Title
Distributed optimal in-network resource allocation algorithm design via a control theoretic approach

Author
Kia, Solmaz S

Publication Date
2017-09-01

DOI
10.1016/j.sysconle.2017.07.012

Peer reviewed

Systems & Control Letters 107 (2017) 49–57

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Distributed optimal in-network resource allocation algorithm design via a control theoretic approach

Solmaz S. Kia

Department of Mechanical and Aerospace Engineering, University of California, Irvine 4200 Engineering Gateway, Irvine, CA 92697, United States

A R T I C L E I N F O

Article history:
Received 30 July 2016
Received in revised form 21 June 2017
Accepted 14 July 2017
Available online 10 August 2017

Keywords:
Cooperative control
Distributed convex optimization
Distributed optimal resource allocation
Strongly connected and weight-balanced digraphs

A B S T R A C T

In this paper, we consider an in-network optimal resource allocation problem with multiple demand equations. We propose a novel distributed continuous-time algorithm that solves the problem over strongly connected and weight-balanced digraph network topologies when the local cost functions are strongly convex. We also discuss the extension of our convergence guarantees to dynamically changing topologies. Finally, we show that if the network is an undirected connected graph, we can guarantee stability and convergence of our algorithm for problems involving local convex functions. This convergence guarantee is to a point in the set of minimizers of our optimal resource allocation problem. The design and analysis of our algorithm are carried out using a control theoretic approach. We demonstrate our results through a numerical example.

1. Introduction

This paper considers the problem of designing a distributed algorithm for an optimal resource allocation problem subject to a set of affine equality constraints over a network of N agents with communication and processing capabilities. In particular, each agent $i \in \{1, \dots, N\}$ has a convex and differentiable local cost function $f^i : \mathbb{R} \to \mathbb{R}$. These agents are meeting some demands $\mathsf{b}_j \in \mathbb{R}, j \in \{1, \dots, p\}$, through weighted contributions, in a way that the total cost $f(\mathbf{x}) = \Sigma_{i=1}^{N} f^i(x^i)$ is at its minimum. In other words, each agent $i \in \{1, \dots, N\}$ seeks $\mathsf{x}_i^\star$, the ith element of $\mathbf{x}^\star$ given by

$$\mathbf{x}^\star = \arg\min_{\mathbf{x}\in\mathbb{R}^N} \sum_{i=1}^{N} f^i(x^i), \quad \text{subject to} \tag{1a}$$

$$\omega_j^1 x^1 + \cdots + \omega_j^N x^N - \mathsf{b}_j = 0, \quad j \in \{1, \dots, p\}, \tag{1b}$$

where, $\omega_j^i \in \mathbb{R}, i \in \{1, \dots, N\}$, is the weight on the contribution of agent i to demand equation $j \in \{1, \dots, p\}$. The weights $\{\omega_j^i\}_{j=1}^p$ of each agent $i \in \{1, \dots, N\}$ are known to that agent. The aforementioned problem appears in many optimal decision making tasks such as economic dispatch over power networks [1,2], optimal routing [3] and economic systems [4].

Literature review: Our paper is related to a large recent literature on distributed algorithm design for solving a multi-agent optimization problem where the global cost function is a sum of local convex functions, each representing a private local cost only available to a single agent, subject to some convex constraints. Some of the recent literature on distributed optimization algorithm design includes distributed algorithms implemented both in discrete-time [5–9] and continuous-time [10–14]. Although some of these algorithms can solve the optimal resource allocation problem (1), they require each agent to keep and evolve a copy of the global decision variable of the problem which is of order N, where N is the size of network. Such a requirement is costly and unnecessary for problem (1), as the agents only need to obtain their own respective component of the global decision variable. Distributed optimization algorithms that specifically target the optimal resource allocation problem (1) are presented in [15] in discrete-time form, and [2,16] in continuous-time form. These algorithms require the agents to keep and evolve only their respective component of the global decision variable. However, these algorithms all can solve the optimal allocation problem (1) subject to single unweighted demand equation, i.e., $\omega_1^i = 1$, $i \in \{1, \dots, N\}$ and $p = 1$ in (1b). Also, these algorithms require the agents to transmit the gradient of their local cost functions to their neighbors, which makes these algorithms less favorable for privacy-sensitive applications. The composition of our algorithm is inspired by the multi-time scale singularly perturbed systems in control theory (cf. [17]). Singularly perturbed distributed algorithms are used in [12] for unconstrained in-network convex optimization, and in [18] for dynamic consensus problem over networked systems.

Statement of contributions: We propose a novel continuous-time distributed algorithm to solve the optimal resource allocation problem (1) over networked systems. We show that our algorithm

E-mail address: solmaz@uci.edu.

http://dx.doi.org/10.1016/j.sysconle.2017.07.012

converges over strongly connected and weight-balanced digraphs if the local cost functions are strongly convex. Such guarantees also hold for time-varying strongly connected and weight-balanced digraphs with piecewise constant adjacency matrices if the gradients of all local cost functions are globally Lipschitz. When the communication graph is an undirected connected graph, the convergence is guaranteed for convex local cost functions, as well. Our convergence guarantee is to a point in the optimizer set. The composition of our algorithm is inspired by the singular perturbation systems in control theory. The idea behind this composition is that an average consensus algorithm creates a local copy of the left hand side of the equality constraint (1b) at each agent. This way, every agent can create a local copy of its respective part in a centralized saddle-point dynamical solver used in the literature to solve the optimization problem (1). In the resulted algorithm, each agent is only required to keep a copy of its own local decision variable. Also, agents are not required to share the gradient of their local cost functions with their neighbors. We use Lyapunov and invariant set analysis to study the convergence and stability of our proposed algorithm. A preliminary work related to our work has appeared in [19].

2. Preliminaries

This section presents our notations, definitions, a review of relevant algebraic graph theory, and the average consensus algorithm of [20].

2.1. Notations

Let $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, and $\mathbb{R}_{>0}$, respectively, be the set of real, non-negative real, and positive real numbers. We let $\mathbf{1}_n$ (resp. $\mathbf{0}_n$) denote the vector of n ones (resp. n zeros), and denote by $\mathbf{I}_n$ the $n \times n$ identity matrix. When clear from the context, we do not specify the matrix dimensions. We denote the standard Euclidean norm of vector $\mathbf{x} \in \mathbb{R}^n$ by $\|\mathbf{x}\| = \sqrt{\mathbf{x}^\top \mathbf{x}}$. We denote the induced 1-norm, ∞-norm and spectral norm of a matrix $\mathbf{A} \in \mathbb{R}^{n\times m}$ by, respectively, $\|\mathbf{A}\|_1$, $\|\mathbf{A}\|_\infty$ and $\|\mathbf{A}\|$. In a network of N agents, to distinguish and emphasize that a variable is local to an agent $i \in \{1,\dots,N\}$, we use superscripts, e.g., $f^i(x^i)$ is the local function of agent i evaluated at its own local state x^i. Moreover, if $\mathbf{p}^i \in \mathbb{R}^d$ is a variable of agent $i \in \{1,\dots,N\}$, the aggregated $\mathbf{p}^i$'s of the network is the vector $\mathbf{p} = [\,\mathbf{p}^{1\top}, \cdots, \mathbf{p}^{N\top}]^\top \in (\mathbb{R}^d)^N$.

A differentiable function $f : \mathbb{R}^d \to \mathbb{R}$ is convex (resp. strictly convex) over a convex set $C \subseteq \mathbb{R}^d$ iff $(\mathbf{z}-\mathbf{x})^\top(\nabla f(\mathbf{z}) - \nabla f(\mathbf{x})) \geq 0$ (resp. $(\mathbf{z}-\mathbf{x})^\top(\nabla f(\mathbf{z}) - \nabla f(\mathbf{x})) > 0$ whenever $\mathbf{x} \neq \mathbf{z}$) for all $\mathbf{x}, \mathbf{z} \in C$, and it is *m-strongly convex* ($m \in \mathbb{R}_{>0}$) iff $(\mathbf{z}-\mathbf{x})^\top(\nabla f(\mathbf{z}) - \nabla f(\mathbf{x})) \geq m\|\mathbf{z}-\mathbf{x}\|^2$, for all $\mathbf{x}, \mathbf{z} \in C$. A function $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz with constant $M \in \mathbb{R}_{>0}$, or simply *M-Lipschitz*, over a set $C \subseteq \mathbb{R}^d$ iff $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq M\,\|\mathbf{x}-\mathbf{y}\|$, for $\mathbf{x}, \mathbf{y} \in C$. Function f is *globally Lipschitz* if it is M-Lipschitz over $\mathbb{R}^d$. Moreover, it is *locally Lipschitz* on $\mathbb{R}^d$ if for every point $\mathbf{x} \in \mathbb{R}^d$ there exists a $M_{\mathbf{x}} \in \mathbb{R}_{>0}$ such that $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq M_{\mathbf{x}}\,\|\mathbf{x}-\mathbf{y}\|$ for all $\mathbf{y}$ in an open and connected neighborhood of $\mathbf{x}$.

2.2. Graph theory

We briefly review basic concepts from algebraic graph theory following [21]. A *digraph*, is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1,\dots,N\}$ is the *node set* and $\mathcal{E} \subseteq \mathcal{V}\times\mathcal{V}$ is the *edge set*. An edge from i to j, denoted by (i,j), means that agent j can send information to agent i. For an edge $(i,j) \in \mathcal{E}$, i is called an *in-neighbor* of j and j is called an *out-neighbor* of i. A graph is *undirected* if $(i,j) \in \mathcal{E}$ anytime $(j,i) \in \mathcal{E}$. A *directed path* is a sequence of nodes connected by edges. A digraph is *strongly connected* if for every pair of nodes there is a directed path connecting them. A *weighted digraph* is a triplet $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A})$, where $(\mathcal{V}, \mathcal{E})$ is a digraph and $\mathbf{A} \in \mathbb{R}^{N\times N}$ is a weighted *adjacency* matrix such that $\mathsf{a}_{ij} > 0$ if $(i,j) \in \mathcal{E}$ and $\mathsf{a}_{ij} = 0$, otherwise. A weighted digraph is *undirected* if $\mathsf{a}_{ij} = \mathsf{a}_{ji}$ for all $i,j \in \mathcal{V}$. A *connected graph* is a strongly connected and undirected graph. The *weighted in-* and *out-degrees* of a node i are, respectively, $\mathsf{d}^i_{\text{in}} = \Sigma_{j=1}^N \mathsf{a}_{ji}$ and $\mathsf{d}^i_{\text{out}} = \Sigma_{j=1}^N \mathsf{a}_{ij}$. A digraph is *weight-balanced* if at each node $i \in \mathcal{V}$ the weighted out-degree and weighted in-degree coincide. Any connected graph is weight-balanced. The *(out-) Laplacian* matrix is $\mathbf{L} = \mathbf{D}^{\text{out}} - \mathbf{A}$, where $\mathbf{D}^{\text{out}} = \text{Diag}(\mathsf{d}^1_{\text{out}}, \cdots, \mathsf{d}^N_{\text{out}}) \in \mathbb{R}^{N\times N}$. Note that $\mathbf{L}\mathbf{1}_N = \mathbf{0}$. A digraph is weight-balanced iff $\mathbf{1}_N^T\mathbf{L} = \mathbf{0}$. We let $\{\lambda_i\}_{i=1}^N$ and $\{\hat{\lambda}_i\}_{i=1}^N$, respectively, be the set of eigenvalues of $\mathbf{L}$ and $\text{Sym}(\mathbf{L}) = (\mathbf{L} + \mathbf{L}^\top)/2$. Based on the structure of $\mathbf{L}$, at least one of the eigenvalues of $\mathbf{L}$ is zero ($\lambda_1 = 0$) and the rest of them have nonnegative real parts. For a strongly connected and weight-balanced digraph, zero is a simple eigenvalue of both $\mathbf{L}$ and $\text{Sym}(\mathbf{L})$. Moreover, we have

$$0 < \hat{\lambda}_2\mathbf{I} \leq \mathbf{R}^\top \text{Sym}(\mathbf{L})\mathbf{R} \leq \hat{\lambda}_N\mathbf{I}, \tag{2}$$

where $\hat{\lambda}_2$ and, $\hat{\lambda}_N$ are, respectively, the smallest non-zero eigenvalue and maximum eigenvalue of $\text{Sym}(\mathbf{L})$. Here, $\mathbf{R} \in \mathbb{R}^{N\times(N-1)}$ along with $\mathbf{r} \in \mathbb{R}^N$ satisfies

$$\mathbf{r} = \frac{1}{\sqrt{N}}\mathbf{1}_N,\ \ \mathbf{r}^\top\mathbf{R} = \mathbf{0},\ \ \mathbf{R}^\top\mathbf{R} = \mathbf{I}_{N-1},\ \ \mathbf{R}\mathbf{R}^\top = \mathbf{I}_N - \mathbf{r}\mathbf{r}^\top. \tag{3}$$

For connected graphs $\hat{\lambda}_i = \lambda_i$, $i \in \mathcal{V}$, therefore, $0 < \lambda_2\mathbf{I} \leq \mathbf{R}^\top\mathbf{L}\mathbf{R} \leq \lambda_N\mathbf{I}$.

2.3. Average consensus algorithm

Let $\mathcal{G}$ be a strongly connected and weight-balanced digraph of N agents. Assume each node $i \in \mathcal{V}$ has access to a static reference input $\mathsf{r}^i \in \mathbb{R}^p$. [20] shows that for $\beta \in \mathbb{R}_{>0}$, if each agent $i \in \mathcal{V}$, implements

$$\dot{\mathbf{v}}^i = \beta\sum_{j=1}^N \mathsf{a}_{ij}(\mathbf{y}^i - \mathbf{y}^j), \tag{4}$$

$$\dot{\mathbf{y}}^i = -(\mathbf{y}^i - \mathsf{r}^i) - \beta\sum_{j=1}^N \mathsf{a}_{ij}(\mathbf{y}^i - \mathbf{y}^j) - \mathbf{v}^i,$$

starting at $\mathbf{y}^i(0), \mathbf{v}^i(0) \in \mathbb{R}^p$, $\sum_{j=1}^N \mathbf{v}^j(0) = \mathbf{0}$, then as $t \to \infty$, its state $\mathbf{y}^i$ converges to $\frac{1}{N}\sum_{j=1}^N \mathsf{r}^j$ exponentially fast.

3. Problem statement

We consider the optimal resource allocation problem (1) over a network of N agents interacting over a digraph $\mathcal{G}$, and under the following assumption.

Assumption 1. Matrix $\Omega = [\boldsymbol{\omega}^1, \cdots, \boldsymbol{\omega}^N]$, where $\boldsymbol{\omega}^i = [\omega_1^i, \cdots, \omega_p^i]^\top$, $i \in \mathcal{V}$, is full row rank. Moreover, the optimization problem (1) has a finite optimum $f^\star = f(\mathbf{x}^\star)$. Finally, ∇f^i, $i \in \mathcal{V}$, is locally Lipschitz on $\mathbb{R}$.

The first part of Assumption 1 ensures that the feasible set of optimization problem (1) is non-empty and the problem has a finite minimizer in the feasible set. Local Lipschitzness of ∇f^i, $i \in \mathcal{V}$, guarantees existence and uniqueness of the solutions of the dynamical solvers that we study in this paper for problem (1) (cf. [22, Theorem 3.3])—these solvers use ∇f^i, $i \in \mathcal{V}$.

The Karush–Kuhn–Tucker (KKT) conditions give a set of necessary and sufficient conditions to characterize the solution set of the convex optimization problem (1) as follows (cf. [23] for proof).

Lemma 3.1 (*Solution Set of (1)*). *Let $f^i:\mathbb{R}\to\mathbb{R}$, $i\in\mathcal{V}$, in constrained optimization problem (1) be a differentiable and convex function on $\mathbb{R}$ and Ω be full row rank. A point $\mathbf{x}^\star\in\mathbb{R}^N$ is a solution of (1) iff there exists a $\boldsymbol{\mu}^\star\in\mathbb{R}^p$, such that*

$$\nabla f^i(\mathsf{x}_i^\star)+\sum_{j=1}^{p}\omega_j^i\,\mu_j^\star=0,\quad i\in\mathcal{V}, \tag{5a}$$

$$\omega_j^1\mathsf{x}_1^\star+\cdots+\omega_j^N\mathsf{x}_N^\star-\mathsf{b}_j=0,\quad j\in\{1,\dots,p\}. \tag{5b}$$

When the local costs are all strictly convex, KKT equation (5) has a unique solution. □

We represent the set of points satisfying the KKT condition (5) by

$$\mathcal{O}^\star=\Big\{(\boldsymbol{\mu},\mathbf{x})\in\mathbb{R}^p\times\mathbb{R}^N\,\Big|\,\nabla f^i(x^i)+\boldsymbol{\mu}^\top\boldsymbol{\omega}^i=0,\\ i\in\mathcal{V},\ \Omega\,\mathbf{x}-\mathbf{b}=0\Big\}, \tag{6}$$

where $\mathbf{b}=\begin{bmatrix}\mathsf{b}_1 & \cdots & \mathsf{b}_p\end{bmatrix}^\top$ is the aggregate demand vector.

To define our objective, consider the integrator dynamics

$$\dot{x}^i=\psi^i(t), \tag{7}$$

at each agent $i\in\mathcal{V}$. Our aim is to design the driving command $\psi^i:\mathbb{R}_{\geq0}\to\mathbb{R}$, $i\in\mathcal{V}$, such that (a) $t\mapsto\mathbf{x}(t)$ converges asymptotically to a minimizer of the optimization problem (1), (b) the structure of and the inputs to the dynamics that generates ψ^i depend only on the local variables of agent i and information it receives from its out-neighbors.

Our proposed distributed algorithm is inspired by the saddle-point dynamics

$$\dot{\boldsymbol{\mu}}=\boldsymbol{\omega}^1x^1+\cdots+\boldsymbol{\omega}^Nx^N-\mathbf{b},\qquad \boldsymbol{\mu}(0)\in\mathbb{R}^p, \tag{8a}$$

$$\dot{x}^i=-\nabla f^i(x^i)-\boldsymbol{\omega}^{i\,\top}\boldsymbol{\mu},\quad i\in\mathcal{V},\quad x^i(0)\in\mathbb{R}, \tag{8b}$$

in which when the local cost functions are strictly convex, every trajectory $t\mapsto(\boldsymbol{\mu}(t),\mathbf{x}(t))$ converges to $\mathcal{O}^\star$ (see [24]). Our distributed solver uses the average consensus algorithm (4) to distribute the coupled equation (8a).

4. Distributed continuous-time algorithm for optimal resource allocation

In this section, we present our distributed solver of the in-network optimal resource allocation problem (1). We show that when the communication topology is a strongly connected and weight-balanced digraph, this algorithm converges to the solution of (1) if the local cost functions are strongly convex. Next, we show that when the communication topology is a connected graph, the convergence guarantees extend to the class of optimal resource allocation problems with convex local cost functions. We close the section with discussion on convergence guarantees over dynamically changing topologies.

The following novel continuous-time distributed algorithm is our solution to the optimization problem (1),

$$\dot{\mathbf{v}}^i=\beta\sum_{j=1}^{N}\mathsf{a}_{ij}(\mathbf{y}^i-\mathbf{y}^j), \tag{9a}$$

$$\dot{\mathbf{y}}^i=-(\mathbf{y}^i-(\boldsymbol{\omega}^ix^i+\boldsymbol{\mu}^i-\tilde{\mathbf{b}}^i))-\beta\sum_{j=1}^{N}\mathsf{a}_{ij}(\mathbf{y}^i-\mathbf{y}^j)-\mathbf{v}^i, \tag{9b}$$

$$\dot{\boldsymbol{\mu}}^i=-\boldsymbol{\mu}^i+\mathbf{y}^i, \tag{9c}$$

$$\dot{x}^i=-\nabla f^i(x^i)-\boldsymbol{\omega}^{i\,\top}\mathbf{y}^i, \tag{9d}$$

for $i\in\mathcal{V}$, where $\sum_{i=1}^{N}\tilde{\mathbf{b}}^i=\mathbf{b}$ and $\beta\in\mathbb{R}_{>0}$. Algorithm (9) is a distributed algorithm which generates the driving command ψ^i in (7) according to $\psi^i(t)=-\nabla f^i(x^i)-\boldsymbol{\omega}^{i\,\top}\mathbf{y}^i$, the output of dynamics (9a)–(9c) whose inputs are $(\tilde{\mathbf{b}}^i,\{\mathbf{y}^j\}_{j\in\mathcal{N}_{\text{out}}^i},x^i)$. Here, $\mathcal{N}_{\text{out}}^i$ is the set of the out-neighbors of agent $i\in\mathcal{V}$. We assume that every agent $i\in\mathcal{V}$ knows its own weights $\{\omega_j^i\}_{j=1}^p$ and also $\tilde{\mathbf{b}}^i$ (possible cases for $\tilde{\mathbf{b}}^i$ are (a) $\tilde{\mathbf{b}}^i=\mathbf{b}/N$, $i\in\mathcal{V}$, i.e., every agent knows the demand vector and the size of the network, (b) $\tilde{\mathbf{b}}^1=\mathbf{b}$ and $\tilde{\mathbf{b}}^i=\mathbf{0}_p$, $i\in\{2,\cdots,N\}$, i.e., only, without loss of generality, agent 1 knows the demand vector). It is worth noting that the dimension of the local variables $(\mathbf{v}^i,\mathbf{y}^i,\boldsymbol{\mu}^i,x^i)\in\mathbb{R}^{3p+1}$ of each agent $i\in\mathcal{V}$ in (9) is regardless of the size of the network, i.e., the solution is achieved without requiring each agent to keep a copy of the entire minimizer vector which is of dimension N.

Let $\Omega_{\text{D}}=\text{Diag}(\boldsymbol{\omega}^1,\cdots,\boldsymbol{\omega}^N)$ and $\mathbf{L}=\mathbf{L}\otimes\mathbf{I}_p$. Then, in the network aggregated variables $\mathbf{v},\mathbf{y},\boldsymbol{\mu}\in\mathbb{R}^{pN}$ and $\mathbf{x}\in\mathbb{R}^N$, the algorithm reads as (recall $f(\mathbf{x})=\Sigma_{i=1}^{N}f^i(x^i)$, therefore $\nabla f(\mathbf{x})=[\nabla f^1(x^1),\cdots,\nabla f^N(x^N)]^\top$)

$$\dot{\mathbf{v}}=\beta\,\mathbf{L}\,\mathbf{y}, \tag{10a}$$

$$\dot{\mathbf{y}}=-(\mathbf{y}-(\Omega_{\text{D}}\,\mathbf{x}+\boldsymbol{\mu}-\tilde{\mathbf{B}}))-\beta\,\mathbf{L}\,\mathbf{y}-\mathbf{v}, \tag{10b}$$

$$\dot{\boldsymbol{\mu}}=-\boldsymbol{\mu}+\mathbf{y}, \tag{10c}$$

$$\dot{\mathbf{x}}=-\nabla f(\mathbf{x})-\Omega_{\text{D}}^\top\,\mathbf{y}, \tag{10d}$$

where $\tilde{\mathbf{B}}=\left[\tilde{\mathbf{b}}^{1\top},\cdots,\tilde{\mathbf{b}}^{N\top}\right]^\top$. The composition of algorithm (9) is inspired by the central solver (8) and use of multi-time scale analysis approach in the singular perturbation theory. Note that ((9a), (9b)) has the form of the average consensus algorithm (4) with $\mathbf{r}^i=(\boldsymbol{\omega}^ix^i+\boldsymbol{\mu}^i-\tilde{\mathbf{b}}^i)$. Now, assume that (9a), (9b) run in a faster time scale than the rest of the dynamics. Thus, with appropriate initialization, in this fast dynamics for $i\in\mathcal{V}$, every $\mathbf{y}^i$ converges to the common value $\frac{1}{N}\sum_{j=1}^{N}(\boldsymbol{\omega}^jx^j+\boldsymbol{\mu}^j-\tilde{\mathbf{b}}^j)$. Substituting this value for $\mathbf{y}^i$ in (10c) and left multiplying (10c) by $(\mathbf{1}_N^\top\otimes\mathbf{I}_p)$ result in $\sum_{i=1}^{N}\dot{\boldsymbol{\mu}}^i=(\boldsymbol{\omega}^1x^1+\cdots+\boldsymbol{\omega}^Nx^N-\mathbf{b})$. If we postulate that (10c) converges in a faster time scale than (10d), we can use $\boldsymbol{\mu}^i\to\mathbf{y}^i$ as $t\to\infty$ to obtain the slow dynamics $\dot{x}^i=-\nabla f^i(x^i)-\boldsymbol{\omega}^{i\,\top}\boldsymbol{\mu}^i$. Next, note that in the fast scale eventually we can write $\dot{\boldsymbol{\mu}}^i=\frac{1}{N}(\boldsymbol{\omega}^1x^1+\cdots+\boldsymbol{\omega}^Nx^N-\mathbf{b})$. Therefore, every agent $i\in\mathcal{V}$ eventually reconstructs locally (almost) a copy of saddle-point dynamics (8) (here right hand side of (8a) is scaled by $1/N$). The preceding discussion sketches the inspiration behind the composition of the algorithm. In the following, we provide a rigorous study of the stability and convergence properties of algorithm (9) using the Lyapunov stability analysis. We start by characterizing the equilibrium points of algorithm (9).

Lemma 4.1 (*Equilibrium Points of Algorithm (9) Over Strongly Connected and Weight-Balanced Digraphs*). *Let $\mathcal{G}$ be a strongly connected and weight-balanced digraph. Assume f^i, $i\in\mathcal{V}$, is convex and differentiable. Then, the set of equilibrium points of (9) is given by*

$$\underline{\mathcal{U}}=\Big\{(\mathbf{v},\mathbf{y},\boldsymbol{\mu},\mathbf{x})\in\mathbb{R}^{pN}\times\mathbb{R}^{pN}\times\mathbb{R}^{pN}\times\mathbb{R}^N\Big|\mathbf{v}=\Omega_D\,\mathbf{x}-\tilde{\mathbf{B}},\\ \mathbf{y}=\boldsymbol{\mu}=\mathbf{1}_N\otimes\boldsymbol{\theta},\\ \sum_{i=1}^{N}\mathbf{v}^i=\sum_{i=1}^{N}\boldsymbol{\omega}^ix^i-\mathbf{b},\ \nabla f^i(x^i)+\boldsymbol{\omega}^{i\top}\boldsymbol{\theta}=0,\ i\in\mathcal{V},\ \boldsymbol{\theta}\in\mathbb{R}^p\Big\}. \tag{11}$$

Proof. Let $(\underline{\mathbf{v}},\underline{\mathbf{y}},\underline{\boldsymbol{\mu}},\underline{\mathbf{x}})$ be an equilibrium point of (9), i.e.,

$$\mathbf{0}=\beta\,\mathbf{L}\,\underline{\mathbf{y}}, \tag{12a}$$

$$\mathbf{0}=-(\underline{\mathbf{y}}-(\Omega_{\text{D}}\,\underline{\mathbf{x}}+\underline{\boldsymbol{\mu}}-\tilde{\mathbf{B}}))-\beta\,\mathbf{L}\,\underline{\mathbf{y}}-\underline{\mathbf{v}}, \tag{12b}$$

$$\mathbf{0}=-\underline{\boldsymbol{\mu}}+\underline{\mathbf{y}}, \tag{12c}$$

$$\mathbf{0}=-\nabla f(\underline{\mathbf{x}})-\Omega_{\text{D}}^\top\,\underline{\mathbf{y}}. \tag{12d}$$

For the given network topology, the rank of $\mathbf{L}$ is $N-1$ and its null-space is spanned by $\mathbf{1}_N$. Therefore, from (12a), we obtain $\underline{\mathbf{y}} = \mathbf{1}_N \otimes \boldsymbol{\theta}, \boldsymbol{\theta} \in \mathbb{R}^p$. Subsequently, from (12c) and (12d), we have $\underline{\boldsymbol{\mu}} = \underline{\mathbf{y}} = \mathbf{1}_N \otimes \boldsymbol{\theta}$ and $\nabla f^i(\underline{x}^i) + \boldsymbol{\omega}^{i\top}\boldsymbol{\theta} = 0, i \in \mathcal{V}$. Moreover, we can use $\underline{\boldsymbol{\mu}} = \underline{\mathbf{y}}$ together with (12a) and (12b) to establish $\underline{\mathbf{v}} = \Omega_{\mathrm{D}}\,\underline{\mathbf{x}} - \tilde{\mathbf{B}}$. On the other hand, because the network topology is weight-balanced, we have $\mathbf{1}_N^\top\mathbf{L} = \mathbf{0}$. Then, by left multiplying (12b) by $\mathbf{1}_N^\top \otimes \mathbf{I}_p$, we obtain $\sum_{i=1}^N \underline{\mathbf{v}}^i = \sum_{i=1}^N \omega^i \underline{x}^i - \mathbf{b}$, which completes the proof. □

Next, we point out few remarks about the limiting state of algorithm (9) if it converges to a point in its equilibrium set (under the assumptions stated in the statement of Lemma 4.1).

Remark 4.1 (*Dependence of the Limiting Points of Algorithm (9) on Initial Conditions and Their Relation to the Minimizer(s) of Optimization Problem (1)*). First, note that starting from any initial conditions at $t = 0$, if algorithm (9) converges to an equilibrium point in (11), that point depends on value of $\sum_{i=1}^N \mathbf{v}^i(0)$. To see this connection note that over weight-balanced digraphs, because of $\mathbf{1}_N^\top\mathbf{L} = \mathbf{0}$, from (10a) we can write

$$\sum_{i=1}^N \dot{\mathbf{v}}^i = \beta\,(\mathbf{1}_N^\top \otimes \mathbf{I}_p)\mathbf{L}\mathbf{y} = \mathbf{0} \;\Rightarrow\; \sum_{i=1}^N \mathbf{v}^i(t) = \sum_{i=1}^N \mathbf{v}^i(0),\;\; \forall t \in \mathbb{R}_{\geq 0}, \tag{13}$$

which implies $\sum_{i=1}^N \lim_{t\to\infty}\mathbf{v}^i(t) = \sum_{i=1}^N \mathbf{v}^i(0)$. Our second remark is that if algorithm (9) is initialized such that $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}$, then if it converges to $(\underline{\mathbf{v}}, \underline{\mathbf{y}}, \underline{\boldsymbol{\mu}}, \underline{\mathbf{x}}) \in \underline{\mathcal{U}}$, we will have $(\underline{\mathbf{v}}^i, \underline{\mathbf{y}}^i, \underline{\boldsymbol{\mu}}^i, \underline{x}^i) = (\boldsymbol{\omega}^i\,\mathsf{x}_i^\star - \tilde{\mathbf{b}}^i, \boldsymbol{\mu}^\star, \boldsymbol{\mu}^\star, \mathsf{x}_i^\star)$, $i \in \mathcal{V}$ where $(\boldsymbol{\mu}^\star, \mathbf{x}^\star)$ is a KKT point in $\mathcal{O}^\star$ (see (6)). This relationship is the consequence of invoking (10a) to write that $(\underline{\mathbf{v}}, \underline{\mathbf{y}}, \underline{\boldsymbol{\mu}}, \underline{\mathbf{x}}) \in \underline{\mathcal{U}}$ now satisfies

$$\underline{\mathbf{v}}^i = \boldsymbol{\omega}^i\underline{x}^i - \tilde{\mathbf{b}}^i,\;\; \underline{\mathbf{y}}^i = \underline{\boldsymbol{\mu}}^i = \boldsymbol{\theta} \in \mathbb{R}^p, \qquad i \in \mathcal{V} \tag{14a}$$

$$\boldsymbol{\omega}^1\underline{x}^1 + \cdots + \boldsymbol{\omega}^N\underline{x}^N - \mathbf{b} = \sum_{i=1}^N \underline{\mathbf{v}}^i = \mathbf{0},$$
$$\nabla f(\underline{x}^i) + \boldsymbol{\omega}^{i\top}\boldsymbol{\theta} = 0,\;\; i \in \mathcal{V} \tag{14b}$$

which, by comparing to the KKT condition (5), shows $(\boldsymbol{\theta}, \underline{\mathbf{x}}) \in \mathcal{O}^\star$. We close this remark by pointing out that the dependence of limiting states on the initial state is seen in numerous dynamical systems including biomedical [25], chemical kinetics [26] systems and also in network algorithms such as average static Laplacian consensus [27] (for more examples see [28, p. 260]).

The first result below shows that when the communication topology is a strongly connected and weight-balanced digraph and the local cost functions are strongly convex, under an appropriate initialization, the trajectories $t \mapsto x^i(t)$, $i \in \mathcal{V}$, of algorithm (9), as $t \to \infty$, converge to the minimizer of problem (1). To simplify notation, hereafter, given r and R in (3), we define and use

$$\mathbf{T} = [\mathbf{r}\;\; \mathbf{R}],\;\; \mathbf{r} = \mathsf{r} \otimes \mathbf{I}_p,\;\; \mathbf{R} = \mathsf{R} \otimes \mathbf{I}_p,\;\; \mathbf{T} = \mathsf{T} \otimes \mathbf{I}_p. \tag{15}$$

Theorem 4.1 (*Asymptotic Convergence of (9) Over Strongly Connected and Weight-Balanced Digraphs with Strongly Convex Local Cost Functions*). *Let Assumption 1 hold. Assume that $\mathcal{G}$ is a strongly connected and weight-balanced digraph, and each f^i, $i \in \mathcal{V}$, is differentiable, and m^i-strongly convex ($m^i \in \mathbb{R}_{>0}$). Let $m = \min\{m^1, m^2, \cdots, m^N\}$. Then, for each $i \in \mathcal{V}$, starting from $\mathbf{v}^i(0), \mathbf{y}^i(0), \boldsymbol{\mu}^i(0) \in \mathbb{R}^p$ and $x^i(0) \in \mathbb{R}$ with $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}$, algorithm (9) over $\mathcal{G}$ makes $(\boldsymbol{\mu}^i(t), x^i(t))$ converge asymptotically to $(\boldsymbol{\mu}^\star, \mathsf{x}_i^\star)$, where $\mathbf{x}^\star$ is the unique minimizer of problem (1) and $\boldsymbol{\mu}^\star$ is its corresponding Lagrange multiplier, (see Lemma 3.1), provided*

$$\beta \geq \frac{(\phi+1)^2}{\hat{\lambda}_2\phi} \text{ such that } \|\Omega_{\mathrm{D}}^\top\Big(\Big(\mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top\Big) \otimes \mathbf{I}_p\Big)\Omega_{\mathrm{D}}\| < m(\phi+1),\;\; \phi \in \mathbb{R}_{>0}. \tag{16}$$

Proof. Because of m^i-strong convexity of the local cost functions, the global cost function in (1) is m-strongly convex, with m as defined in the statement. Therefore, the optimization problem (1) has a unique minimizer, i.e., $\mathcal{O}^\star$ is a singleton (see Lemma 3.1). Next, for convenience in convergence analysis, we apply the following change of variables to the states of (9)

$$\begin{aligned} \mathbf{u} &= \mathbf{T}^\top(\mathbf{v} - (\Omega_{\mathrm{D}}\,\mathbf{x}^\star - \tilde{\mathbf{B}})), & \mathbf{z} &= \mathbf{T}^\top(\mathbf{y} - \mathbf{1}_N \otimes \boldsymbol{\mu}^\star),\\ \boldsymbol{\chi} &= \mathbf{x} - \mathbf{x}^\star, & \boldsymbol{\eta} &= \boldsymbol{\mu} - \mathbf{1}_N \otimes \boldsymbol{\mu}^\star, \end{aligned} \tag{17}$$

where $\mathbf{T}$ is defined in (15), and $(\boldsymbol{\mu}^\star, \mathbf{x}^\star) \in \mathcal{O}^\star$. We write $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_{2:N})$ and $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_{2:N})$, where $\mathbf{u}_1, \mathbf{z}_1 \in \mathbb{R}^p$. Notice that given (3), $\mathbf{T} = [\mathsf{r}\;\; \mathsf{R}] \otimes \mathbf{I}_p$ is an orthogonal matrix, i.e., $\mathbf{T}\mathbf{T}^\top = \mathbf{T}^\top\mathbf{T} = \mathbf{I}_{pN}$. Then, in the new variables, the algorithm (9) reads as

$$\dot{\mathbf{u}}_1 = \mathbf{0}_p, \tag{18a}$$
$$\dot{\mathbf{u}}_{2:N} = \beta\,\mathbf{R}^\top\mathbf{L}\mathbf{R}\,\mathbf{z}_{2:N}, \tag{18b}$$
$$\dot{\mathbf{z}}_1 = -\mathbf{z}_1 + \mathbf{r}^\top(\Omega_{\mathrm{D}}\boldsymbol{\chi} + \boldsymbol{\eta}), \tag{18c}$$
$$\dot{\mathbf{z}}_{2:N} = -\mathbf{z}_{2:N} + \mathbf{R}^\top(\Omega_{\mathrm{D}}\boldsymbol{\chi} + \boldsymbol{\eta}) - \beta\,\mathbf{R}^\top\mathbf{L}\mathbf{R}\,\mathbf{z}_{2:N} - \mathbf{u}_{2:N}, \tag{18d}$$
$$\dot{\boldsymbol{\eta}} = -\boldsymbol{\eta} + \mathbf{T}\,\mathbf{z}, \tag{18e}$$
$$\dot{\boldsymbol{\chi}} = -\mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star) - \Omega_{\mathrm{D}}^\top\mathbf{T}\,\mathbf{z}, \tag{18f}$$

where $\mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star) = (\nabla f(\boldsymbol{\chi} + \mathbf{x}^\star) - \nabla f(\mathbf{x}^\star))$. Here, we used $\mathbf{r}^\top\mathbf{v} = \mathbf{0}_p$ which is the result of (13) together with the given initial conditions. Also, given that $(\boldsymbol{\mu}^\star, \mathbf{x}^\star) \in \mathcal{O}^\star$, we used $\mathbf{r}^\top(\Omega_{\mathrm{D}}\,\mathbf{x}^\star - \tilde{\mathbf{B}}) = \frac{1}{\sqrt{N}}(\boldsymbol{\omega}^1\mathsf{x}_1^\star + \cdots + \boldsymbol{\omega}^N\mathsf{x}_N^\star - \mathbf{b}) = \mathbf{0}_p$ and $\nabla f(\mathbf{x}^\star) = -\Omega_{\mathrm{D}}^\top(\mathbf{1}_N \otimes \boldsymbol{\mu}^\star)$.

Let $(\bar{\mathbf{u}}_{2:N}, \bar{\mathbf{z}}, \bar{\boldsymbol{\eta}}, \bar{\boldsymbol{\chi}})$ be an equilibrium point of (18b)–(18f). Then, thanks to $\mathbf{R}^\top\mathbf{L}\mathbf{R}$ being an invertible matrix for strongly connected and weight-balanced digraphs, from (18b) we obtain $\bar{\mathbf{z}}_{2:N} = 0$. Subsequently, from (18e) we obtain $-\bar{\boldsymbol{\eta}} + \mathbf{r}\,\bar{\mathbf{z}}_1 = 0$, and thanks to orthogonality of $\mathbf{T}$, we also obtain $-\mathbf{r}^\top\bar{\boldsymbol{\eta}} + \bar{\mathbf{z}}_1 = 0$ and $-\mathbf{R}^\top\bar{\boldsymbol{\eta}} = 0$. Then, from (18c) we can write $\mathbf{r}^\top\,\Omega_{\mathrm{D}}\,\bar{\boldsymbol{\chi}} = 0$, and from (18d) we can write $\mathbf{R}^\top\,\Omega_{\mathrm{D}}\bar{\boldsymbol{\chi}} - \bar{\mathbf{u}}_{2:N} = 0$. Next, from (18f), thanks to orthogonal matrix, we can write $-\nabla f(\bar{\boldsymbol{\chi}} + \mathbf{x}^\star) - \Omega_{\mathrm{D}}^\top(\mathbf{r}\bar{\mathbf{z}}_1 + \mathbf{1}_N \otimes \boldsymbol{\mu}^\star) = \mathbf{0}$, which equivalently reads as $\nabla f^i(\bar{\chi}^i + \mathsf{x}_i^\star) + \boldsymbol{\omega}^{i\top}(\frac{1}{\sqrt{N}}\bar{\mathbf{z}}_1 + \boldsymbol{\mu}^\star) = 0$, $i \in \mathcal{V}$. Finally note that from $\mathbf{r}^\top\,\Omega_{\mathrm{D}}\,\bar{\boldsymbol{\chi}} = 0$ together with $\mathbf{r}^\top(\Omega_{\mathrm{D}}\,\mathbf{x}^\star - \tilde{\mathbf{B}}) = \mathbf{0}$, we obtain $\mathbf{r}^\top(\Omega_{\mathrm{D}}(\bar{\boldsymbol{\chi}} + \mathbf{x}^\star) - \tilde{\mathbf{B}}) = \mathbf{0}$, which is equivalent to $\frac{1}{\sqrt{N}}(\boldsymbol{\omega}^i(\bar{\chi}^i + \mathsf{x}_1^\star) + \cdots + \boldsymbol{\omega}^N(\bar{\chi}^i + \mathsf{x}_N^\star) - \mathbf{b}) = \mathbf{0}_p$. Recalling the KKT condition (5), we can conclude that $(\frac{1}{\sqrt{N}}\bar{\mathbf{z}}_1 + \boldsymbol{\mu}^\star, \bar{\boldsymbol{\chi}} + \mathbf{x}^\star) \in \mathcal{O}^\star$. However, $\mathcal{O}^\star$ is a singleton set with only one member $(\boldsymbol{\mu}^\star, \mathbf{x}^\star)$. As a result, we obtain $\bar{\mathbf{z}}_1 = \mathbf{0}$ and $\bar{\boldsymbol{\chi}} = \mathbf{0}$. Consequently, from the preceding relations, we obtain $\bar{\mathbf{u}}_{2:N} = \mathbf{0}$ and $\bar{\boldsymbol{\eta}} = \mathbf{0}$. Therefore, (18b)–(18f) have a unique equilibrium point which is located at the origin.

Note that (18a) corresponds to the constant of motion (13). To study the stability in the other variables, consider the radially unbounded and positive-definite candidate Lyapunov function

$$V = \frac{\phi+1}{2}\,\mathbf{z}_1^\top\mathbf{z}_1 + \frac{\phi}{2}\,\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N} + \frac{\phi+1}{2}\,\boldsymbol{\eta}^\top\boldsymbol{\eta} + \frac{\phi+1}{2}\boldsymbol{\chi}^\top\boldsymbol{\chi} + \frac{1}{2}(\mathbf{z}_{2:N} + \mathbf{u}_{2:N})^\top(\mathbf{z}_{2:N} + \mathbf{u}_{2:N}), \tag{19}$$

with $\phi \in \mathbb{R}_{>0}$ as in the statement. The Lie derivative of V along (18b)–(18f) is given by (after some manipulations)

$$\begin{aligned}\dot{V} = &-(\phi+1)\mathbf{z}^\top\mathbf{z} - \phi\beta\mathbf{z}_{2:N}^\top\mathbf{R}^\top\mathbf{L}\mathbf{R}\mathbf{z}_{2:N}\\ &- (\phi+2)\mathbf{z}_{2:N}^\top\mathbf{u}_{2:N} - (\phi+1)\boldsymbol{\eta}^\top\boldsymbol{\eta}\\ &+ 2(\phi+1)\boldsymbol{\eta}^\top[\mathbf{r}\quad \mathbf{R}]\mathbf{z} - (\phi+1)\boldsymbol{\chi}^\top\mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star)\end{aligned}$$

$$
\begin{aligned}
&- \mathbf{u}_{2:N}^\top \mathbf{u}_{2:N} + \mathbf{u}_{2:N}^\top \mathbf{R}^\top \Omega_{\mathrm{D}} \boldsymbol{\chi} + \mathbf{u}_{2:N}^\top \mathbf{R}^\top \boldsymbol{\eta} \\
&= -\phi(\boldsymbol{\eta} - \mathbf{T}\mathbf{z})^\top (\boldsymbol{\eta} - \mathbf{T}\mathbf{z}) - \left(\mathbf{T}^\top \boldsymbol{\eta} - \mathbf{z} - \frac{1}{2}\tilde{\mathbf{u}}\right)^\top \\
&\times \left(\mathbf{T}^\top \boldsymbol{\eta} - \mathbf{z} - \frac{1}{2}\tilde{\mathbf{u}}\right) \\
&- ((\phi+1)\mathbf{z}_{2:N} + \frac{1}{2}\mathbf{u}_{2:N})^\top ((\phi+1)\mathbf{z}_{2:N} + \frac{1}{2}\mathbf{u}_{2:N}) \\
&- \phi\beta \mathbf{z}_{2:N}^\top \mathbf{R}^\top \mathbf{L}\mathbf{R}\mathbf{z}_{2:N} \\
&+ (\phi+1)^2 \mathbf{z}_{2:N}^\top \mathbf{z}_{2:N} - (\phi+1)\boldsymbol{\chi}^\top \mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star) + \boldsymbol{\chi}^\top \Omega_{\mathrm{D}}^\top \mathbf{R}\mathbf{R}^\top \Omega_{\mathrm{D}} \boldsymbol{\chi} \\
&- (\mathbf{R}^\top \Omega_{\mathrm{D}} \boldsymbol{\chi} - \frac{1}{2}\mathbf{u}_{2:N})^\top \left(\mathbf{R}^\top \Omega_{\mathrm{D}} \boldsymbol{\chi} - \frac{1}{2}\mathbf{u}_{2:N}\right) - \frac{1}{2}\mathbf{u}_{2:N}^\top \mathbf{u}_{2:N},
\end{aligned}
$$

where $\tilde{\mathbf{u}} = [\mathbf{0}_p^\top \ \ \mathbf{u}_{2:N}^\top]^\top$. Next, we establish an upper bound on $\dot{V}$. Note that due to m-strong convexity of the global cost function we can write $-\boldsymbol{\chi}^\top \mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star) = -\boldsymbol{\chi}^\top(\nabla f(\boldsymbol{\chi} + \mathbf{x}^\star) - \nabla f(\mathbf{x}^\star)) \leq -m\boldsymbol{\chi}^\top \boldsymbol{\chi}$. Moreover, using (2), we can write $-\mathbf{z}_{2:N}^\top \mathbf{R}^\top \mathbf{L}\mathbf{R}\mathbf{z}_{2:N} \leq -\hat{\lambda}_2 \mathbf{z}_{2:N}^\top \mathbf{z}_{2:N}$. Therefore, we have $\dot{V} \leq \overline{W}$, where

$$
\begin{aligned}
\overline{W} &= -\phi(\boldsymbol{\eta} - \mathbf{T}\mathbf{z})^\top (\boldsymbol{\eta} - \mathbf{T}\mathbf{z}) \\
&- \left(\mathbf{T}^\top \boldsymbol{\eta} - \mathbf{z} - \frac{1}{2}\tilde{\mathbf{u}}\right)^\top \left(\mathbf{T}^\top \boldsymbol{\eta} - \mathbf{z} - \frac{1}{2}\tilde{\mathbf{u}}\right) \\
&- ((\phi+1)\mathbf{z}_{2:N} + \frac{1}{2}\mathbf{u}_{2:N})^\top ((\phi+1)\mathbf{z}_{2:N} + \frac{1}{2}\mathbf{u}_{2:N}) \\
&- (\phi\beta\hat{\lambda}_2 - (\phi+1)^2)\mathbf{z}_{2:N}^\top \mathbf{z}_{2:N} \\
&- \boldsymbol{\chi}^\top ((\phi+1)m\,\mathbf{I} - \Omega_{\mathrm{D}}^\top \mathbf{R}\mathbf{R}^\top \Omega_{\mathrm{D}})\boldsymbol{\chi} \\
&- (\mathbf{R}^\top \Omega_{\mathrm{D}} \boldsymbol{\chi} - \frac{1}{2}\mathbf{u}_{2:N})^\top (\mathbf{R}^\top \Omega_{\mathrm{D}} \boldsymbol{\chi} - \frac{1}{2}\mathbf{u}_{2:N}) - \frac{1}{2}\mathbf{u}_{2:N}^\top \mathbf{u}_{2:N}. \qquad (20)
\end{aligned}
$$

Since $\Omega_{\mathrm{D}}^\top \mathbf{R}\mathbf{R}^\top \Omega_{\mathrm{D}} = \Omega_D^\top((\mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top) \otimes \mathbf{I}_p)\Omega_D$, by virtue of (16), we have $-\boldsymbol{\chi}^\top((\phi+1)m\,\mathbf{I} - \Omega_{\mathrm{D}}^\top \mathbf{R}\mathbf{R}^\top \Omega_{\mathrm{D}})\boldsymbol{\chi} < 0$ for $\boldsymbol{\chi} \neq \mathbf{0}$, and $-(\phi\beta\hat{\lambda}_2 - (\phi+1)^2)\mathbf{z}_{2:N}^\top \mathbf{z}_{2:N} \leq 0$. As a result, $\overline{W}$ is a sum of nonpositive quadratic terms, guaranteeing $\dot{V} \leq \overline{W} \leq 0$.

Because $\dot{V} \leq 0$, the trajectories of (18) under the stated assumptions are bounded. Next, we invoke the invariant set stability analysis theorem to complete our proof. Let $\mathcal{S} = \{(\mathbf{u}_{2:N}, \mathbf{z}, \boldsymbol{\eta}, \boldsymbol{\chi}) \in \mathbb{R}^{p(N-1)} \times \mathbb{R}^{pN} \times \mathbb{R}^{pN} \times \mathbb{R}^N | \dot{V} \equiv 0\}$. Because $\dot{V} \leq \overline{W} \leq 0$, we have $\mathcal{S} \subseteq \overline{\mathcal{S}} = \{(\mathbf{u}_{2:N}, \mathbf{z}, \boldsymbol{\eta}, \boldsymbol{\chi}) \in \mathbb{R}^{p(N-1)} \times \mathbb{R}^{pN} \times \mathbb{R}^{pN} \times \mathbb{R}^N | \overline{W} \equiv 0\}$. By inspecting $\overline{W}$ we obtain $\overline{\mathcal{S}} = \Big\{(\mathbf{u}_{2:N}, \mathbf{z}, \boldsymbol{\eta}, \boldsymbol{\chi}) \in \mathbb{R}^{(N-1)p} \times \mathbb{R}^{pN} \times \mathbb{R}^{pN} \times \mathbb{R}^N \Big| \mathbf{z}_{2:N} \equiv \mathbf{0},\ \boldsymbol{\chi} \equiv \mathbf{0},\ \mathbf{u}_{2:N} \equiv \mathbf{0},\ \boldsymbol{\eta} - \mathbf{T}\mathbf{z} \equiv \mathbf{0}\Big\}$. Then trajectories of (18b)–(18f) that belong to $\overline{\mathcal{S}}$ for all $t \geq 0$ must satisfy ($\dot{\mathbf{u}}_{2:N} \equiv \mathbf{0}$, $\dot{\mathbf{z}}_1 \equiv \mathbf{0}$, $\dot{\mathbf{z}}_{2:N} \equiv \mathbf{0}$, $\dot{\boldsymbol{\eta}} \equiv \mathbf{0}$, $\dot{\boldsymbol{\chi}} \equiv \mathbf{0}$). Because (18b)–(18f) have a unique equilibrium point located at the origin, the only trajectory of (18b)–(18f) that belongs to $\overline{\mathcal{S}}$ for all $t \in \mathbb{R}_{\geq 0}$ is ($\mathbf{u}_{2:N} \equiv \mathbf{0}$, $\mathbf{z} \equiv \mathbf{0}$, $\boldsymbol{\eta} \equiv \mathbf{0}$, $\boldsymbol{\chi} \equiv \mathbf{0}$). Because $\dot{V} \leq \overline{W} \leq 0$ and $\mathcal{S} \subseteq \overline{\mathcal{S}}$, then using proof by contradiction we can show that the only trajectory of (18b)–(18f) that belongs to $\mathcal{S}$ for all $t \in \mathbb{R}_{\geq 0}$ is also ($\mathbf{u}_{2:N} \equiv \mathbf{0}$, $\mathbf{z} \equiv \mathbf{0}$, $\boldsymbol{\eta} \equiv \mathbf{0}$, $\boldsymbol{\chi} \equiv \mathbf{0}$). As a result the largest invariant set in $\mathcal{S}$ is the equilibrium point of (18b)–(18f). Invoking the invariant set theorem [28, Theorem 3.4], we conclude that starting at any initial condition, the trajectories of (18b)–(18f) converge to its unique equilibrium point as $t \to \infty$. Then, given the state equation (18a) and the affine and static state transformation (17), we conclude that starting from any initial condition given in the statement, as $t \to \infty$, the trajectories $t \mapsto (\mathbf{v}^i(t), \mathbf{y}^i(t), \boldsymbol{\mu}^i(t), x^i(t))$, $i \in \mathcal{V}$ converge to $(\boldsymbol{\omega}^i x_i^\star - \tilde{\mathbf{b}}^i, \boldsymbol{\mu}^\star, \boldsymbol{\mu}^\star, x_i^\star)$, where $(\boldsymbol{\mu}^\star, \mathbf{x}^\star) \in \mathcal{O}^\star$. This completes the proof. □

In Theorem 4.1, the requirement $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}$ is trivially satisfied if each agent $i \in \mathcal{V}$ starts at $\mathbf{v}^i(0) = \mathbf{0}$. The determination of admissible value of β by individual agents can be achieved if they know a lower bound on m, have knowledge of $\hat{\lambda}_2$, either through a lower bound on it (see e.g., [29]) or dedicated algorithms to compute it (see, e.g., [30]), and know an upper bound on $\|\Omega_D^\top((\mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top) \otimes \mathbf{I}_p)\Omega_D\|$. Let each agent choose its local weights as nonnegative scalars that satisfy $\sum_{j=1}^p \omega_j^i \leq 1$. Then, we can rely on $\|\Omega_{\mathrm{D}}^\top((\mathbf{I} - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top) \otimes \mathbf{I}_p)\Omega_{\mathrm{D}}\| \leq \|((\mathbf{I} - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top) \otimes \mathbf{I}_p)\|\|\Omega_{\mathrm{D}}\|^2 = \|\Omega_{\mathrm{D}}\|^2 \leq \|\Omega_{\mathrm{D}}\|_1\|\Omega_{\mathrm{D}}\|_\infty$, $\|\Omega_{\mathrm{D}}\|_1 \leq 1$, and $\|\Omega_{\mathrm{D}}\|_\infty \leq 1$ to use 1 as an upper bound on $\|\Omega_D^\top((\mathbf{I} - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top) \otimes \mathbf{I}_p)\Omega_D\|$. It is also interesting to observe that to guarantee convergence over strongly connected and weight-balanced digraph topologies, other distributed optimization algorithms also require bounds similar to (16) for a scalar parameter of the algorithm (see e.g., [11,14] for unconstrained convex optimization and [16] for a constrained optimization problems).

Next, we show that if strongly convex local cost functions have globally Lipschitz gradients, then convergence of (9) over strongly connected and weight-balanced digraphs is exponentially fast for problems with single demand equation.

Proposition 4.1 (*Exponential Convergence of (9) Over Strongly Connected and Weight-Balanced Digraphs With Strongly Convex Local Cost Functions and Globally Lipschitz Gradients*). *Let Assumption 1 hold. Assume that $\mathcal{G}$ is a strongly connected and weight-balanced digraph, and each f^i, $i \in \mathcal{V}$, is differentiable, m^i-strongly convex and has M^i-Lipschitz gradient ($M^i, m^i \in \mathbb{R}_{>0}$). Let $m = \min\{m^1, m^2, \cdots, m^N\}$, and $M = \max\{M^1, M^2, \cdots, M^N\}$. In the optimal resource allocation problem (1) let $p = 1$ (single demand) and $\omega^i = 1$, $i \in \mathcal{V}$. For each $i \in \mathcal{V}$, starting from $v^i(0), y^i(0), \mu^i(0), x^i(0) \in \mathbb{R}$ with $\sum_{i=1}^N v^i(0) = 0$, the algorithm (9) over $\mathcal{G}$ makes $(\mu^i(t), x^i(t))$, $i \in \mathcal{V}$, converge exponentially fast to $(\mu^\star, x_i^\star)$, where $\mathbf{x}^\star$ is the unique minimizer of problem (1) and $\mu^\star$ is its corresponding Lagrange multiplier, (see Lemma 3.1), provided*

$$
\beta \geq \frac{(\phi+1)(\phi+2)}{\hat{\lambda}_2 \phi}, \ \text{such that} \ \phi + 2 > \frac{1.25 + (M)^2}{m}, \quad \phi \in \mathbb{R}_{>0}. \qquad (21)
$$

Proof. We follow the proof of Theorem 4.1 until the choice of Lyapunov function to analyze the stability of (18b)–(18f). Here, we use the radially unbounded and positive-definite candidate Lyapunov function

$$
\begin{aligned}
V &= \frac{\phi+1}{2} z_1^\top z_1 + \frac{\phi}{2}\mathbf{z}_{2:N}^\top \mathbf{z}_{2:N} + \frac{\phi+1}{2}\boldsymbol{\eta}^\top \boldsymbol{\eta} + \frac{\phi+1}{2}\boldsymbol{\chi}^\top \boldsymbol{\chi} \\
&+ \frac{1}{2}(\mathbf{z}_{2:N} + \mathbf{u}_{2:N})^\top (\mathbf{z}_{2:N} + \mathbf{u}_{2:N}) + \frac{1}{2}\boldsymbol{\eta}^\top \mathbf{r}\mathbf{r}^\top \boldsymbol{\eta} \\
&+ \frac{1}{2}(\boldsymbol{\chi} + \mathbf{r} z_1)^\top (\boldsymbol{\chi} + \mathbf{r} z_1) = \boldsymbol{\zeta}^\top \mathbf{E} \boldsymbol{\zeta},
\end{aligned}
$$

where $\phi \in \mathbb{R}_{>0}$ as in (26), and $\boldsymbol{\zeta} = [z_1, \mathbf{z}_2^\top, \mathbf{u}_{2:N}^\top, \boldsymbol{\eta}^\top, \boldsymbol{\chi}^\top]^\top$, with $\mathbf{E} > 0$ being the obvious matrix describing the quadratic V. Note that this Lyapunov function consists of the Lyapunov candidate function (19) plus the last two new quadratic terms. The Lie derivative of these last two terms along trajectories of (18b)–(18f) is given by

$$
\begin{aligned}
&\boldsymbol{\eta}^\top \mathbf{r}\mathbf{r}^\top \dot{\boldsymbol{\eta}} + (\boldsymbol{\chi} + \mathbf{r} z_1)^\top (\dot{\boldsymbol{\chi}} + \mathbf{r}\dot{z}_1) = \boldsymbol{\eta}^\top \mathbf{r}\mathbf{r}^\top(-\boldsymbol{\eta} + \mathbf{r} z_1 + \mathbf{R}\mathbf{z}_{2:N}) \\
&+ (\boldsymbol{\chi} + \mathbf{r} z_1)^\top (-\mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star) \\
&- \mathbf{r} z_1 - \mathbf{R}\mathbf{z}_{2:N} - \mathbf{r} z_1 + \mathbf{r}\mathbf{r}^\top \boldsymbol{\chi} + \mathbf{r}\mathbf{r}^\top \boldsymbol{\eta}) \\
&= -\left(\mathbf{r}^\top \boldsymbol{\eta} - \frac{1}{2}\mathbf{r}^\top \boldsymbol{\chi} - z_1\right)^2 - \boldsymbol{\chi}^\top \mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star) \\
&- \frac{3}{4} z_1^2 - (\frac{1}{2}\mathbf{r} z_1 + \mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star))^\top (\frac{1}{2}\mathbf{r} z_1 + \mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star)) \\
&+ \mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star)^\top \mathbf{h}(\boldsymbol{\chi}, \mathbf{x}^\star) + \frac{5}{4}\boldsymbol{\chi}^\top \boldsymbol{\chi} \\
&- \boldsymbol{\chi}^\top \mathbf{R}\mathbf{R}^\top \boldsymbol{\chi} - \left(\frac{1}{2}\mathbf{R}^\top \boldsymbol{\chi} + \mathbf{z}_{2:N}\right)^\top \left(\frac{1}{2}\mathbf{R}^\top \boldsymbol{\chi} + \mathbf{z}_{2:N}\right) + \mathbf{z}_{2:N}^\top \mathbf{z}_{2:N},
\end{aligned}
$$

where given $p = 1$ and $\omega^i = 1$, $i \in \mathcal{V}$, we used $\mathbf{r} = \mathsf{r}$, $\mathbf{R} = \mathsf{R}$ and $\Omega_{\mathrm{D}} = \mathbf{I}_N$. We also invoked $\mathsf{r}\mathsf{r}^\top = \mathbf{I} - \mathsf{R}\mathsf{R}^\top$. Using this manipulation, $\dot{V}$ along trajectories of (18b)–(18f) is

$$\begin{aligned}\dot{V} = &-\phi(\boldsymbol{\eta} - \mathsf{T}\mathbf{z})^\top(\boldsymbol{\eta} - \mathsf{T}\mathbf{z}) - \Big(\mathsf{T}^\top\boldsymbol{\eta} - \mathbf{z} - \frac{1}{2}\tilde{\mathbf{u}}\Big)^\top\Big(\mathsf{T}^\top\boldsymbol{\eta} - \mathbf{z} - \frac{1}{2}\tilde{\mathbf{u}}\Big)\\
&-((\phi+1)\mathbf{z}_{2:N} + \frac{1}{2}\mathbf{u}_{2:N})^\top\Big((\phi+1)\mathbf{z}_{2:N} + \frac{1}{2}\mathbf{u}_{2:N}\Big)\\
&-(\phi\beta\mathbf{z}_{2:N}^\top\mathsf{R}^\top\mathsf{L}\mathsf{R}\mathbf{z}_{2:N} - (\phi+1)^2\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N})\\
&-(\phi+1)\boldsymbol{\chi}^\top\mathbf{h}(\boldsymbol{\chi},\mathbf{x}^\star) - \frac{1}{4}\mathbf{u}_{2:N}^\top\mathbf{u}_{2:N}\\
&+\mathbf{u}_{2:N}^\top\mathsf{R}^\top\boldsymbol{\chi} - \frac{1}{2}\mathbf{u}_{2:N}^\top\mathbf{u}_{2:N} - (\mathsf{r}^\top\boldsymbol{\eta} - \frac{1}{2}\mathsf{r}^\top\boldsymbol{\chi} - z_1)^2\\
&-\boldsymbol{\chi}^\top\mathbf{h}(\boldsymbol{\chi},\mathbf{x}^\star) - \frac{3}{4}z_1^2\\
&-\Big(\frac{1}{2}\mathsf{r}z_1 + \mathbf{h}(\boldsymbol{\chi},\mathbf{x}^\star)\Big)^\top\Big(\frac{1}{2}\mathsf{r}z_1 + \mathbf{h}(\boldsymbol{\chi},\mathbf{x}^\star)\Big)\\
&+\mathbf{h}(\boldsymbol{\chi},\mathbf{x}^\star)^\top\mathbf{h}(\boldsymbol{\chi},\mathbf{x}^\star) + \frac{5}{4}\boldsymbol{\chi}^\top\boldsymbol{\chi} - \boldsymbol{\chi}^\top\mathsf{R}\mathsf{R}^\top\boldsymbol{\chi}\\
&-\Big(\frac{1}{2}\mathsf{R}^\top\boldsymbol{\chi} + \mathbf{z}_{2:N}\Big)^\top(\frac{1}{2}\mathsf{R}^\top\boldsymbol{\chi} + \mathbf{z}_{2:N}) + \mathbf{z}_{2:N}^\top\mathbf{z}_{2:N}.\end{aligned}$$

Next, we establish an upper bound on $\dot{V}$. We start by using the M^i-Lipschitzness property of local gradients to write $\mathbf{h}(\boldsymbol{\chi},\mathbf{x}^\star)^\top\mathbf{h}(\boldsymbol{\chi},\mathbf{x}^\star) \le (M)^2\boldsymbol{\chi}^\top\boldsymbol{\chi}$ (recall that $\mathbf{h}(\boldsymbol{\chi},\mathbf{x}^\star) = (\nabla f(\boldsymbol{\chi}+\mathbf{x}^\star) - \nabla f(\mathbf{x}^\star))$. We also write $-\boldsymbol{\chi}^\top\mathbf{h}(\boldsymbol{\chi},\mathbf{x}^\star) \le -m\boldsymbol{\chi}^\top\boldsymbol{\chi}$ due to the m-strong convexity of global cost function and $-\mathbf{z}_{2:N}^\top\mathsf{R}^\top\mathsf{L}\mathsf{R}\mathbf{z}_{2:N} \le -\hat{\lambda}_2\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N}$, which is true because the communication topology is a strongly connected and weight-balanced digraph. Then, we have

$$\begin{aligned}\dot{V} \le &-\phi(\boldsymbol{\eta} - \mathsf{T}\mathbf{z})^\top(\boldsymbol{\eta} - \mathsf{T}\mathbf{z}) - \Big(\mathsf{T}^\top\boldsymbol{\eta} - \mathbf{z} - \frac{1}{2}\tilde{\mathbf{u}}\Big)^\top\Big(\mathsf{T}^\top\boldsymbol{\eta} - \mathbf{z} - \frac{1}{2}\tilde{\mathbf{u}}\Big)\\
&-((\phi+1)\mathbf{z}_{2:N} + \frac{1}{2}\mathbf{u}_{2:N})^\top((\phi+1)\mathbf{z}_{2:N} + \frac{1}{2}\mathbf{u}_{2:N})\\
&-(\phi\beta\hat{\lambda}_2 - (\phi+1)(\phi+2))\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N}\\
&-\boldsymbol{\chi}^\top((\phi+2)m - \frac{5}{4} - (M)^2)\boldsymbol{\chi}\\
&-\Big(\mathsf{R}^\top\boldsymbol{\chi} - \frac{1}{2}\mathbf{u}_{2:N}\Big)^\top\Big(\mathsf{R}^\top\boldsymbol{\chi} - \frac{1}{2}\mathbf{u}_{2:N}\Big) - \Big(\frac{1}{2}\mathsf{R}^\top\boldsymbol{\chi} + \mathbf{z}_{2:N}\Big)^\top\\
&\times\Big(\frac{1}{2}\mathsf{R}^\top\boldsymbol{\chi} + \mathbf{z}_{2:N}\Big)\\
&-\frac{1}{2}\mathbf{u}_{2:N}^\top\mathbf{u}_{2:N} - \frac{3}{4}z_1^2 = -\boldsymbol{\zeta}^\top\mathbf{F}\,\boldsymbol{\zeta},\end{aligned}$$

where $\mathbf{F}$ is the obvious matrix describing the derived quadratic upper bound on $\dot{V}$. Given the conditions on ϕ and β in the statement, we have $-\boldsymbol{\zeta}^\top\mathbf{F}\,\boldsymbol{\zeta} < 0$ (or $\mathbf{F} > 0$), and consequently $\dot{V} < 0$ along the trajectories of (18b)–(18f). Because V is a quadratic positive definite function and the upper bound on $\dot{V}$ is a quadratic negative definite function, (18b)–(18f) are exponentially stable, and its trajectories converge to origin with the rate no worse than $\frac{\lambda_{\min}(\mathbf{F})}{2\lambda_{\max}(\mathbf{E})}$, where $\lambda_{\min}(\mathbf{F})$ is the minimum eigenvalue of $\mathbf{F}$ and $\lambda_{\max}(\mathbf{E})$ is the maximum eigenvalue of $\mathbf{E}$ (cf. [28, Theorem 3.1]). Consequently, we conclude that starting from any initial condition given in the statement, as $t \to \infty$, trajectories $t \mapsto (v^i(t), y^i(t), \mu^i(t), x^i(t))$, $i \in \mathcal{V}$ converge to $(\omega^i\mathsf{x}_i^\star - \mathsf{b}^i, \mu^\star, \mu^\star, \mathsf{x}_i^\star)$ exponentially fast with the rate given above. This completes the proof. □

Below, we show that the correctness of algorithm (9) over connected graphs is guaranteed for any $\beta \in \mathbb{R}_{>0}$. Such a result is obtained by invoking the positive definiteness of $\mathsf{R}^\top\mathsf{L}\mathsf{R}$ to construct a new Lyapunov function to study stability of (18b)–(18f).

We also show that the convergence guarantees of algorithm (9) over connected graphs extend to optimal allocation problems with local convex functions. Recall that when local cost functions are convex, the optimization problem (1) can have infinite number of minimizers. Given the discussions in Remark 4.1 which connect the equilibrium points of algorithm (9) to the minimizers of optimization problem (1), algorithm (9) for convex local cost functions has infinite number of equilibrium points. Next, we use results from semi-stability analysis to show that, under proper initialization, algorithm (9) converges to one of the minimizers of the optimization problem (1).

Theorem 4.2 (*Convergence of (9) to a Point in the Minimizer Set of Optimization Problem (1) Over Connected Graphs*)**.** *Let Assumption 1 hold. Let $\mathcal{G}$ be a connected graph. Assume each f^i, $i \in \mathcal{V}$, is differentiable and convex. Then, for each $i \in \mathcal{V}$, starting from $\mathbf{y}^i(0), \mathbf{v}^i(0), \boldsymbol{\mu}^i(0) \in \mathbb{R}^p$, and $x^i(0) \in \mathbb{R}$ with $\sum_{i=1}^N \mathbf{v}^i(0) = \mathbf{0}$, as $t \to \infty$, algorithm (9) over $\mathcal{G}$ makes $(\boldsymbol{\mu}^i(t), \mathbf{x}(t))$ converge to $(\boldsymbol{\mu}^\star, \mathsf{x}_i^\star)$, where $\mathbf{x}^\star$ is a minimizer of problem (1) and $\boldsymbol{\mu}^\star$ is its corresponding Lagrange multiplier, (see Lemma 3.1).*

Proof. For convenience in analysis, here, once again we apply the change of state (17) to obtain (18), an equivalent representation of (9). Once again, (18a) corresponds to the constant of motion (13). To study the stability in the other variables, here we take advantage of structural properties of the Laplacian matrix for connected graphs (recall (2)) to consider the following radially unbounded and positive-definite candidate Lyapunov function

$$V = \frac{1}{2\beta}\mathbf{u}_{2:N}^\top(\mathsf{R}^\top\mathsf{L}\mathsf{R})^{-1}\mathbf{u}_{2:N} + \frac{1}{2}(\mathbf{z}^\top\mathbf{z} + \boldsymbol{\eta}^\top\boldsymbol{\eta} + \boldsymbol{\chi}^\top\boldsymbol{\chi}). \tag{22}$$

The Lie derivative of V along (18b)–(18f) is given by (after some manipulations)

$$\begin{aligned}\dot{V} = &-\beta\,\mathbf{z}_{2:N}^\top\mathsf{R}^\top\mathsf{L}\mathsf{R}\mathbf{z}_{2:N}^\top - (\boldsymbol{\eta} - \mathsf{T}\,\mathbf{z})^\top(\boldsymbol{\eta} - \mathsf{T}\,\mathbf{z})\\
&-((\boldsymbol{\chi}+\mathbf{x}^\star) - \mathbf{x}^\star)^\top(\nabla f(\boldsymbol{\chi}+\mathbf{x}^\star) - \nabla f(\mathbf{x}^\star)) \le 0.\end{aligned} \tag{23}$$

Here, we invoked the convexity of the local cost functions and (2) for connected graphs to establish $-((\boldsymbol{\chi}+\mathbf{x}^\star) - \mathbf{x}^\star)^\top(\nabla f(\boldsymbol{\chi}+\mathbf{x}^\star) - \nabla f(\mathbf{x}^\star)) \le 0$, and $-\mathbf{z}_{2:N}^\top\mathsf{R}^\top\mathsf{L}\mathsf{R}\mathbf{z}_{2:N} \le -\lambda_2\mathbf{z}_{2:N}^\top\mathbf{z}_{2:N} < 0$. respectively.

So far, by virtue of $\dot{V} \le 0$, we have shown that the trajectories of (18) and as a result (9) under the stated assumptions are bounded. Next, we use the invariant set stability analysis theorems to complete our proof. Let $\mathcal{S} = \{(\mathbf{u}_{2:N}, \mathbf{z}, \boldsymbol{\eta}, \boldsymbol{\chi}) \in \mathbb{R}^{p(N-1)} \times \mathbb{R}^{pN} \times \mathbb{R}^{pN} \times \mathbb{R}^N \,|\, \dot{V} \equiv 0\}$. For a connected graph, we have

$$\begin{aligned}\mathcal{S} = \big\{&(\mathbf{u}_{2:N}, \mathbf{z}, \boldsymbol{\eta}, \boldsymbol{\chi}) \in \mathbb{R}^{(N-1)p} \times \mathbb{R}^{pN} \times \mathbb{R}^{pN} \times \mathbb{R}^N \,\big|\\
&\mathbf{z}_{2:N} \equiv \mathbf{0},\ \boldsymbol{\eta} - \mathbf{r}\,\mathbf{z}_1 \equiv \mathbf{0},\ \boldsymbol{\chi}^\top(\nabla f(\boldsymbol{\chi}+\mathbf{x}^\star) - \nabla f(\mathbf{x}^\star)) \equiv \mathbf{0}\big\}.\end{aligned}$$

Next, we identify the largest invariant set of (18b)–(18f) in $\mathcal{S}$. A trajectory $t \mapsto (\mathbf{u}_{2:N}(t), \mathbf{z}(t), \boldsymbol{\eta}(t), \boldsymbol{\chi}(t))$ of (18b)–(18f) belonging to $\mathcal{S}$ for $t \ge 0$ must satisfy

$$\dot{\mathbf{u}}_{2:N} \equiv \mathbf{0}, \tag{24a}$$

$$\mathbf{0} \equiv \mathbf{r}^\top\Omega_{\mathrm{D}}\boldsymbol{\chi}, \tag{24b}$$

$$\mathbf{0} \equiv \mathbf{R}^\top\Omega_{\mathrm{D}}\boldsymbol{\chi} - \mathbf{u}_{2:N}, \tag{24c}$$

$$\dot{\boldsymbol{\eta}} \equiv \mathbf{0}, \tag{24d}$$

$$\dot{\boldsymbol{\chi}} = -(\nabla f(\boldsymbol{\chi}+\mathbf{x}^\star) - \nabla f(\mathbf{x}^\star)) - \Omega_{\mathrm{D}}^\top\mathbf{r}\mathbf{z}_1. \tag{24e}$$

From (24a)–(24c), for trajectories in $\mathcal{S}$, we obtain $[\mathbf{r}\ \ \mathbf{R}]^\top\Omega_{\mathrm{D}}\dot{\boldsymbol{\chi}} \equiv \mathbf{0}$ which results in $\dot{\boldsymbol{\chi}} \equiv \mathbf{0}$ (recall that $[\mathbf{r}\ \ \mathbf{R}]^\top$ is a full rank matrix and Ω_{D} is full column rank). Therefore, we can establish that the trajectories in $\mathcal{S}$ should satisfy ($\dot{\mathbf{u}}_{2:N}\equiv\mathbf{0}$, $\dot{\mathbf{z}}\equiv\mathbf{0}$, $\dot{\boldsymbol{\eta}}\equiv\mathbf{0}$, $\dot{\boldsymbol{\chi}}\equiv\mathbf{0}$). Therefore, the largest invariant set in $\mathcal{S}$ is the set of equilibrium points of (18b)–(18f). Invoking the invariant set theorem [28, Theorem 3.4], we conclude that starting at any initial condition, as $t\to\infty$ the trajectories of (18b)–(18f) approach the set of its equilibrium points. Next, we show that all the equilibrium points of (18b)–(18f) are Lyapunov stable. Then, because the set that trajectories of (18b)–(18f) are asymptotically approaching is the set of stable equilibrium points, we can invoke [28, Theorem 4.20] to establish that (18b)–(18f) are semi-stable and the trajectories of (18b)–(18f) will converge to a point in this set.

To study the stability of any equilibrium point ($\underline{\mathbf{u}}_{2:N}$, $\underline{\mathbf{z}}_1$, $\underline{\mathbf{z}}_{2:N}$, $\underline{\boldsymbol{\eta}}$, $\underline{\boldsymbol{\chi}}$), we transfer that equilibrium point to the origin using $\mathbf{p}=\mathbf{u}_{2:N}-\underline{\mathbf{u}}_{2:N}$, $\mathbf{q}_1=\mathbf{z}_1-\bar{\mathbf{z}}_1$, $\mathbf{q}_{2:N}=\mathbf{z}_{2:N}-\underline{\mathbf{z}}_{2:N}$, $\boldsymbol{\zeta}=\boldsymbol{\eta}-\underline{\boldsymbol{\eta}}$, and $\boldsymbol{\aleph}=\boldsymbol{\chi}-\underline{\boldsymbol{\chi}}$ to write (18b)–(18f) in the following equivalent form

$$\dot{\mathbf{p}}=\beta\,\mathbf{R}^\top\mathbf{L}\mathbf{R}\mathbf{q}_{2:N},\quad \dot{q}_1=-\mathbf{q}_1+\mathbf{r}^\top(\Omega_{\rm D}\boldsymbol{\aleph}+\boldsymbol{\zeta}), \tag{25a}$$

$$\dot{\mathbf{q}}_{2:N}=-\mathbf{z}_{2:N}+\mathbf{R}^\top(\Omega_{\rm D}\boldsymbol{\aleph}+\boldsymbol{\zeta})-\beta\,\mathbf{R}^\top\mathbf{L}\mathbf{R}\,\mathbf{q}_{2:N}-\mathbf{p}, \tag{25b}$$

$$\dot{\boldsymbol{\zeta}}=-\boldsymbol{\zeta}+\mathbf{T}\mathbf{q},\quad \dot{\boldsymbol{\aleph}}=-\tilde{\mathbf{h}}(\boldsymbol{\aleph},\mathbf{x}^\star)-\Omega_{\rm D}^\top\mathbf{T}\,\mathbf{q}, \tag{25c}$$

where $\tilde{\mathbf{h}}(\boldsymbol{\aleph})=\mathbf{h}(\boldsymbol{\aleph}+\underline{\boldsymbol{\chi}})-\mathbf{h}(\underline{\boldsymbol{\chi}})=\nabla f(\boldsymbol{\aleph}+\underline{\boldsymbol{\chi}}+\mathbf{x}^\star)-\nabla f(\underline{\boldsymbol{\chi}}+\mathbf{x}^\star)$. To study the stability of the origin in (25), we use the Lyapunov candidate function (22) in which ($\mathbf{u}_{2:N}$, $\mathbf{z}$, $\boldsymbol{\eta}$, $\boldsymbol{\chi}$) is replaced, respectively, with ($\mathbf{p}$, $\mathbf{q}$, $\boldsymbol{\zeta}$, $\boldsymbol{\aleph}$). Taking the derivative of this V along the trajectories of (25), gives, similar to (23),

$$\dot{V}=-\beta\,\mathbf{q}_{2:N}^\top\mathbf{R}^\top\mathbf{L}\mathbf{R}\mathbf{q}_{2:N}^\top-(\boldsymbol{\zeta}-\mathbf{T}\,\mathbf{q})^\top(\boldsymbol{\zeta}-\mathbf{T}\,\mathbf{q})-\boldsymbol{\aleph}^\top\tilde{\mathbf{h}}(\boldsymbol{\aleph})\leq 0.$$

Convexity of local cost functions gives $(\aleph^i+\underline{\chi}^i+\mathsf{x}_i^\star-(\underline{\chi}^i+\mathsf{x}_i^\star))(\nabla f^i(\aleph^i+\underline{\chi}^i+\mathsf{x}_i^\star)-\nabla f^i(\underline{\chi}^i+\mathsf{x}_i^\star))\geq 0$, $i\in\mathcal{V}$, or equivalently $\boldsymbol{\aleph}^\top\tilde{\mathbf{h}}(\boldsymbol{\aleph})\geq 0$ for all $\boldsymbol{\aleph}\in\mathbb{R}^N$. Moreover, because the network topology is a connected graph we have $\mathsf{R}^\top\mathsf{L}\mathsf{R}>0$. Therefore, we conclude that along the trajectories of (25) we have $\dot{V}\leq 0$. Therefore, any ($\underline{\mathbf{u}}_{2:N}$, $\underline{\mathbf{z}}_1$, $\underline{\mathbf{z}}_{2:N}$, $\underline{\boldsymbol{\eta}}$, $\underline{\boldsymbol{\chi}}$) is a Lyapunov stable equilibrium point.

So far, we have shown that starting from any initial condition, the trajectories of (18b)–(18f) converge to a point in its equilibrium set. Then, given the state equation (18a) and the affine and static state transformation (17), we conclude that starting from any initial condition given in the statement, the trajectories $t\mapsto(\mathbf{v}^i(t),\mathbf{y}^i(t),\boldsymbol{\mu}^i(t),x^i(t))$, $i\in\mathcal{V}$, of (9) converge to one of its equilibrium points as $t\to\infty$. Then, with a discussion similar to that made in Remark 4.1, we can conclude that as $t\to\infty$ the trajectories $t\mapsto(\mathbf{v}^i(t),\mathbf{y}^i(t),\boldsymbol{\mu}^i(t),x^i(t))$, $i\in\mathcal{V}$, converge to $(\boldsymbol{\omega}^i\mathsf{x}_i^\star-\tilde{\mathbf{b}}^i,\boldsymbol{\mu}^\star,\boldsymbol{\mu}^\star,\mathsf{x}_i^\star)$ where $(\boldsymbol{\mu}^\star,\mathbf{x}^\star)$ is a KKT point in $\mathcal{O}^\star$. This completes the proof. □

We close this section by discussing how the convergence of (9) can be extended to the dynamically changing topologies. Such extension is immediate as the proof of Theorem 4.1 relies on a Lyapunov function that has no dependency on the systems parameters and its derivative is negative definite with a quadratic upper bound. The proof details are omitted for brevity.

Proposition 4.2 (*Convergence of (9) Over Dynamically Changing Strongly Connected and Weight-Balanced Digraphs*)**.** *Let Assumption 1 hold. Assume that $\mathcal{G}$ is a time-varying digraph which is strongly connected and weight-balanced at all times and whose adjacency matrix is uniformly bounded and piecewise constant. Let each f^i, $i\in\mathcal{V}$, be differentiable, m^i-strongly convex and have M^i-Lipschitz gradient (M^i, $m^i\in\mathbb{R}_{>0}$). In the optimal resource allocation problem (1) let $p=1$ (single demand) and $\omega^i=1$, $i\in\mathcal{V}$. For each $i\in\mathcal{V}$, starting from $v^i(0)$, $y^i(0)$, $\mu^i(0)$, $x^i(0)\in\mathbb{R}$ with $\sum_{i=1}^N v^i(0)=0$, algorithm (9) over $\mathcal{G}$ makes $(\mu^i(t),x^i(t))$, $i\in\mathcal{V}$, converge exponentially fast to $(\mu^\star,\mathsf{x}_i^\star)$, where $\mathbf{x}^\star$ is the minimizer of problem (1) and $\mu^\star$ is its corresponding Lagrange multiplier, (see Lemma 3.1), provided*

$$\beta\geq\frac{(\phi+1)(\phi+2)}{(\hat{\lambda}_2)_{\min}\phi},\ \text{such that}\ \phi+2>\frac{1.25+(M)^2}{m},\qquad \phi\in\mathbb{R}_{>0}, \tag{26}$$

where $(\hat{\lambda}_2)_{\min}=\min_{p\in\mathcal{P}}\{\hat{\lambda}_2(\mathbf{L}_p)\}$ where $\mathcal{P}$ is the index set of all possible realizations of $\mathcal{G}$ and $m=\min\{m^1,m^2,\cdots,m^N\}$, $M=$ $\max\{M^1,M^2,\cdots,M^N\}$.

5. Simulations

We consider an in-network resource allocation problem for a group of 7 agents interacting over a strongly connected and weight-balanced digraph depicted in Fig. 1. The local cost function for each agent is given by $f^i(x^i)=\alpha^i x^{i2}+\beta^i x^i+\gamma^i$, where $(\alpha^i,\beta^i,\gamma^i)$, $i\in\{1,\dots,7\}$, are selected randomly in the intervals, respectively, ([0.10875, 0.06967], [10.76, 74], [6.78, 32.96]). The cost functions are selected according to the cost for power generators in IEEE 118 bus. In this problem, agents {1, 2, 3, 4, 7} are meeting a demand $\mathsf{b}_1=850$ and agents {4, 5, 6, 7} are meeting another demand $\mathsf{b}_2=750$. Agents 4 and 7 contribute equally to each of these demands by splitting their allocated value into half between them. Our objective here is to meet these demands with least possible cost for the group. The optimization problem we solve is

$$\begin{aligned}\mathbf{x}^\star&=\arg\min_{\mathbf{x}\in\mathbb{R}^7}\ \sum\nolimits_{i=1}^7 f^i(x^i),\quad\text{subject to}\\ &\qquad x^1+x^2+x^3+0.5\,x^4+0.5\,x^7=850,\\ &\qquad 0.5x^4+x^5+x^6+0.5\,x^7=750.\end{aligned} \tag{27}$$

Fig. 2(a)–(c) show the time history of x^i's generated by implementing the distributed optimization algorithm (9) over the network shown in Fig. 1(a) for different values of β, and compares it to the solution obtained when (27) is solved as a centralized problem by Matlab's constraint optimization solver 'fmincon'. As expected the decision variable x^i of each agent $i\in\{1,\dots,7\}$ converges to its corresponding solution of the optimization problem (27). Fig. 2(d) shows the execution of algorithm (9) when network topology changes every 20 seconds from network in Fig. 1(a) to the one in Fig. 1(b). In this example, we have observed that the algorithm converges for any positive value of β.

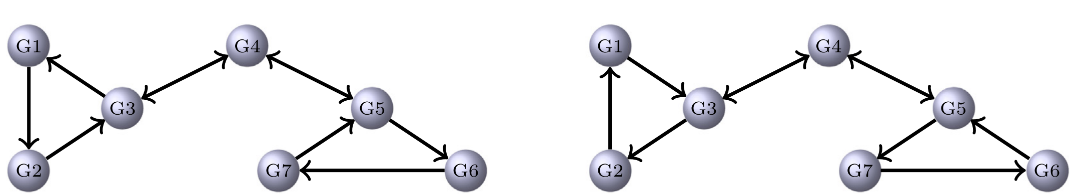
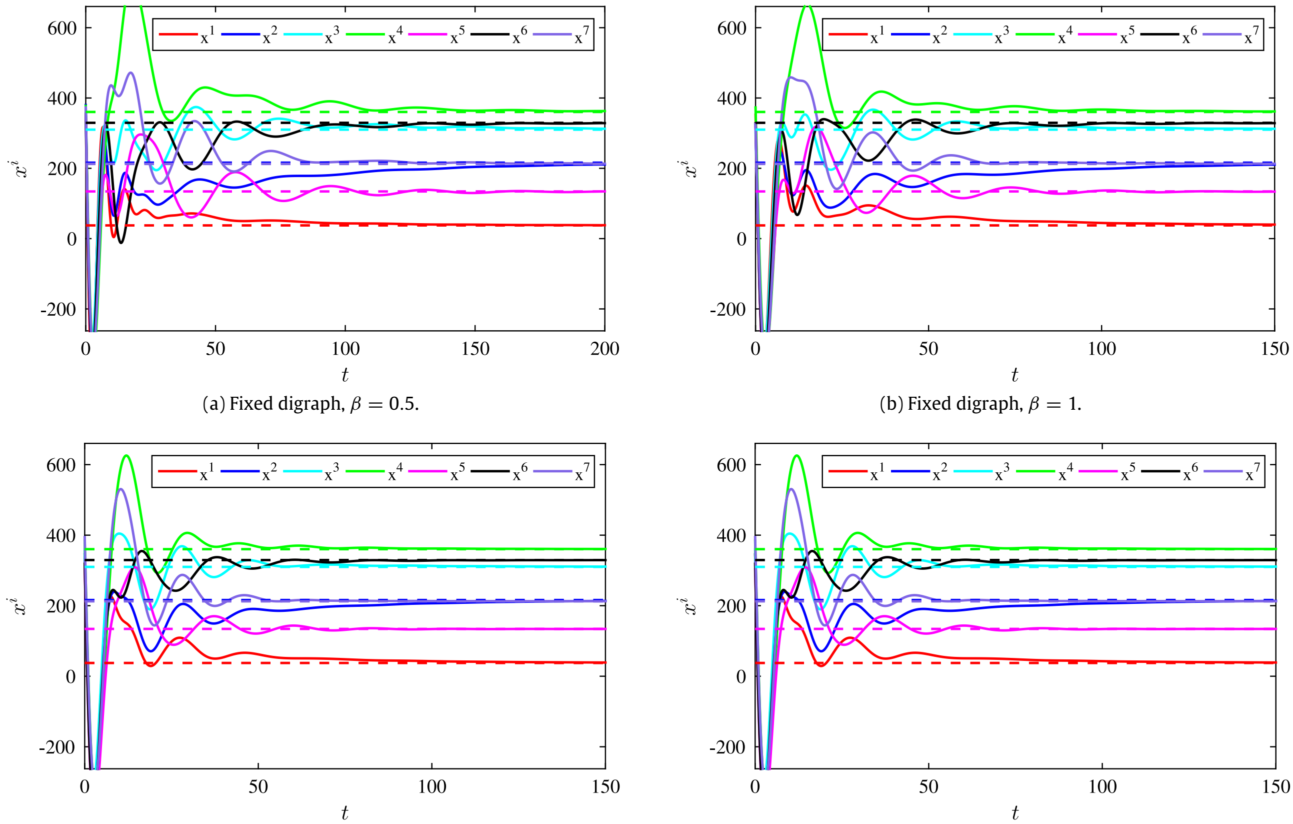


Fig. 1. The digraphs used in the simulation study (adjacency weights are 1).



(a) Fixed digraph, $\beta = 0.5$.

(b) Fixed digraph, $\beta = 1$.

(c) Fixed digraph, $\beta = 2$.

(d) Time-varying digraph, $\beta = 1$.

Fig. 2. Plots (a)–(c) show the result of execution of algorithm (9) over the network depicted in Fig. 1(a) for three different values of β. Plot (d) shows the execution of (9) over a network whose topology changes every 20 s from the network depicted in Fig. 1(a) to the one in Fig. 1(b). In all plots, the colored solid curved plots depict the time history of decision variable of each agent generated by algorithm (9). Horizontal dashed lines depict the centralized solution obtained using Matlab's constraint optimization solver 'fmincon'. As plot (a)–(c) show, by increasing β the rate of convergence of the algorithm increases.

6. Conclusions

We presented a novel distributed algorithm to solve an optimal in-network resource allocation problem where the total cost is the sum of local cost functions and the constraint is the weighted sum of contribution of each agent. We showed that when the communication topology is a strongly connected and weight-balanced digraph and the local cost functions are strongly convex, with an appropriate initialization, the proposed algorithm converges to the minimizer of the optimal resource allocation of interest. We also showed that if the local cost functions all have globally Lipschitz gradients, then the convergence is exponential. Finally we showed that if the communication topology is a connected graph, the convergence to a point in the set of minimizers of the problem of interest can be extended to convex local cost functions. Future work includes rigorous treatment of use of exact penalty methods to extend the proposed algorithm to solve allocation problems with limited local resources at each agent. As future work we will also study the event-triggered communication implementation of our proposed algorithm and characterize its privacy preservation properties.

Acknowledgment

This work is supported by NSF CAREER award ECCS-1653838.

References

[1] A.J. Wood, F. Wollenberg, G.B. Sheble, Power Generation, Operation and Control, third ed., John Wiley, New York, 2013.
[2] A. Cherukuri, J. Cortés, Distributed generator coordination for initialization and anytime optimization in economic dispatch, IEEE Trans. Control Netw. Syst. 2 (3) (2015) 226–237.
[3] R. Madan, S. Lall, Distributed algorithms for maximum lifetime routing in wireless sensor networks, IEEE Trans. Wirel. Commun. 5 (8) (2006) 2185–2193.
[4] G.M. Heal, Planning without prices, Rev. Econom. Stud. 36 (1969) 347–362.
[5] A. Nedić, A. Ozdaglar, Distributed subgradient methods for multi-agent optimization, IEEE Trans. Automat. Control 54 (1) (2009) 48–61.
[6] B. Johansson, M. Rabi, M. Johansson, A randomized incremental subgradient method for distributed optimization in networked systems, SIAM J. Optim. 20 (3) (2009) 1157–1170.
[7] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. 3 (1) (2010) 1–122.
[8] M. Zhu, S. Martínez, On distributed convex optimization under inequality and equality constraints, IEEE Trans. Automat. Control 57 (1) (2012) 151–164.
[9] J. Duchi, A. Agarwal, M. Wainwright, Dual averaging for distributed optimization: Convergence analysis and network scaling, IEEE Trans. Automat. Control 57 (3) (2012) 592–606.
[10] J. Wang, N. Elia, A control perspective for centralized and distributed convex optimization, in: IEEE Int. Conf. on Decision and Control, Florida, USA, 2011.
[11] B. Gharesifard, J. Cortés, Distributed continuous-time convex optimization on weight-balanced digraphs, IEEE Trans. Automat. Control 59 (3) (2014) 781–786.
[12] F. Zanella, D. Varagnolo, A. Cenedese, G. Pillonetto, L. Schenato, Newton-Raphson consensus for distributed convex optimization, in: IEEE Int. Conf. on Decision and Control, Florida, USA, 2011, pp. 5917–5922.
[13] J. Lu, C.Y. Tang, Zero-gradient-sum algorithms for distributed convex optimization: The continuous-time case, IEEE Trans. Automat. Control 57 (9) (2012) 2348–2354.
[14] S.S. Kia, J. Cortés, S. Martínez, Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication, Automatica 55 (2015) 254–264.
[15] L. Xiao, S. Boyd, Optimal scaling of a gradient method for distributed resource allocation, J. Optim. Theory Appl. 129 (3) (2006) 469–488.

[16] A. Cherukuri, J. Cortés, Initialization-free distributed coordination for economic dispatch under varying loads and generator commitment, Automatica 74 (2016) 183–193.

[17] P.V. Kokotović, H.K. Khalil (Eds.), Singular Perturbation Methods in Control: Analysis and Design, SIAM, 1999.

[18] S.S. Kia, J. Cortés, S. Martínez, Singularly perturbed algorithms for dynamic average consensus, in: European Control Conference, Zürich, Switzerland, 2013, pp. 1758–1763.

[19] S.S. Kia, A distributed dynamical solver for an optimal resource allocation problem over networked systems, in: IEEE Int. Conf. on Decision and Control, Osaka, Japan, 2015.

[20] S.S. Kia, J. Cortés, S. Martínez, Dynamic average consensus under limited control authority and privacy requirements, Internat. J. Robust Nonlinear Control 25 (13) (2015) 1941–1966.

[21] F. Bullo, J. Cortés, S. Martínez, Distributed Control of Robotic Networks, in: Applied Mathematics Series, Princeton University Press, 2009.

[22] H. Khalil, Nonlinear Systems, Prentice Hall, 2002.

[23] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[24] K.J. Arrow, L. Hurwicz, H. Uzawa, Studies in Linear and Nonlinear Programming, Stanford University Press, 1958.

[25] J.A. Jacquez, Compartmental analysis in biology and medicine, 1985.

[26] D.S. Bernstein, S.P. Bhat, Nonnegativity, reducibility, and semistability of mass action kinetics, in: IEEE Int. Conf. on Decision and Control, vol. 3, Phoenix, AZ, 1999, pp. 2206–2211.

[27] R. Olfati-Saber, J.A. Fax, R.M. Murray, Consensus and cooperation in networked multi-agent systems, in: Proceedings of the IEEE vol. 95, no.1, 2007, pp. 215–233.

[28] W. Haddad, V. Chellaboina, Nonlinear Dynamical Systems and Control, Princeton University Press, Princeton, NJ, 2008.

[29] B. Mohar, Eigenvalues, diameter, and mean distance in graphs, Graphs Combin. 7 (1991) 53–64.

[30] P. Yang, R.A. Freeman, G.J. Gordon, K.M. Lynch, S.S. Srinivasa, R. Sukthankar, Decentralized estimation and control of graph connectivity for mobile sensor networks, Automatica 46 (2) (2010) 390–396.