal{I}$. Substituting (33) to (34), we obtain that

$$\begin{aligned}\dot{V} &= (\nabla_{\hat{z}}\mathcal{L}(\hat{z}, t))^T\cdot\begin{bmatrix}[-\alpha\nabla_{\hat{x}_i}f_i(\hat{x}_i, t) - \alpha\hat{\lambda}]_N\\ -\alpha\left(\sum_{i=1}^N\hat{x}_i - b\right)\end{bmatrix}\\ &= -\alpha\left(\nabla_{\hat{z}}\mathcal{L}(\hat{z}, t)\right)^T\left(\nabla_{\hat{z}}\mathcal{L}(\hat{z}, t)\right) \leq 0.\end{aligned}$$

It follows from Lemma 4 that the variable $\hat{z}$ will converge to the optimal solution z^* for the optimization problem defined by (12) and (16). Thus, the state $\hat{x}_i$ and Lagrange multiplier $\hat{\lambda}_i$ of the ith agent with dynamics (31) using the estimator (30) will converge to the corresponding optimal state and Lagrange multiplier, respectively. ■

Remark 5: The algorithm (31) is able to solve the case with nonidentical constant Hessians at the cost of increasing communication over the network because the variables $\omega_i(t)$ and $\theta_i(t)$ for $i \in \mathcal{I}$ associated with the ith agent must be communicated to its neighbors besides $\hat{\lambda}_i$. In contrast, the algorithm (19) just needs to exchange the $\hat{\lambda}_i$.

Remark 6: The algorithm (31) can be modified to handle the situation in [15], where the resource assigned to each agent is time varying and bounded, by redesigning the estimator (30b) as $\omega_i(t) = \xi_i(t) + a_i^{-1}(\alpha c_i(t) + \dot{c}_i(t)) + \alpha b_i(t) + \dot{b}_i(t)$. The proof is similar to Theorem 3 and hence is omitted here. Note again that [15] requires time-invariant quadratic cost functions.

IV. Numerical Simulations

In this section, numerical examples are provided to illustrate the effectiveness of the proposed distributed continuous-time

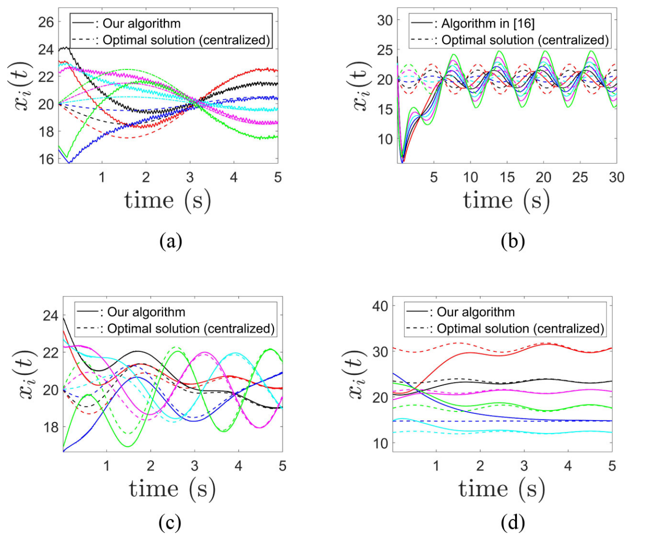

Fig. 1. State trajectories in four scenarios. (a) Scenario I. (b) Scenario II. (c) Scenario III. (d) Scenario IV.

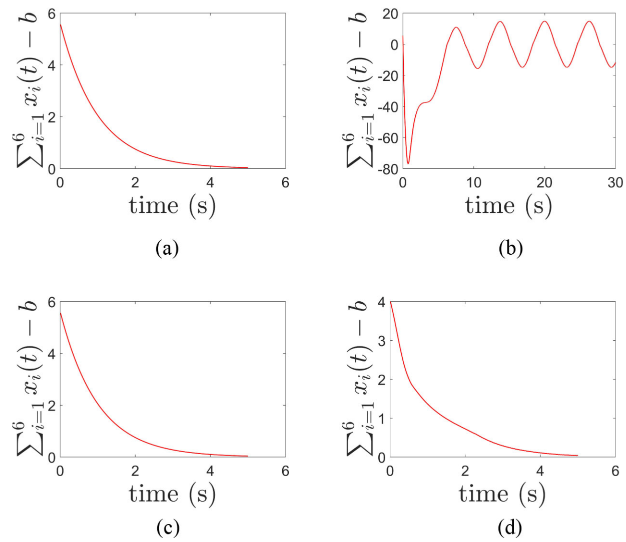

Fig. 2. Demand violation $\sum_{i=1}^{N} x_i - b$ in four scenarios. (a) Scenario I. (b) Scenario II. (c) Scenario III. (d) Scenario IV.

optimization algorithms for the optimal resource allocation problem defined by (12) and (16). We consider a multiagent system composed of $N = 6$ agents communicating over a fixed ring graph.

In the following four scenarios, the common demand that all agents should satisfy cooperatively is set as $b = 120$. In scenario I, each agent has identical constant Hessians. The local cost functions are given by $f_i(x_i, t) = x_i^2 - 2i\sin(t)x_i + i^2\sin^2(t)$ for $i \in \mathcal{I}$. The parameters of the algorithm (19) are set as $\alpha = 1$ and $\tau = 8$, and b_i is set as b/N for $i \in \mathcal{I}$. In scenario II, for comparison, we employ the method proposed in [15] to the case in scenario I.In scenario III, each agent has identical time-varying Hessians. The local cost functions are given by $f_i(x_i, t) = 0.5a(t)x_i + c_i(t)x_i + g_i(t)$, where $a(t) = 8 + 2\sin(0.34t)$, $c_i(t) = -4i\cos(0.5it)$, and $g_i(t) = i^2\sin(0.5it)$ for $i \in \mathcal{I}$. The parameters of the algorithm (24) are set as $\alpha = 1$, $\mu_o = 10$, $\zeta_o = 25$, and $b_i(0)$ is set as b/N for $i \in \mathcal{I}$. In scenario IV, each agent has nonidentical constant Hessians. The local cost functions are given by $f_i(x_i, t) = (a_i x_i - i\sin(\frac{\pi}{4}t))^2$, where $a_1 = 3.1211$, $a_2 = 3.5716$, $a_3 = 4.5144$, $a_4 = 4.9527$, $a_5 = 3.7608$, and $a_6 = 4.1356$. The parameters in (30) and (31) are set as $\alpha = 1$, $\eta_o = 5$, $\gamma = 10$, and $\beta = 0.1$, and we set $b_i = b/N$ for $i \in \mathcal{I}$.

The trajectories of the states in all four scenarios are presented in Fig. 1 and compared with the results obtained by Matlab's constraint optimization solver (centralized approach). The optimal solutions using the Matlab tool are depicted in dashed lines, and the states using our algorithms and the method in [15] are depicted in solid lines. The corresponding time evolution of the demand violation, i.e., $\sum_{i=1}^{N} x_i - b$, is shown in Fig. 2. As shown in Fig. 1(a), (c), and (d), each agent's state converges to its corresponding optimal state of the optimization problem defined by (12) and (16) using our proposed algorithms (19), (24), and (31), respectively, and as shown in Fig. 2(a), (c), and (d), the corresponding demand violations approach zero. In contrast, Fig. 1(b) shows that the states using the method in [15] cannot converge to the optimal solutions while the corresponding demand violation does not approach zero in the case of identical constant Hessians.

V. Conclusion

This article studied the distributed continuous-time optimization algorithms for the time-varying optimal resource allocation problem with certain quadratic time-varying cost functions. The proposed distributed dynamical systems were proved to solve the problem with time-varying cost functions under the assumptions that all local cost functions have identical Hessians and nonidentical constant Hessians. We showed that the state of each agent converges to the corresponding optimal trajectory as the Lagrange multipliers reach consensus. The performance of the distributed algorithms has been demonstrated in simulations.

Appendix A
Proof of Lemma 1

Define $e_i(t) = r_i(t) - \bar{r}(t)$, where $\bar{r}(t) \triangleq \frac{1}{N}\sum_{j=1}^{N} r_j(t)$. Motivated by [20], we consider the Lyapunov function candidate $V = \frac{1}{2}\sum_{i=1}^{N} e_i^2(t)$. We then obtain the time derivative of V along (4) and (5) as

$$\dot{V} = \sum_{i=1}^{N} e_i(t)\dot{e}_i(t) = \sum_{i=1}^{N} e_i(t)\left(\dot{r}_i(t) - \dot{\bar{r}}(t)\right).$$

Since the graph $G(\mathcal{A})$ is undirected, we have $\dot{\bar{r}}(t) = \frac{1}{N}\sum_{j=1}^{N}\dot{r}_j = \frac{1}{N}\sum_{j=1}^{N}(-\nu r_j(t) + d_j(t))$. We then have

$$\dot{V} = \sum_{i=1}^{N} e_i(t)\left(-\nu\left(r_i(t) - \frac{1}{N}\sum_{j=1}^{N} r_j(t)\right) + d_i(t)\right.$$
$$\left. - \frac{1}{N}\sum_{j=1}^{N} d_j(t) + \eta\sum_{j\in N_i}\operatorname{sign}\left(r_j(t) - r_i(t)\right)\right).$$

Note that $\sum_{i=1}^{N} e_i(t) = 0$. We then have $\sum_{i=1}^{N} e_i(t)(-\frac{1}{N}\sum_{j=1}^{N} d_j(t)) = 0$. Because $G(\mathcal{A})$ is undirected, it follows from [21, Lemma 6.1] that $\sum_{i=1}^{N} e_i(t)\sum_{j\in N_i}\text{sign}(r_j(t) - r_i(t)) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j\in N_i}(e_j(t) - e_i(t))\text{sign}(e_j(t) - e_i(t))$. Hence, we have $\dot{V} = -\nu\sum_{i=1}^{N} e_i^2(t) + \sum_{i=1}^{N} e_i(t)d_i(t) - \frac{1}{2}\eta\sum_{i=1}^{N}\sum_{j\in N_i}|e_j(t) - e_i(t)|$. Define $e(t) = [e_1(t), ..., e_N(t)]^T$. We have $||e(t)||_1 \leq \sqrt{N}||e(t)||_2$, i.e., $\sum_{i=1}^{N}|e_i(t)| \leq \sqrt{N}(\sum_{i=1}^{N} e_i^2(t))^{\frac{1}{2}}$. Define $\epsilon(t) = [\epsilon_{11}(t), ..., \epsilon_{1N}(t), ..., \epsilon_{N1}(t), ..., \epsilon_{NN}(t)]^T$, where $\epsilon_{ij}(t) = a_{ij}|e_j(t) - e_i(t)|$ for $i, j \in \mathcal{I}$. We have $||\epsilon(t)||_2 \leq ||\epsilon(t)||_1$, i.e., $(\sum_{i=1}^{N}\sum_{j\in N_i}|e_j(t) - e_i(t)|^2)^{\frac{1}{2}} \leq \sum_{i=1}^{N}\sum_{j\in N_i}|e_j(t) - e_i(t)|$. By leveraging the property of (1) and (2) for the connected undirected network $G(\mathcal{A})$ and $\sum_{i=1}^{N} e_i(t)d_i(t) < D\sum_{i=1}^{N}|e_i(t)|$, we obtain that

$$\begin{aligned}\dot{V} &< -\nu\sum_{i=1}^{N} e_i^2(t) + \sqrt{N}D\left(\sum_{i=1}^{N} e_i^2(t)\right)^{\frac{1}{2}}\\ &\quad - \frac{1}{2}\eta\left(\sum_{i=1}^{N}\sum_{j\in N_i}|e_j(t) - e_i(t)|^2\right)^{\frac{1}{2}}\\ &\leq -2\nu V - (\eta\sqrt{\lambda_2} - D\sqrt{2N})V^{\frac{1}{2}} \leq 0.\end{aligned}$$

Then, it follows that the variables r_i for $i \in \mathcal{I}$ with dynamics (4) will reach consensus as $t \to \infty$.

Appendix B
Proof of Lemma 2

Define $P(t) = [p_1(t), ..., p_N(t)]^T$, $Q(t) = [q_1(t), ..., q_N(t)]^T$, and $\Phi(t) = [\phi_1(t), ..., \phi_N(t)]^T$. We then write the system (6) in the following vector form:

$$\dot{P}(t) = \sigma(t)P(t) - \zeta LQ(t) - \mu B\text{sign}\left(B^T Q(t)\right), \quad (35)$$

$$Q(t) = \Phi(t) + P(t). \quad (36)$$

Let $\widetilde{Q}(t) = \Pi Q(t)$ and $Y(t) = B^T\widetilde{Q}(t)$. We have $Y(t) = B^T Q(t)$ due to the fact that $B^T\Pi = B^T$. Note that $L = BB^T$. It follows from (35) and (36) that

$$\begin{aligned}\dot{\widetilde{Q}}(t) &= \Pi\dot{Q}(t) = \Pi\left(\dot{\Phi}(t) + \dot{P}(t)\right)\\ &= \Pi\left(\dot{\Phi}(t) - \sigma(t)\Phi(t)\right) + \sigma(t)\widetilde{Q}(t) - \zeta BB^T Q(t)\\ &\quad - \mu B\text{sign}\left(B^T Q(t)\right), \quad (37)\end{aligned}$$

$$\begin{aligned}\dot{Y}(t) &= B^T\dot{\widetilde{Q}}(t)\\ &= B^T\left(\dot{\Phi}(t) - \sigma(t)\Phi(t)\right) + \sigma(t)Y(t) - \zeta B^T BY(t)\\ &\quad - \mu B^T B\text{sign}\left(Y(t)\right). \quad (38)\end{aligned}$$

Motivated by [17], we consider the Lyapunov function candidate $V = \frac{1}{2}Y(t)^T(B^T B)^+ Y(t)$. We then obtain the time derivative of V along (37) and (38) as

$$\begin{aligned}\dot{V} &= Y(t)^T(B^T B)^+\dot{Y}(t)\\ &= Y(t)^T(B^T B)^+ B^T\left(\dot{\Phi}(t) - \sigma(t)\Phi(t)\right)\\ &\quad + \sigma(t)Y(t)^T(B^T B)^+ Y(t)\\ &\quad - \zeta Y(t)^T(B^T B)^+ B^T BY(t)\\ &\quad - \mu Y(t)^T(B^T B)^+ B^T B\text{sign}\left(Y(t)\right). \quad (39)\end{aligned}$$

Replace $Y(t)$ with $B^T Q(t)$ in (39), invoke the property (3) that $\Pi = B(B^T B)^+ B^T$ for the fixed connected undirected network $G(\mathcal{A})$, and we then obtain

$$\begin{aligned}\dot{V} &= Y(t)^T(B^T B)^+ B^T\left(\dot{\Phi}(t) - \sigma(t)\Phi(t)\right)\\ &\quad + \sigma(t)Y(t)^T(B^T B)^+ Y(t)\\ &\quad - \zeta Y(t)^T Y(t) - \mu Y(t)^T\text{sign}\left(Y(t)\right)\\ &\leq ||Y(t)||_1\left|\left|(B^T B)^+\right|\right|_\infty\left|\left|B^T\left(\dot{\Phi}(t) - \sigma(t)\Phi(t)\right)\right|\right|_\infty\\ &\quad + \sigma_{\max}\cdot\lambda_{\max}\{(B^T B)^+\}Y(t)^T Y(t)\\ &\quad - \zeta Y(t)^T Y(t) - \mu||Y(t)||_1.\end{aligned}$$

It follows from (7) that $\sup_{t\geq 0}\ ||B^T\Phi(t)||_\infty \leq \phi_{\max}$ and $\sup_{t\geq 0}\ ||B^T\dot{\Phi}(t)||_\infty \leq \varphi_{\max}$, where $\Phi(t) = [\phi_1(t), ..., \phi_N(t)]^T$. Therefore, we have $\dot{V} \leq -||Y(t)||_1 \leq -||Y(t)||_2$, which guarantees the boundness of $Y(t)$. We then obtain that $\dot{Y}(t)$ is bounded according to (38) with (7). In addition, $Y(t)$ is square-integrable since $\int_{t_0}^{\infty}(Y(t)^T Y(t))^{1/2}dt < \infty$. According to Barbalat's lemma, we have $\lim_{t\to\infty}Y(t) = \mathbf{0}_\ell$, where ℓ is defined as the cardinality of set $\mathcal{E}$ in Section II-A, which implies that the entries of $\widetilde{Q}(t)$ will reach consensus as $t \to \infty$ due to the fact that $Y(t) = B^T\widetilde{Q}(t)$. Note that $\mathbf{1}_n^T\widetilde{Q}(t) = 0$ always hods. We then obtain $\lim_{t\to\infty}\widetilde{Q}(t) = \mathbf{0}_N$ and hence the consensus of variables q_i for $i \in \mathcal{I}$ is reached.

References

[1] R. Mudumbai, S. Dasgupta, and B. B. Cho, "Distributed control for optimal economic dispatch of a network of heterogeneous power generators," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1750–1760, Nov. 2012.
[2] G. Binetti, A. Davoudi, F. L. Lewis, D. Naso, and B. Turchiano, "Distributed consensus-based economic dispatch with transmission losses," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1711–1720, Jul. 2014.
[3] 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, Jan. 2011.
[4] 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.
[5] 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.
[6] 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, Dec. 2016.
[7] S. S. Kia, "Distributed optimal in-network resource allocation algorithm design via a control theoretic approach," *Syst. Control Lett.*, vol. 107, pp. 49–57, Sep. 2017.
[8] D. Wang, Z. Wang, C. Wen, and W. Wang, "Second-order continuous-time algorithm for optimal resource allocation in power systems," *IEEE Trans. Ind. Inform.*, vol. 15, no. 2, pp. 626–637, Feb. 2019.
[9] Y. Zhu, W. Yu, G. Wen, G. Chen, and W. Ren, "Continuous-time distributed subgradient algorithm for convex optimization with general constraints," *IEEE Trans. Autom. Control*, vol. 64, no. 4, pp. 1694–1701, Apr. 2019.

[10] 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.

[11] A. Simonetto, A. Koppel, A. Mokhtari, G. Leus, and A. Ribeiro, "Decentralized prediction-correction methods for networked time-varying convex optimization," *IEEE Trans. Autom. Control*, vol. 62, no. 11, pp. 5724–5738, Nov. 2017.

[12] A. Simonetto, "Dual prediction-correction methods for linearly constrained time-varying convex programs," *IEEE Trans. Autom. Control*, vol. 64, no. 8, pp. 3355–3361, Aug. 2019.

[13] M. Fazlyab, S. Paternain, V. M. Preciado, and A. Ribeiro, "Prediction-correction interior-point method for time-varying convex optimization," *IEEE Trans. Autom. Control*, vol. 63, no. 7, pp. 1973–1986, Jul. 2018.

[14] S. Rahili and W. Ren, "Distributed continuous-time convex optimization with time-varying cost functions," *IEEE Trans. Autom. Control*, vol. 62, no. 4, pp. 1590–1605, Apr. 2017.

[15] L. Bai, C. Sun, Z. Feng, and G. Hu, "Distributed continuous-time resource allocation with time-varying resources under quadratic cost functions," in *Proc. IEEE Conf. Decis. Control*, Miami Beach, FL, USA, Dec. 2018, pp. 823–828.

[16] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Apr. 2004.

[17] J. George and R. Freeman, "Robust dynamic average consensus algorithms," *IEEE Trans. Autom. Control*, vol. 64, no. 11, pp. 4615–4622, Nov. 2019.

[18] F. Chen, Y. Cao, and W. Ren, "Distributed average tracking of multiple time-varying reference signals with bounded derivatives," *IEEE Trans. Autom. Control*, vol. 57, no. 12, pp. 3169–3174, Dec. 2012.

[19] Y. Cao and W. Ren, "Finite-time consensus for multi-agent networks with unknown inherent nonlinear dynamics," *Automatica*, vol. 50, no. 10, pp. 2648–2656, Oct. 2014.

[20] H. Hong, W. Yu, G. Wen, and X. Yu, "Distributed robust fixed-time consensus for nonlinear and disturbed multiagent systems," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 47, no. 7, pp. 1464–1473, Jul. 2017.

[21] W. Ren and Y. Cao, *Distributed Coordination of Multi-Agent Networks: Emergent Problems, Models, and Issues*. London, U.K.: Springer, 2010.

Bo Wang received the B.S. degree in automation from the School of Automation, Beijing Institute of Technology, Beijing, China, in 2014. He is currently working toward the Ph.D. degree in control theory and control engineering with the School of Automation.

From 2018 to 2019, he was a Visiting Scholar at the University of California at Riverside, Riverside, CA, USA. His research focuses on robotics, multiagent system control, and distributed optimization.

Shan Sun received the B.S. and M.Sc. degrees in guidance, navigation and control from the School of Astronautics, Beihang University, Beijing, China, in 2014 and 2017, respectively. She is currently working toward the Ph.D. degree in electrical and computer engineering with the University of California, Riverside, CA, USA.

Her research focuses on distributed optimization and multiagent systems.

Wei Ren (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Brigham Young University, Provo, UT, USA, in 2004.

He is currently a Professor with the Department of Electrical and Computer Engineering, University of California, Riverside, CA, USA. Prior to joining UC Riverside, he was a Faculty Member at Utah State University and a Postdoctoral Research Associate at the University of Maryland, College Park, MD, USA. His research focuses on distributed control of multiagent systems and autonomous control of unmanned vehicles. Dr. Ren was a recipient of the IEEE Control Systems Society Antonio Ruberti Young Researcher Prize in 2017 and the National Science Foundation CAREER Award in 2008. He is an IEEE Control Systems Society Distinguished Lecturer.