Systems & Control Letters 83 (2015) 45–52

Contents lists available at ScienceDirect

Systems & Control Letters

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

Distributed gradient algorithm for constrained optimization with application to load sharing in power systems[☆]

Peng Yi[a], Yiguang Hong[a,*], Feng Liu[b]

[a] Key Lab of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China

[b] Electrical Engineering Department, Tsinghua University, China

ARTICLE INFO

Article history:
Received 15 September 2014
Received in revised form
13 May 2015
Accepted 25 June 2015
Available online 14 July 2015

Keywords:
Distributed optimization
Constrained optimization
Gradient flow
Multi-agent system
Load sharing control

ABSTRACT

In this paper, a distributed constrained optimization problem is discussed to achieve the optimal point of the sum of agents' local objective functions while satisfying local constraints. Here neither the local objective function nor local constraint functions of each agent can be shared with other agents. To solve the problem, a novel distributed continuous-time algorithm is proposed by using the KKT condition combined with the Lagrangian multiplier method, and the convergence is proved with the help of Lyapunov functions and an invariance principle for hybrid systems. Furthermore, this distributed algorithm is applied to optimal load sharing control problem in power systems. Both theoretical and numerical results show that the optimal load sharing can be achieved within both generation and delivering constraints in a distributed way.

1. Introduction

Recent years have witnessed increasing research attention on distributed optimization and its applications in many engineering systems (see [1–10]). In fact, distributed optimization algorithms with a global objective function as the sum of agents' local objective functions have been proposed in the multi-agent systems where each agent's local objective/constraint function cannot be known or shared by other agents, maybe due to the privacy concern, computational burden, or communication cost/failure.

Actually, optimization problems often involve certain constraints, and great efforts keep devoted to solve the constrained optimization problems in a distributed way (see [2–8]). Projection-based distributed algorithm was proposed in [2], and further investigated in [5] and [7] for set constrained optimization. Lagrangian multiplier method was investigated in [4], while a penalty-based method was proposed in [8], both for function constrained problems. Meanwhile, dual decomposition was applied to separable problems with affine constraints in [3,6]. However, those results (with discrete-time algorithms) mainly addressed the problems where all the agents have the same constraints, which may be restrictive in some situations.

The continuous-time dynamics for distributed optimization attract more and more attention by taking advantage of the well-developed continuous-time control techniques and by concerning the implementation of physical systems (see [11–17]). The continuous-time optimization algorithm was studied in the seminal work [18] and then investigated with various backgrounds (referring to [19,20]). Recently, there have been discussions on continuous-time distributed optimization. For example, a continuous-time dynamics was proposed to show connectivity conditions for the convex set intersection computation in [12]. A second-order distributed dynamics was proposed to solve an unconstrained optimization in [11], while a similar algorithm was also constructed with non-smooth objective functions in [16]. Moreover, a distributed optimization algorithm based on proportional–integral control was given in [15], and later internal model principle was employed to achieve exact optimization with capability of rejecting external disturbance in [13]. However, to our knowledge, very few continuous-time distributed algorithms for constrained optimization have ever been documented.

The distributed load sharing optimization problem has been widely investigated in power systems (see [21–24]). It aims to find the optimal generation allocation to share the loads within both the generation and the transmission capacity bounds, which can be formulated as a class of distributed constrained optimization

☆ This work is supported partially by NNSF of China under Grants. 61174071 and 61333001, and partially by 973 Program under Grant 2014CB845301. This work is also partly supported by the Special grand from EPRI of China (XTB51201303968).

* Corresponding author.
E-mail addresses: yipeng@amss.ac.cn (P. Yi), yghong@iss.ac.cn (Y. Hong), lfeng@mail.tsinghua.edu.cn (F. Liu).

http://dx.doi.org/10.1016/j.sysconle.2015.06.006
0167-6911/

problems. [21] provided an interesting insight that the frequency and power flow dynamics are closely related with a primal–dual optimization algorithm. [22–24] considered the distributed load sharing control with or without generation capacity constraints. However, such works have not taken the transmission line capacity constraints into consideration, which is crucial for the practicality of the load sharing control. Theoretically, the distributed load sharing optimization considering both the generation and the transmission constraints can lead to a new class of distributed constrained optimization problem which is the focus of this paper.

In this paper, we study a distributed optimization problem for agents with their local inequality constraints, and its application to the load sharing optimization. Different from the existing results such as those given in [2–8] and [25,11–17], our distributed algorithm enables the agents to find the optimal point with respect to the sum of the local objective functions while satisfying all the local constraints. Note that the optimal solution must be within the intersection set of each agent's local private feasible set specified by the local constraints, while neither local objective function nor local constraint functions of each agent can be known or shared by other agents.

To solve the complicated optimization problem, we propose a novel continuous-time distributed algorithm by using the KKT condition and saddle point property, and analyze the algorithm with constructed Lyapunov functions and hybrid LaSalle invariance principle. Then we apply the proposed algorithm to the optimal load sharing problem considering both the generation limits of the power generators and the delivering limits of the transmission lines. This extends the existing results presented in [21–24], leading to a more practical optimal distributed load sharing control. We also provide the simulation experiments to show the effectiveness of the proposed method.

The paper is organized as follows. The distributed constrained optimization problem is formulated in Section 2, while the continuous-time algorithm is proposed in Section 3. Then the convergence of the algorithm is proved along with a numerical experiment in Section 4. Moreover, the application to the distributed load sharing optimization in power systems is shown in Section 5. Finally, the concluding remarks are given in Section 6.

Notations: Denote $\mathbf{1}_m = (1, \dots, 1)^T \in \mathbf{R}^m$ and $\mathbf{0}_m = (0, \dots, 0)^T \in \mathbf{R}^m$. For a column vector $x \in \mathbf{R}^m$, x^T denotes its transpose. I_n denotes the identity matrix in $\mathbf{R}^{n\times n}$. For a matrix $A = [a_{ij}]$, a_{ij} or A_{ij} stands for the matrix entry in the ith row and jth column of A.

2. Problem formulation

In this section, we give the formulation of a distributed optimization problem with local inequality constraints.

Consider a group of agents, $\mathcal{N} = \{1, \dots, N\}$, where each has **local objective function** $f_i(x)$ and **local inequality constraints** $g_j^i(x) \le 0, j = 1, \dots, J^i$. The agents need to optimize the sum of their local objective functions $f_i(x)$ under all the agents' local constraints. Since both the local objective functions $f_i(x)$ and local constraints $g_j^i(x), j = 1, \dots, J^i$ are only known by agent i and cannot be shared with other agents, the optimization has to be achieved with the cooperation of all the agents in a distributed way. To be strict, we consider

Problem 1.

$$\min f(x), \quad f(x) = \sum_{i=1}^{N} f_i(x)$$
$$\text{subject to} \quad g_j^i(x) \le 0, \quad j = 1, \dots, J^i, i = 1, \dots, N,$$

where $x \in \mathbf{R}^m$ is the decision variable, and $f_i(x), g_j^i(x), j = 1, \dots, J^i, i = 1, \dots, N$ are twice continuously differentiable and convex functions over $\mathbf{R}^m$, which are only known by agent i.

Remark 2.1. Problem 1 is different from the formulations given in [2,4,5,7,8], because each agent has local private convex feasible set $X_i = \{x \in \mathbf{R}^m | g_j^i(x) \le 0, j = 1, \dots, J^i\}$ specified by local constraint functions. Moreover, Problem 1 is not the separable one considered in [3,6] and [25], because the decision variable x is common for all the agents. Therefore, the agents need to find one common point within the intersection set $X = \bigcap_{i=1}^N X_i$ in order to minimize the sum of local objective functions, without knowing other agents' feasible sets.

Then we give the following assumptions for Problem 1:

Assumption 1. At least one of the local objective functions has positive definite Hessian $\nabla^2 f_i(x)$ over $x \in \mathbf{R}^m$.

Assumption 1 implies that at least one of the local objective functions is *strictly* convex, hence the uniqueness of the optimal solution (see Theorem 2.69 in [26]).

Assumption 2. Problem 1 has finite optimal solution, and $X = \bigcap_{i=1}^N X_i$ has nonempty interior point.

Assumption 2 guarantees Slater's constraint qualification condition, and moreover, this assumption implies that there exists finite x^* such that $x^* = \arg\min f(x) = \sum_{i=1}^N f_i(x), x \in X$, and there exists at least one interior point x_0 of X with $g_j^i(x_0) < 0, j = 1, \dots, J^i, i = 1, \dots, N$.

From Theorems 3.25, 3.26 and 3.27 in [26], we have the following result.

Lemma 2.2. *With Assumptions 1 and 2, the point x^* is the optimal solution of Problem 1 if and only if there exist Lagrangian multipliers $\lambda_{ij}^* \ge 0, j = 1, \dots, J^i, i = 1, \dots, N$ (or denoted as $\{\lambda_{ij}^*\}$) satisfying the following KKT condition.*

$$\sum_{i=1}^N \nabla f_i(x^*) + \sum_{i=1}^N \sum_{j=1}^{J^i} \lambda_{ij}^* \nabla g_j^i(x^*) = \mathbf{0}$$
$$g_j^i(x^*) \le 0, \qquad \lambda_{ij}^* g_j^i(x^*) = 0, \quad j = 1, \dots, J^i, \; i = 1, \dots, N. \tag{1}$$

Furthermore, the set of multipliers $\{\lambda_{ij}^*\}$ satisfying KKT condition (1) is closed, convex, and bounded.

By Lemma 2.2, the distributed optimization task is to cooperatively find the point x^* with multipliers $\{\lambda_{ij}^*\}$ to satisfy (1).

In the multi-agent network, each agent can exchange information only with some neighbor agents. The network topology can be described by a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with $\mathcal{N} = \{1, \dots, N\}$ representing the agents set and $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ containing all the information interactions between agents. If agent i can get information from agent j, then $(j, i) \in \mathcal{E}$. The graph $\mathcal{G}$ is undirected when $(i, j) \in \mathcal{E}$ if and only if $(j, i) \in \mathcal{E}$. A path of graph $\mathcal{G}$ is a sequence of distinct agents in $\mathcal{N}$ such that any consecutive agents in the sequence corresponding to an edge of the graph $\mathcal{G}$. Agent j is said to be connected to agent i if there is a path from j to i. Graph $\mathcal{G}$ is said to be connected if any two agents are connected. Define the adjacency matrix $A = [a_{ij}]$ associated with $\mathcal{G}$ with $a_{ij} = 1$ if $(j, i) \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. Then the Laplacian of graph $\mathcal{G}$ is $L = Deg - A$ with the degree matrix $Deg = diag\{\sum_{j=1}^N a_{1j}, \dots, \sum_{j=1}^N a_{Nj}\}$. More details about graph theory for multi-agent network can be found in [27]. The following assumption is about the connectivity of graph $\mathcal{G}$, which guarantees that any agent's information can reach any other agents.

Assumption 3. The undirected graph $\mathcal{G}$ is connected.

Clearly, Assumption 3 implies that 0 is a simple eigenvalue of Laplacian L with the eigenspace $\{\alpha \mathbf{1}_N | \alpha \in \mathbf{R}\}$, and $L\mathbf{1}_N = \mathbf{0}_N$, $\mathbf{1}_N^T L = \mathbf{0}_N^T$ (see [27]).

3. Distributed optimization algorithm

In this section, we propose a distributed continuous-time algorithm for Problem 1.

Denote $x_i \in \mathbf{R}^m$ as agent i's estimation of the optimal solution x^*, and $\bar{x} = (x_1^T, \ldots, x_N^T)^T$. Then we consider the following problem:

Problem 2.

$$\begin{aligned} &\min\ \bar{f}(\bar{x}) = \sum_{i=1}^N f_i(x_i) + \frac{1}{2}\bar{x}^T L \otimes I_m \bar{x} \\ &\text{subject to}\quad L \otimes I_m \bar{x} = \mathbf{0}_{mN}, \\ &\qquad g_j^i(x_i) \le 0, \quad j = 1, \ldots, J^i, i = 1, \ldots, N \end{aligned}$$

where L is the Laplacian of graph $\mathcal{G}$.

Remark 3.1. The equality constraint with Laplacian matrix L ensures that all the estimations must lie in the null space of L, which is exactly the consensus space $\{x_1 = x_2 =, \ldots, = x_N\}$. The quadratic penalty term in the objective function utilizes the augmented Lagrangian method (see section 4.7 of [26]) for the consensus, and it plays a damping role in the algorithm (as in [15]).

Then we show a relationship between Problems 1 and 2.

Lemma 3.2. *Suppose Assumptions 1–3 hold. Then Problem 2 has the optimal solution as $\bar{x}^* = (x^*, \ldots, x^*)$ where x^* is the optimal solution of Problem 1.*

Denote $v_i \in \mathbf{R}_m$ as the multiplier for constraint $\sum_{j=1}^N a_{ij}(x_i - x_j) = \mathbf{0}_m$, and $\lambda_{ij} \ge 0$ as the multiplier for constraint $g_j^i(x_i) \le 0$. For the same reason of Lemma 2.2, $\bar{x}^* = (x_1^*, \ldots, x_N^*)$ is the optimal solution of Problem 2 if and only if there exist Lagrangian multipliers $v_i^*, \lambda_{ij}^* \ge 0, j = 1, \ldots, J^i, i = 1, \ldots, N$ (denoting $\bar{v}^* = (v_1^{*T}, \ldots, v_N^{*T})^T$ and $\bar{\lambda}^* = (\lambda_{ij}^*, j = 1 \ldots, J^i, i = 1, \ldots, N)$ for simplicity in the sequel) such that the following KKT condition holds:

$$\begin{aligned} &\nabla f_i(x_i^*) + \sum_{j=1}^N a_{ij}(x_i^* - x_j^*) + \sum_{j=1}^N a_{ij}(v_i^* - v_j^*) \\ &\quad + \sum_{j=1}^{J^i} \lambda_{ij}^* \nabla g_j^i(x_i^*) = \mathbf{0}, \quad i = 1, \ldots, N; \\ &g_j^i(x_i^*) \le 0, \qquad \lambda_{ij}^* g_j^i(x_i^*) = 0, \quad j = 1, \ldots, J^i,\ i = 1, \ldots, N; \\ &L \otimes I_m \bar{x}^* = \mathbf{0}_{mN}. \end{aligned} \tag{2}$$

Remark 3.3. Under Assumptions 2 and 3, Slater's constraint qualification condition also holds for Problem 2. There exist multipliers $(\bar{v}^*, \bar{\lambda}^*)$ satisfying KKT condition (2) and belonging to a compact convex set, from Theorems 3.25, 3.26 and 3.27 in [26].

Since graph $\mathcal{G}$ is connected, KKT condition (2) is also equivalent to the following one:

$$\begin{aligned} &\nabla f_i(x^*) + \sum_{j=1}^N a_{ij}(v_i^* - v_j^*) + \sum_{j=1}^{J^i} \lambda_{ij}^* \nabla g_j^i(x^*) = \mathbf{0}; \\ &\lambda_{ij}^* \ge 0, \qquad g_j^i(x^*) \le 0, \qquad \lambda_{ij}^* g_j^i(x^*) = 0, \quad j = 1, \ldots, J^i. \end{aligned} \tag{3}$$

Hence, if all the agents can find the point $\bar{x}^*$ with multiplier $\bar{v}^*, \bar{\lambda}^*$ satisfying (3) in a distributed way, then all the agents' estimations converge to the same optimal solution of Problem 1.

We design the distributed algorithm by the observation that the point satisfying (3) is also the saddle point of the following Lagrangian function.

$$\begin{aligned} \mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}) = &\sum_{i=1}^N f_i(x_i) + \bar{v}^T L \otimes I_m \bar{x} + \frac{1}{2}\bar{x}^T L \otimes I_m \bar{x} \\ &+ \sum_{i=1}^N \sum_{j=1}^{J^i} \lambda_{ij} g_j^i(x_i) \quad \lambda_{ij} \ge 0. \end{aligned} \tag{4}$$

By the Saddle Point Theorem (Theorem 4.7 in [26]), we have:

Lemma 3.4. *Given Assumptions 1–3, an optimal solution $(x_1^*, \ldots, x_N^*)$ of Problem 2 satisfies the KKT condition (3) with Lagrangian multiplier $\bar{v}^*, \bar{\lambda}^*$ if and only if $(\bar{x}^*, \bar{v}^*, \bar{\lambda}^*)$ is a saddle point of the Lagrangian function $\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda})$ in (4).*

To find one saddle point of Lagrangian function (4) with a continuous-time algorithm, we adopt the primal–dual gradient dynamics as follows, which was first proposed in a centralized version in [18] and further investigated in [25],

$$\begin{aligned} &\dot{x}_i = -\nabla_{x_i} \mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}); \\ &\dot{v}_i = \nabla_{v_i} \mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}), \qquad \dot{\lambda}_{ij} = [\nabla_{\lambda_{ij}} \mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda})]^+_{\lambda_{ij}}. \end{aligned} \tag{5}$$

Hence, the dynamics for agent i is:

$$\begin{aligned} \dot{x}_i = &-\nabla f_i(x_i) - \sum_{j=1}^N a_{ij}(x_i - x_j) \\ &- \sum_{j=1}^N a_{ij}(v_i - v_j) - \sum_{j=1}^{J^i} \lambda_{ij} \nabla g_j^i(x_i); \\ \dot{v}_i = &\sum_{j=1}^N a_{ij}(x_i - x_j); \\ \dot{\lambda}_{ij} = &[g_j^i(x_i)]^+_{\lambda_{ij}}, \quad j = 1, \ldots, J^i. \end{aligned} \tag{6}$$

Here $[p]^+_\lambda = p$ if $p > 0$ or $\lambda > 0$, and $[p]^+_\lambda = 0$ otherwise. Notice that $x(t) \ge 0$, $\forall t \ge 0$ for the dynamics $\dot{x} = [f(x)]^+_x$ with the initial condition $x(0) \ge 0$. Thus, the multipliers in $\bar{\lambda}$ are nonnegative all the time due to (6).

In our formulation, agent i has its private variables $x_i, v_i, \lambda_{ij}, j = 1, \ldots, J^i$ in (6), and only needs to exchange state x_i and multiplier v_i with its neighbor agents through graph $\mathcal{G}$. Therefore, the algorithm (6) is **fully distributed** because each agent only needs to manipulate its private local objective function and local constraint functions, and only needs partial information of its neighbors.

Remark 3.5. Note that (6) contains an additional second-order consensus dynamics to ensure all the agents to reach the same optimal point, differing from that given in [25]. Moreover, if there were no constraints in our problem formulation, our algorithm would be consistent with those algorithms for the unconstrained optimization in [11,15–17].

4. The convergence analysis

In this section, we prove the convergence of algorithm (6), and then show an illustrative simulation example. Without loss of generality, we assume the dimension of the decision variable $x \in \mathbf{R}^m$ to be $m = 1$ here.

Let us first analyze the equilibrium point $(\bar{x}^*, \bar{v}^*, \bar{\lambda}^*)$ of (6). Since graph $\mathcal{G}$ is connected, $\dot{\bar{v}} = 0$ yields $x_1^* = x_2^* =, \dots, = x_N^* = x^*$. When $[g_j^i(x^*)]^+_{\lambda_{ij}^*} = 0$, we have

$$\lambda_{ij}^* g_j^i(x^*) = 0, \quad g_j^i(x^*) \le 0. \tag{7}$$

From the first equation of (6), we get

$$\nabla f_i(x^*) + \sum_{j=1}^{N} a_{ij}(v_i^* - v_j^*) + \sum_{j=1}^{J^i} \lambda_{ij}^* \nabla g_j^i(x^*) = \mathbf{0}, \quad i = 1, \dots, N. \tag{8}$$

Hence, the equilibrium point of (6) satisfies KKT condition (3). Moreover, adding those N equations in (8) and combined with (7), we obtain KKT condition (1) since the graph $\mathcal{G}$ is also undirected. Therefore, the equilibrium point of (6) contains the optimal solution to Problem 1.

We will prove that all the agents converge to the set of equilibrium points of (6) satisfying $\dot{\bar{x}} = \mathbf{0}$, $\dot{\bar{v}} = \mathbf{0}$, $\dot{\bar{\lambda}} = \mathbf{0}$, and therefore, find the same optimal solution to Problem 1.

At first, the following result shows that the trajectories are bounded.

Lemma 4.1. *Under Assumptions 1–3, the trajectories of the dynamics (6) with any finite initial points are bounded.*

Proof. Define the function

$$W(\bar{x}, \bar{v}, \bar{\lambda}) = \sum_{i=1}^{N} \frac{1}{2}\left[(x_i - x_i^*)^2 + (v_i - v_i^*)^2 + \sum_{j=1}^{J^i}(\lambda_{ij} - \lambda_{ij}^*)^2\right]$$

where $(\bar{x}^*, \bar{v}^*, \bar{\lambda}^*)$ is one finite point satisfying KKT condition (2).

Then along the trajectories of the dynamics (6),

$$\begin{aligned}\frac{dW(\bar{x}, \bar{v}, \bar{\lambda})}{dt} &= \sum_{i=1}^{N} -(x_i - x_i^*)\nabla_{x_i}\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}) \\ &\quad + (v_i - v_i^*)\nabla_{v_i}\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}) + \sum_{j=1}^{J^i}(\lambda_{ij} - \lambda_{ij}^*)[\nabla_{\lambda_{ij}}\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda})]^+_{\lambda_{ij}} \\ &\le \sum_{i=1}^{N}\Bigg[-(x_i - x_i^*)\nabla_{x_i}\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}) + (v_i - v_i^*)\nabla_{v_i}\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}) \\ &\quad + \sum_{j=1}^{J^i}(\lambda_{ij} - \lambda_{ij}^*)\nabla_{\lambda_{ij}}\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda})\Bigg]. \end{aligned} \tag{9}$$

In the last step of (9), if the projection is active at some constraint index j for agent i, $(\lambda_{ij} - \lambda_{ij}^*)g_j^i(x_i) \ge 0$, due to $g_j^i(x_i) < 0$, $\lambda_{ij} = 0$ and $\lambda_{ij}^* \ge 0$. Because $\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda})$ is convex in x_i and concave in λ_{ij}, v_i (in fact, linear in λ_{ij}, v_i), $\sum_{i=1}^{N}[-(x_i - x_i^*)\nabla_{x_i}\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda})] \le \mathcal{L}(\bar{x}^*, \bar{v}, \bar{\lambda}) - \mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda})$, and $((\bar{v}, \bar{\lambda}) - (\bar{v}^*, \bar{\lambda}^*))^T(\nabla_{\bar{v}}\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}), \nabla_{\bar{\lambda}}\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda})) \le \mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}) - \mathcal{L}(\bar{x}, \bar{\lambda}^*, \bar{v}^*)$.

Since $(\bar{x}^*, \bar{v}^*, \bar{\lambda}^*)$ is one saddle point of the Lagrangian function $\mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda})$ by Lemma 3.4, after manipulation, we have

$$\begin{aligned}\frac{dW(\bar{x}, \bar{v}, \bar{\lambda})}{dt} &\le \mathcal{L}(\bar{x}^*, \bar{\lambda}, \bar{v}) - \mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}) \\ &\quad + \mathcal{L}(\bar{x}, \bar{v}, \bar{\lambda}) - \mathcal{L}(\bar{x}, \bar{\lambda}^*, \bar{v}^*) \\ &\le \mathcal{L}(\bar{x}^*, \bar{\lambda}, \bar{v}) - \mathcal{L}(\bar{x}^*, \bar{v}^*, \bar{\lambda}^*) \\ &\quad + \mathcal{L}(\bar{x}^*, \bar{v}^*, \bar{\lambda}^*) - \mathcal{L}(\bar{x}, \bar{\lambda}^*, \bar{v}^*) \le 0. \end{aligned} \tag{10}$$

From Assumption 2 and Remark 3.3, $(\bar{x}^*, \bar{v}^*, \bar{\lambda}^*)$ is a finite point. Thus, there is a positive invariant compact set for algorithm (6): $\{(\bar{x}, \bar{v}, \bar{\lambda}) | W(\bar{x}, \bar{v}, \bar{\lambda}) \le W(\bar{x}(0), \bar{v}(0), \bar{\lambda}(0))\}$, which also implies that the trajectories are bounded. □

To show that $\bar{x}, \bar{v}, \bar{\lambda}$ converge to the set of points satisfying $\dot{\bar{x}} = \mathbf{0}$, $\dot{\bar{v}} = \mathbf{0}$, $\dot{\bar{\lambda}} = \mathbf{0}$, we will adopt a LaSalle invariance principle of hybrid systems (see [28] and [25]). Define an index set of local constraint functions for agent i as $\sigma_i = \{j | \lambda_{ij} = 0, g_j^i(x_i) < 0\}$. The constraints corresponding to the set σ_i are irrelevant with the optimization for agent i. Take $\sigma = \{\sigma_1, \dots, \sigma_N\}$, and then different σ's indicate different multipliers dynamics. Hence, algorithm (6) can be regarded as a hybrid system, where σ indicates which dynamics the multi-agent system is performing with. Because the numbers of the constraints and agents are finite, there are only a finite number of different index sets σ.

Construct a Lyapunov function:

$$V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma) = \frac{1}{2}\sum_{i=1}^{N}\left\{\dot{x}_i^2 + \dot{v}_i^2 + \sum_{j\notin\sigma_i}\dot{\lambda}_{ij}^2\right\}. \tag{11}$$

Since σ is totally determined by the state $(\bar{x}, \bar{v}, \bar{\lambda})$ and $\dot{\lambda}_{ij} = 0$ for $j \in \sigma_i$, Lyapunov function (11) only depends on $(\bar{x}, \bar{v}, \bar{\lambda})$. When the index set σ changes as the changing of the state $(\bar{x}, \bar{v}, \bar{\lambda})$ (taking some $g_i^j(x_i) \le 0$ leads λ_{ij} decreasing to zero), the function $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)$ may be discontinuous. However, the following result shows that $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)$ is non-increasing along the dynamics (6).

Lemma 4.2. *With Assumptions 1–3, the function $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)$ in (11) is non-increasing along the dynamics (6) all the time.*

Proof. With fixed σ, since the matrix L is symmetric,

$$\begin{aligned}\frac{dV}{dt} &= \sum_{i=1}^{N}\left\{\dot{x}_i\ddot{x}_i + \dot{v}_i\ddot{v}_i + \sum_{j\notin\sigma_i}\dot{\lambda}_{ij}\ddot{\lambda}_{ij}\right\} \\ &= \sum_{i=1}^{N}\Bigg(\dot{x}_i\Bigg\{-\nabla^2 f_i(x_i)\dot{x}_i - \sum_{j=1}^{N}a_{ij}(\dot{x}_i - \dot{x}_j) - \sum_{j=1}^{N}a_{ij}(\dot{v}_i - \dot{v}_j) \\ &\quad - \sum_{j\notin\sigma_i}[\nabla g_j^i(x)\dot{\lambda}_{ij} + \lambda_{ij}\nabla^2 g_j^i(x_i)\dot{x}_i]\Bigg\} \\ &\quad + \dot{v}_i\left(\sum_{j=1}^{N}a_{ij}(\dot{x}_i - \dot{x}_j)\right) + \sum_{j\notin\sigma_i}\dot{\lambda}_{ij}\nabla g_j^i(x)\dot{x}_i\Bigg) \\ &= \sum_{i=1}^{N}\left(-\dot{x}_i\nabla^2 f_i(x_i)\dot{x}_i - \sum_{j\notin\sigma_i}\lambda_{ij}\dot{x}_i\nabla^2 g_j^i(x_i)\dot{x}_i\right) \\ &\quad - \dot{\bar{x}}^T L\dot{\bar{x}} - \dot{\bar{x}}^T L\dot{\bar{v}} + \dot{\bar{v}}^T L\dot{\bar{x}} \\ &= -\dot{\bar{x}}^T\left(diag\{\nabla^2 f_1(x_1)\dots\nabla^2 f_N(x_N)\} + L\right)\dot{\bar{x}} \\ &\quad - \sum_{i=1}^{N}\left(\sum_{j\notin\sigma_i}\lambda_{ij}\dot{x}_i\nabla^2 g_j^i(x_i)\dot{x}_i\right) \le 0 \end{aligned} \tag{12}$$

where $\nabla^2 f_i(x_i)$ and $\nabla^2 g_j^i(x_i)$ are positive semidefinite for all $x_i \in \mathbf{R}^m$ (see Theorem 2.69 in [26]) because $f_i(x)$ and $g_j^i(x)$ are convex functions with Slater's condition, and the last step in (12) holds because the Laplacian L is positive semidefinite and $\lambda_{ij} \ge 0$.

For the state $(\bar{x}, \bar{v}, \bar{\lambda})$ is changing, the index set σ may change accordingly. The following analysis shows that the Lyapunov function $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)$ keeps non-increasing even when the index set σ changes.

(i) The first case is that the change of the states results in that some index set σ_i is reduced from time t_- to time t_+. For some agent i and corresponding constraint $g_j^i(x_i)$, the function $g_j^i(x_i)$

goes through zero from negative to positive. Then $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)$ will change from $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma(t_-))$ to $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma(t_+))$ with $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma(t_+))$ owning an extra term related to $\dot{\lambda}_{ij}^2$. However, for this term $\dot{\lambda}_{ij}^2(t_-) = \dot{\lambda}_{ij}^2(t_+) = 0$. Thus the function $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)$ is continuous at time t. With the same argument as before, $\frac{dV}{dt} \leq 0$ holds for $t < t_-$ and $t > t_+$ and $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)(t_-) = V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)(t_+)$.

(ii) The second case is that from time t_- to time t_+, for some agent i and some constraint $g_j^i(x_i)$, the corresponding multiplier λ_{ij} decreases from positive to zero. This causes function $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma(t_+))$ to lose a nonnegative term compared with $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma(t_-))$. Thus, $\frac{dV}{dt} \leq 0$ for $t < t_-$ and $t > t_+$ and $V(t_+) \leq V(t_-)$. □

Next we introduce a LaSalle invariance principle for hybrid systems obtained in [28]:

Lemma 4.3. *Suppose the hybrid dynamics (6) have a compact, positively invariant set Ω (i.e., trajectories starting in Ω stay in Ω) and a function $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)$ decreasing along trajectories in Ω. Then every trajectory in Ω converges to $\mathcal{I}$, the maximal positively invariant set within Ω with trajectory satisfying*

- $\frac{dV(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)}{dt} \equiv 0$ *in intervals of fixed σ;*
- $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma_-) = V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma_+)$ *if σ switches at time t between σ_- and σ_+.*

Based on the above analysis, we give our main result.

Theorem 4.4. *Suppose Assumptions 1–3 hold. With algorithm (6), all the agents converge to the same optimal solution of Problem 1.*

Proof. By the invariance principle given in Lemma 4.3, for a fixed σ,

$$\frac{dV(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)}{dt} \equiv 0.$$

With Assumption 1, one of the local objective functions has positive definite Hessian $\nabla^2 f_i(x)$. Then the matrix $diag\{\nabla^2 f_1(x_1), \ldots, \nabla^2 f_N(x_N)\} + L$ is positive definite. (With Assumption 3, if a nonzero vector satisfies $x^T L x = \mathbf{0}_N$, it must be of the form $x = \alpha \mathbf{1}_N$. Because $\nabla^2 f_i(x_i)$ is positive semidefinite for all i and one of $\nabla^2 f_i(x_i)$ is positive definite, the matrix $diag\{\nabla^2 f_1(x_1), \ldots, \nabla^2 f_N(x_N)\} + L$ is positive definite.) Thus, $\dot{\bar{x}} = \mathbf{0}$ from (12).

With $\dot{x}_i = 0$, we claim that all x_i must reach consensus (that is, $x_1 = x_2 = \cdots = x_N = x^*$). Otherwise, there is v_i such that v_i goes to infinity from $\dot{\bar{v}} = L\bar{x}$, because $\{\alpha \mathbf{1}_N | \alpha \in \mathbf{R}\}$ is the null space of L. This contradicts the boundedness of trajectories in Lemma 4.1, which implies $\dot{\bar{v}} = \mathbf{0}$.

If there exists $g_j^i(x^*) > 0$ at the point x^*, then the corresponding λ_{ij} goes to infinity with $\dot{\lambda}_{ij} = g_j^i(x^*)$, which also contradicts the boundedness of the trajectories in Lemma 4.1. Therefore, $g_j^i(x^*) \leq 0, j = 1, \ldots, J^i$. If there is $g_j^i(x^*) < 0$, then the corresponding $\lambda_{ij}^* = 0$. Otherwise, λ_{ij} will decrease to zero, which contradicts the continuity of function $V(\dot{\bar{x}}, \dot{\bar{v}}, \dot{\bar{\lambda}}; \sigma)$. Therefore, $\dot{\bar{\lambda}} = \mathbf{0}$.

As a result, the trajectories converge to one point such that

$$\dot{\bar{x}} = \mathbf{0}, \qquad \dot{\bar{v}} = \mathbf{0}, \qquad \dot{\bar{\lambda}} = \mathbf{0}$$

which also satisfies the KKT condition (1). Thus, all the agents converge to the same optimal solution of Problem 1. □

Here is a numerical example to illustrate our algorithm.

Example 4.5. There are five agents with local objective functions as follows:

$$f_1(y_1, y_2) = 2y_1 - 10y_2;$$
$$f_2(y_1, y_2) = 3y_1^2 \ln(y_1^2 + 1) + 2y_2^2;$$
$$f_3(y_1, y_2) = 3(y_1 - 10)^2 + (y_2 - 8)^2;$$
$$f_4(y_1, y_2) = \frac{4y_1^2}{\sqrt{2y_1^2 + 1}} + 0.1(y_1 + y_2)^2;$$
$$f_5(y_1, y_2) = (y_1 + y_2)^2 + 2(y_1 + y_2).$$

Agent 1 has local constraints as $g_1^1(y_1, y_2) = (y_1 - 1)^2 + y_2^2 - 1 \leq 0$ and $g_2^1(y_1, y_2) = y_1^2 + y_2^2 - 1 \leq 0$. Agent 2 has local constraint as $g^2(y_1, y_2) = y_1 - \frac{1}{2} \leq 0$. Agent 3 has local constraint as $g^3(y_1, y_2) = -y_2 \leq 0$. Agent 4 has local constraint as $g^4(y_1, y_2) = y_1^2 + (y_2 - 2)^2 - 4 \leq 0$, while there is no constraint for agent 5.

Notice that agent 1 has just a linear objective function, while with $f_5(y_1, y_2)$ Assumption 1 is satisfied. Moreover, different agents have different inequality constraints, which implies that they have different local feasible sets. The intersection set of the local feasible sets is shown in Fig. 1.

The information sharing graph $\mathcal{G}$ in algorithm (6) is set as a ring graph: $1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow 4 \leftrightarrow 5 \leftrightarrow 1$. Set the initial values of the five agents' states as $(-2, 4), (2, 3), (4, -3), (1, -2), (-1, 2)$, respectively. Set the initial values of the multiplies as zero. The algorithm (6) is solved with Matlab Simulink ODE1 solver.

The state trajectories of the five agents are shown in Fig. 1, and the initial points are marked with circles. It can be seen that all the agents asymptotically approach the same optimal solution $(\frac{1}{2}, \frac{\sqrt{3}}{2})$, which satisfies all the local constraints and minimizes the sum of local objective functions, without knowing other agents' constraints or feasible sets.

5. Distributed optimal load sharing control

In this section, we apply the proposed algorithm to the distributed optimal load sharing control in power systems.

5.1. Problem statement

In a power grid, the power sources are responsible to decide the most efficient generation allocation to meet the load demand within the generation capacities and transmission line delivering limits, which is referred as the optimal load sharing problem. Recent years, with the high penetration of renewable generations and deregulation of power markets, there is increasing research interest in distributed load sharing control [21–24].

Consider a transmission network $G_P = (\mathcal{N}_P, \mathcal{E}_P)$ with $\mathcal{N}_P = \{1, \ldots, n\}$ as the buses set and $\mathcal{E}_P = \{1, \ldots, m\}$ as the transmission lines set. Each pair of buses $i, k \in \mathcal{N}_P$ that are able to exchange power are connected by a transmission line $l \in \mathcal{E}_P$. After arbitrarily assigning direction to each line l as the reference power flow direction, define the incidence matrix $D \in \mathbf{R}^{n \times m}$ with $D_{il} = 1$ if line l goes to bus i and $D_{il} = -1$ if line l origins from bus i, and $D_{il} = 0$, otherwise.

Each bus has both generator to provide power and local load demand to be met. The load at bus i is constant as P_i^d, and the local cost function at bus i is $f_i(P_i^g)$ with respect to generation P_i^g. The generator at bus i has generation capacity constraint: $\underline{P}_i^g \leq P_i^g \leq \bar{P}_i^g$. The power flow v_l in transmission line l must satisfy $\underline{v}_l \leq v_l \leq \bar{v}_l$. $\underline{P}_i^g, \bar{P}_i^g$ and $\underline{v}_l, \bar{v}_l$ are constants, which should be decided with the consideration of physical systems operation limits. Then we formulate the following optimization problem to determinate the

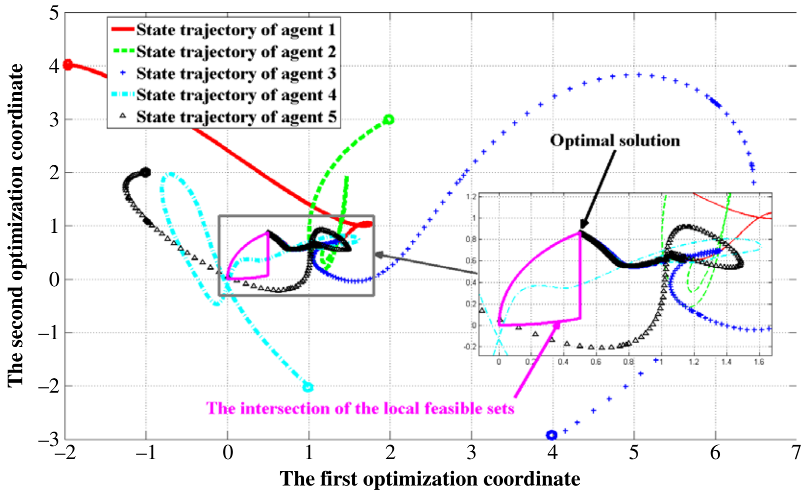

Fig. 1. The state trajectories of agents' estimations of optimal solution.

optimal generation at each bus to meet the overall load demands at the lowest cost subject to generation and transmission constraints.

$$\begin{aligned} &\min_{P^g\in R^n, v\in R^m} f(P^g, v) = \sum_{i\in\mathcal{N}_P} f_i(P_i^g) \\ &s.t \quad P^g - Dv - P^d = \mathbf{0}_n; \\ &\underline{P}_i^g \le P_i^g \le \bar{P}_i^g, \quad \forall i \in \mathcal{N}_P; \\ &\underline{v}_l \le v_l \le \bar{v}_l, \quad \forall l \in \mathcal{E}_P, \end{aligned} \tag{13}$$

where decision vector $P^g = (P_1^g, \ldots, P_n^g)$ represents the generation allocation, and $v = (v_1, \ldots, v_m)$ gives the network flow on each transmission line. $P^d = (P_1^d, \ldots, P_n^d)$ denotes the load demands at all the buses. The power balance is guaranteed by the equality constraint, which is the DC power flow equation.

In practice, problem (13) should be solved in a distributed manner, because each bus's local generation cost $f_i(P_i^g)$ and local generation capacities $\underline{P}_i^g, \bar{P}_i^g$ cannot be known by other buses. Besides that, each bus can only know the power flow capacities $\underline{v}_l, \bar{v}_l$ of the transmission lines that it is directly connected to.

Remark 5.1. The load sharing problem (13) considers both the power grid topology and transmission line capacity limitations, hence is a more realistic model than those proposed in [21–24] which only considered the overall power balance and generation capacity constraints. An equivalent formulation of problem (13) is:

$$\begin{aligned} &\min_{v\in R^m} \sum_{i\in\mathcal{N}_P} \bar{f}_i(v) = \sum_{i\in\mathcal{N}_P} f_i\left(\sum_{l=1}^m D_{il}v_l + P_i^d\right) \\ &\underline{P}_i^g \le \sum_{l=1}^m D_{il}v_l + P_i^d \le \bar{P}_i^g, \quad \forall i \in \mathcal{N}_P; \\ &\underline{v}_l \le v_l \le \bar{v}_l, \quad \forall l \in \mathcal{E}_P. \end{aligned} \tag{14}$$

Because (14) is not separable for the decision variable $v \in R^m$ as those in [21–24], the dual decomposition methods in [3,6,25] and [21–24] may not be directly applied to it, and may also fail for problem (13).

Clearly, (14) matches the formulation of Problem 1 with each bus having local generation cost function and local private inequality constraints $\underline{P}_i^g \le \sum_{l=1}^m D_{il}v_l + P_i^d \le \bar{P}_i^g$. For the security operation of power system, the optimal power flow must locate within the intersection set of all the buses' private feasible sets. In what follows, we apply the ideas and method in Section 3 to derive a distributed load sharing algorithm.

5.2. Algorithm

With the incidence matrix D and abuse of notations, we still define the incidence line set of each bus i as $\mathcal{E}_i = \{l \in \mathcal{E}_P | D_{il} \neq 0\}$ when there is no confusion. We also define the incidence bus set for each line l as $\mathcal{N}_l = \{i \in \mathcal{N}_P | D_{il} \neq 0\}$. Clearly, $|\mathcal{N}_l| = 2, \forall l \in \mathcal{E}_P$. Then we define variables $\{v_i^l, l \in \mathcal{E}_i\}$ as bus i's estimations of the optimal power flow on the transmission lines adjacent to bus i. To solve problem (13) in a distributed way, we first give the following formulation:

$$\begin{aligned} &\min_{P^g\in R^n, \tilde{v}\in R^{2m}} f(P^g, \tilde{v}) = \sum_{i\in\mathcal{N}_P} f_i(P_i^g) \\ &\quad + \frac{1}{2}\sum_{l\in\mathcal{E}_P, i,k\in\mathcal{N}_l} [v_i^l, v_k^l]^T L_c [v_i^l, v_k^l], \\ &s.t \quad P_i^g - \sum_{l\in\mathcal{E}_i} D_{il}v_i^l - P_i^d = 0, \quad \forall i \in \mathcal{N}_P; \\ &\underline{P}_i^g \le P_i^g \le \bar{P}_i^g, \quad \forall i \in \mathcal{N}_P; \\ &\underline{v}_l \le v_i^l \le \bar{v}_l, \quad \forall l \in \mathcal{E}_i,\ i \in \mathcal{N}_P; \\ &L_c[v_i^l, v_k^l] = \mathbf{0}_2, \quad i, k \in \mathcal{N}_l,\ \forall l \in \mathcal{E}_P, \end{aligned} \tag{15}$$

where $L_c = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$. Here the decision variable $\tilde{v} = (\{v_1^l\}, \{v_2^l\}, \ldots, \{v_n^l\})$ becomes $2m$ dimensional, where each sub-vector variable is $\{v_i^l\} = (v_i^l, l \in \mathcal{E}_i)$.

Remark 5.2. In (15), with the last equality constraints, the estimations of each pair of buses on the power flow of the transmission line connecting them have to reach consensus. Hence, denote (P^{g*}, v^*) and $(\bar{P}^g, \bar{v})$ as the optimal solutions of problems (13) and (15), respectively, and there must be $P^{g*} = \bar{P}^g$ and $v_l^* = \bar{v}_i^l, \forall i \in \mathcal{N}_l, \forall l \in \mathcal{E}_P$.

For equality constraint $P_i^g - \sum_{l\in\mathcal{E}_i} D_{il}v_i^l - P_i^d = 0$, γ_i is given as the associated Lagrangian multiplier. $\underline{\lambda}_i, \bar{\lambda}_i$ are given as the multipliers for $\underline{P}_i^g \le P_i^g \le \bar{P}_i^g$, respectively. $\underline{\eta}_i^l, \bar{\eta}_i^l, \theta_i^l, l \in \mathcal{E}_i$, are given as the multipliers for $\underline{v}_l \le v_i^l \le \bar{v}_l$, $v_i^l - v_{\mathcal{N}_l\backslash i}^l = 0, l \in \mathcal{E}_i$, respectively. Then the Lagrangian function of (15) is:

$$\begin{aligned} &\mathcal{L}(P_i^g, \gamma_i, \underline{\lambda}_i, \bar{\lambda}_i; v_i^l, \underline{\eta}_i^l, \bar{\eta}_i^l, \theta_i^l, l \in \mathcal{E}_i; i \in \mathcal{N}_P) \\ &\quad = \sum_{i\in\mathcal{N}_P}\left(f_i(P_i^g) + \gamma_i(P_i^g - \sum_{l\in\mathcal{E}_i} D_{il}v_i^l - P_i^d)\right) \end{aligned}$$

$$+ \sum_{i \in \mathcal{N}_P} (\underline{\lambda}_i (\underline{P}_i^g - P_i^g) + \bar{\lambda}_i (P_i^g - \bar{P}_i^g))$$
$$+ \sum_{i \in \mathcal{N}_P} \sum_{l \in \mathcal{E}_i} (\underline{\eta}_i^l (\underline{v}_l - v_i^l) + \bar{\eta}_i^l (v_i^l - \bar{v}_l))$$
$$+ \sum_{l \in \mathcal{E}_P, i, k \in \mathcal{N}_l} \left((\theta_i^l, \theta_k^l)^T L_c (v_i^l, v_k^l) + \frac{1}{2} (v_i^l, v_k^l)^T L_c (v_i^l, v_k^l) \right).$$

Applying the primal–dual gradient dynamics in (5) with simple calculation, we get the dynamics for bus i as follows:

$$\begin{aligned}
&\dot{P}_i^g = -(\nabla f_i(P_i^g) + \gamma_i - \underline{\lambda}_i + \bar{\lambda}_i),\\
&\dot{\gamma}_i = P_i^g - \sum_{l \in \mathcal{E}_i} D_{il} v_i^l - P_i^d,\\
&\dot{\underline{\lambda}}_i = [\underline{P}_i^g - P_i^g]^+_{\underline{\lambda}_i}, \qquad \dot{\bar{\lambda}}_i = [P_i^g - \bar{P}_i^g]^+_{\bar{\lambda}_i},\\
&\dot{v}_i^l = -(-\gamma_i D_{il} - \underline{\eta}_i^l + \bar{\eta}_i^l) - (\theta_i^l - \theta^l_{\mathcal{N}_l \setminus i}) - (v_i^l - v^l_{\mathcal{N}_l \setminus i}),\\
&\quad l \in \mathcal{E}_i,\\
&\dot{\theta}_i^l = v_i^l - v^l_{\mathcal{N}_l \setminus i}, \quad l \in \mathcal{E}_i\\
&\dot{\underline{\eta}}_i^l = [\underline{v}_l - v_i^l]^+_{\underline{\eta}_i^l}, \quad l \in \mathcal{E}_i, \qquad \dot{\bar{\eta}}_i^l = [v_i^l - \bar{v}_l]^+_{\bar{\eta}_i^l} \quad l \in \mathcal{E}_i.
\end{aligned} \tag{16}$$

The variables $P_i^g, \gamma_i, \underline{\lambda}_i, \bar{\lambda}_i$, and $\{v_i^l; \underline{\eta}_i^l, \bar{\eta}_i^l; \theta_i^l, l \in \mathcal{E}_i\}$ are associated with bus i. Then we observe that only the dynamics of $\dot{v}_i^l, \dot{\theta}_i^l, l \in \mathcal{E}_i$ need the state information $\{v^l_{\mathcal{N}_l \setminus i}, \theta^l_{\mathcal{N}_l \setminus i} | l \in \mathcal{E}_i\}$ from neighbor buses $\{\mathcal{N}_l \setminus i | l \in \mathcal{E}_i\}$. In other words, each pair of buses i, k connected by line l $(i, k \in \mathcal{N}_l)$ only needs to exchange information $\{v_i^l, \theta_i^l\}$ and $\{v_k^l, \theta_k^l\}$.

Remark 5.3. Since each bus i only needs to know the local generation cost function $f_i(P_i^g)$, the local generation capacity bounds $\bar{P}_i^g, \underline{P}_i^g$, and the capacities of the transmission lines it directly connects to, $\{\underline{v}_l, \bar{v}_l | l \in \mathcal{E}_i\}$, and to exchange information with its neighboring buses $\{\mathcal{N}_l \setminus i | l \in \mathcal{E}_i\}$, the algorithm (16) is a **fully distributed** algorithm for the problem (13).

Based on a similar analysis of Theorem 4.4, we have the following result:

Corollary 5.4. *Suppose that each pair of buses connected by the same transmission line can exchange information bi-directionally. Given all the cost functions $f_i(P_i^g)$ being continuously differentiable convex functions with positive definite Hessian $\nabla^2 f_i(P_i^g)$, then with algorithm (16) all the buses can find the optimal generation allocation corresponding to problem (13).*

It is worthy of noting that the load sharing model (13) with distributed algorithm (16) has the following highlights:

(i) Here we consider various physical constraints for optimal load sharing. The generator capacity constraint is crucially important for the security operation of both traditional and nontraditional generators, such as storage device and wind turbine. The transmission line capacity constraints are also critical for load sharing in low-voltage distribution network (e.g., micro-grid). With the consideration of the network topology and transmission line delivering limits, our optimization model (13) is more practical than some existing ones, such as [21–24].

(ii) With the distributed algorithm (16), all the buses can adaptively achieve secure and optimal load sharing among the distributed generators. Then optimal load sharing control is relatively easily realizable with a "peer-to-peer" architecture, in addition to the conventional "master–slave" hierarchal one. Moreover, our algorithm may also enable "plug-and-play" optimal control in power systems with scheduled or unscheduled environmental changes, including load changes, generator/load buses joining-in or leaving-out, generation cost changes, and transmission lines switches.

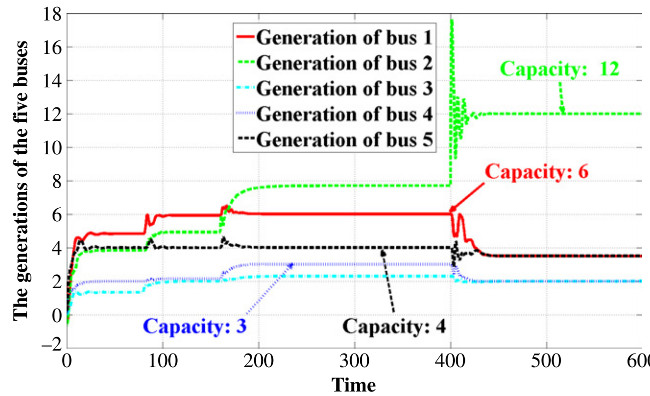

Fig. 2. The state trajectories of generation levels.

5.3. Example

For illustration, we consider a five-bus four-line power grid with the incidence matrix as

$$D = \begin{pmatrix} -1 & -1 & 0 & 0 \\ 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The initial generation cost functions of the five buses are set as follows:

$$f_1(P_1^g) = 2{P_1^g}^2 + 2P_1^g; \qquad f_2(P_2^g) = {P_2^g}^2 + 4P_2^g;$$
$$f_3(P_3^g) = 4{P_3^g}^2 + P_3^g; \qquad f_4(P_4^g) = 3{P_4^g}^2 + P_4^g;$$
$$f_5(P_5^g) = {P_5^g}^2 + 2P_5^g.$$

The generation capacity bounds of the five buses are (0, 6) p.u. (shorthand for per unit), (0, 12) p.u., (0, 3) p.u., (0, 3) p.u., (0, 4) p.u., respectively, while the capacity limits of the four transmission lines are (−10, 10), (−3, 3), (−2, 2), (−2, 2) p.u., respectively.

The initial load vector at the five buses is $P^d = [1, 4, 2, 4, 5]$ p.u. Then, we have P_3^d change to 5 p.u. at time 80 s, and P_2^d change to 8 p.u. at time 160 s. Moreover, the cost function of bus 2 changes to $f_2(P_2^g) = 0.25{P_2^g}^2 + P_2^g$ at time 400 s.

The initial values of all the variables are set with zero. The simulation results are shown in Figs. 2 and 3.

All the results after the transition process are consistent with centralized optimization results, demonstrating that algorithm (16) can correctly find the optimal solution in a distributed manner. From Fig. 2, all the buses can find the optimal generations under different load circumstances and cost functions in fair time. Fig. 3 indicates that the two buses connected to the same transmission line always reach consensus on the optimal power flow of that line. Note that all the settled generations and power flow points satisfy both the generation and the transmission capacities bounds, therefore guarantee the secure operation of the power grids.

6. Conclusions

In this paper, we first formulated a distributed optimization problem with both local objective functions and local constraints private to each agent. Then we proposed a novel continuous-time distributed algorithm for the constrained optimization with a convergence proof. Moreover, we applied the algorithm to optimal load sharing optimization in power grids with simulation analysis.

However, many promising research topics still remain to be investigated. For instance, the convergence rate and communication mechanism may deserve more research attention, and the practical

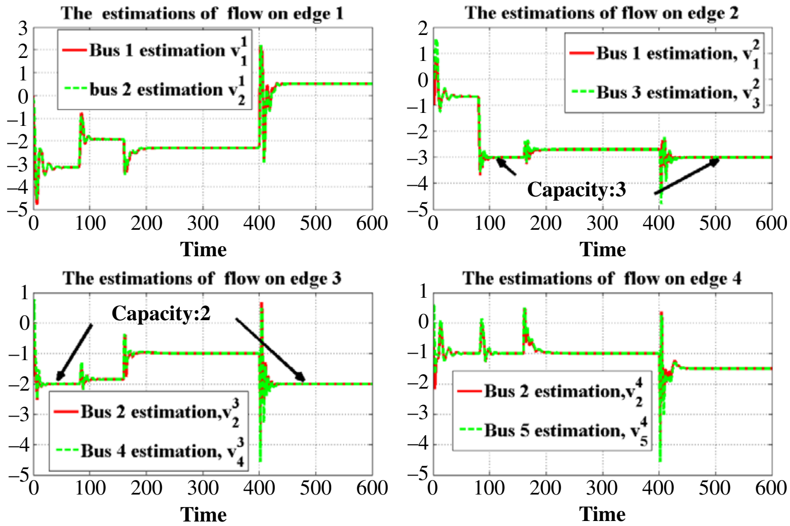

Fig. 3. The state trajectories of power flow estimations.

physical dynamics may also be considered for optimal load sharing control.

References

[1] A. Nedic, A. Ozdaglar, Distributed subgradient methods for multi-agent optimization, IEEE Trans. Automatic Control 54 (1) (2009) 48–61.
[2] A. Nedic, A. Ozdaglar, A.P. Parrilo, Constrained consensus and optimization in multi-agent networks, IEEE Trans. Automat. Control 55 (4) (2010) 922–938.
[3] I. Necoara, Y. Nesterov, F. Glineur, A Random Coordinate Descent Method on Large Optimization Problems with Linear Constraints, Technical Report, University Politehnica Bucharest, 2011, pp. 1–25.
[4] M. Zhu, S. Martinez, On distributed convex optimization under inequality and equality constraints, IEEE Trans. Automat. Control 57 (1) (2012) 151–164.
[5] P. Bianchi, J. Jakubowicz, Convergence of a multi-agent projected stochastic gradient algorithm for non-convex optimization, IEEE Trans. Automat. Control 58 (2) (2013) 391–405.
[6] I. Necoara, Random coordinate descent algorithms for multi-agent convex optimization over networks, IEEE Trans. Automat. Control 58 (8) (2013) 2001–2012.
[7] Y. Lou, G. Shi, K.H. Johansson, Y. Hong, Approximate projected consensus for convex intersection computation: Convergence analysis and critical error angle, IEEE Trans. Automat. Control 59 (7) (2014) 1722–1736.
[8] Z.J. Towfic, A.H. Sayed, Adaptive penalty-based distributed stochastic convex optimization, IEEE Trans. Signal Process. 62 (15) (2014) 3924–3938.
[9] P. Yi, Y. Hong, Quantized subgradient algorithm and data-rate analysis for distributed optimization, IEEE Trans. Control Netw. Syst. 1 (4) (2014) 380–392.
[10] A.H. Sayed, Adaptation, learning, and optimization over networks, in: Foundations and Trends in Machine Learning, Vol. 7, NOW Publishers, Boston-Delft, 2014, pp. 4–5.
[11] J. Wang, N. Elia, Control approach to distributed optimization, in: Allerton Conf. on Communications, Control and Computing, Illinois, USA, 2010, pp. 557–561.
[12] G. Shi, K.H. Johansson, Y. Hong, Reaching an optimal consensus: Dynamical systems that compute intersections of convex sets, IEEE Trans. Automat. Control 58 (3) (2013) 610–622.
[13] X. Wang, P. Yi, Y. Hong, Dynamic optimization for multi-agent systems with external disturbances, Control Theory Technol. 12 (2) (2014) 132–138.
[14] P. Yi, Y. Zhang, Y. Hong, Potential game design for a class of distributed optimization problems, J. Control Decis. 1 (2) (2014) 166–179.
[15] G. Droge, H. Kawashima, M.B. Egerstedt, Continuous-time proportional-integral distributed optimisation for networked systems, J. Control Decis. 1 (3) (2014) 191–213.
[16] B. Gharesifard, J. Cortés, Distributed continuous-time convex optimization on weight-balanced digraphs, IEEE Trans. Automat. Control 59 (3) (2014) 781–786.
[17] S. Kia, J. Cortés, S. Martinez, Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication, Automatica 55 (5) (2015) 254–264.
[18] K. Arrow, L. Huwicz, H. Uzawa, Studies in Linear and Non-linear Programming, Stanford University Press, 1958.
[19] L. Liao, H. Qi, L. Qi, Neurodynamical optimization, J. Global Optim. 28 (1) (2004) 175–195.
[20] A. Bhaya, E. Kaszkurewicz, Control Perspectives on Numerical Algrorithm and Matix Problems, SIAM, 2006.
[21] C. Zhao, U. Topcu, N. Li, S.H. Low, Design and stability of load-side primary frequency control in power systems, IEEE Trans. Automat. Control 59 (5) (2014) 1177–1189.
[22] S. Kar, G. Hug, Distributed robust economic dispatch in power systems: a consensus+ innovations approach, IEEE Power Energy Soc. Gen. Meet. (2012) 1–8.
[23] Z. Zhang, M.Y. Chow, Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid, IEEE Trans. Power Syst. 27 (4) (2012) 1761–1768.
[24] W. Zhang, W. Liu, X. Wang, L. Liu, F. Ferrese, Online optimal generation control based on constrained distributed gradient algorithm, IEEE Trans. Power Syst. 30 (1) (2015) 35–45.
[25] D. Feijer, F. Paganini, Stability of primal–dual gradient dynamics and applications to network optimization, Automatica 46 (12) (2010) 1974–1981.
[26] A.P. Ruszczynski, Nonlinear Optimization, Princeton university press, Princeton, NJ, 2006.
[27] M. Mesbahi, M. Egerstedt, Graph Theoretic Methods for Multiagent Networks, Princeton University Press, Princeton, NJ, 2010.
[28] J. Lygeros, K.H. Johansson, S.N. Simic, J. Zhang, S.S. Sastry, Dynamical properties of hybrid automata, IEEE Trans. Automat. Control 48 (1) (2003) 2–17.