

# Distributed Energy Resources Coordination over Time-varying Directed Communication Networks

Tao Yang, *Member, IEEE*, Di Wu, *Senior Member, IEEE*, Huazhen Fang, *Member, IEEE*, Wei Ren, *Fellow, IEEE*, Hong Wang, *Senior Member, IEEE*, Yiguang Hong, *Fellow, IEEE*, and Karl Henrik Johansson, *Fellow, IEEE*

**Abstract**—In this paper, we consider the optimal coordination problem for distributed energy resources (DERs) including distributed generators and energy storages. We first propose an algorithm based on the push-sum and gradient method to solve the optimal DER coordination problem in a distributed manner. In the proposed algorithm, each DER only maintains a set of variables and updates them through information exchange with a few neighboring DERs over a time-varying directed communication network. We show that the proposed distributed algorithm with appropriately chosen diminishing step-sizes solves the optimal DER coordination problem if the time-varying directed communication network is uniformly jointly strongly connected. Moreover, in order to improve the convergence speed and to reduce the communication burden, we propose an accelerated distributed algorithm with a fixed step-size. We show that the new proposed algorithm exponentially solves the optimal DER coordination problem if the cost functions satisfy an additional assumption and the selected step-size is less than a certain critical value. Both proposed distributed algorithms are validated and evaluated using the IEEE 39-bus system.

**Index Terms**—Distributed coordination, energy storage, multi-agent systems, multi-step optimization, push-sum and gradient method, smart grid.

## I. INTRODUCTION

The infrastructure that defines the U.S. power grid is based largely on pre-digital technologies developed in the first part

This work was supported in part by the Knut and Alice Wallenberg Foundation, and the Swedish Research Council. This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>). (Corresponding author: D. Wu; Tel: +1 509-375-3975.)

Some preliminary work has been presented in the conference version of the paper [1].

T. Yang is with the Department of Electrical Engineering, University of North Texas, Denton, TX 76203 USA (e-mail: Tao.Yang@unt.edu).

D. Wu is with the Pacific Northwest National Laboratory, Richland, WA 99352 USA (e-mail: di.wu@pnnl.gov).

H. Fang is with the Department of Mechanical Engineering, University of Kansas, Lawrence, KS 66045 (e-mail: fang@ku.edu).

W. Ren is with the Department of Electrical Engineering, University of California, Riverside, Riverside, CA 92521 USA (e-mail: ren@ee.ucr.edu).

H. Wang is with the Oak Ridge National Laboratory, Oak Ridge, TN 37932, USA (e-mail: wangh6@ornl.gov).

Y. Hong is with the Key Laboratory of Systems and Control, Institute of Systems Science, Chinese Academy of Science, Beijing 100190, China (e-mail: yghong@iss.ac.cn).

K. H. Johansson is with the ACCESS Linnaeus Center, School of Electrical Engineering, Royal Institute of Technology, 10044 Stockholm, Sweden (e-mail: kallej@kth.se).

of the 20th century. In subsequent decades, grid development has evolved through emphasis on safety, accessibility, and reliability to security and resiliency. The past few years have witnessed acceleration in the deployment and integration of digital smart grid sensing, communication, and control technologies that improve electric grid reliability, security, and efficiency of existing power systems at both transmission and distribution levels. Focusing on distribution system, a great effort has been made in developing distributed generation and energy storage technology. Distributed generation (DG) and energy storage (ES) are important elements of the emerging smart grid paradigm. For ease of reference, these resources are often referred to as distributed energy resources (DERs) [2]. DERs are smaller, highly flexible, and can be aggregated to provide power necessary to meet regular demand. As the electricity grid continues to be modernized, DER such as distributed generators and energy storages can help facilitate the transition of the present power grid to a smarter one.

At the high deployment level, DERs can collectively become a valuable system asset if coordinated with system needs and control processes, as they can respond very fast and are close to the loads. In order to achieve an effective deployment among DERs, one needs to properly design the coordination and control mechanism among them. One approach is through a completely centralized control strategy, where a single control center accesses the entire network's information and provides control signals to the entire system. This centralized control framework may not be effective for large-scale power networks due to performance limitations, such as a single point of failure, high communication and computational burden, limited flexibility, and lack of privacy.

Recently, an alternative distributed approach has been proposed to overcome these limitations. In particular, each DER makes a local decision based on the information received from a few neighboring DERs over the underlying communication network. Most existing distributed DER coordination studies focused on a single type of DERs. On the one hand, for DG coordination, various distributed algorithms have been proposed, among which discrete-time algorithms were given in [3]–[11] and continuous-time algorithms were presented in [12]–[14]. On the other hand, cooperative management for a network of ESs has been considered in [15], [16].

However, only a few studies considered the optimal distributed coordination of distributed generators with energy storages [17]–[21]. To coordinate DGs and ESs over multiple time periods, the authors of [17] proposed a distributed discrete-time algorithm based on the consensus and inno-

vation method, and the authors of [21] developed a distributed continuous-time algorithms based on the Laplacian gradient dynamics and dynamic average consensus. In the above studies, the charging/discharging efficiencies of energy storages were not modeled. As shown in [22] and other existing studies, the optimal charging/discharging operation and the corresponding benefits from a storage device could vary significantly with its efficiencies. The authors of [18], [19] developed distributed DER coordination strategies, where charging/discharging losses have been taken into consideration.

Note that existing studies [17]–[21] focused on the case where the communication network is time-invariant (fixed). However, in practice, the communication network topology may vary due to unexpected loss of communication links. Thus, in this paper, we consider the optimal DER coordination problem over time-varying and directed networks.

The main contributions of this paper are summarized as follows.

- Inspired by the push-sum and gradient method [23], we first propose a distributed algorithm with diminishing step-sizes and show that the proposed distributed algorithm with the properly chosen step-sizes is capable of achieving the optimal DER coordination if the time-varying directed communication network is uniformly jointly strongly connected. Compared with existing distributed DER coordination studies for undirected connected topologies [17]–[20] and strongly connected and weight-balanced directed topologies [21], this requirement is much more general since the network can be disconnected at any time instant as long as the joint graph over a period of time is strongly connected but not necessarily weight-balanced.
- The distributed algorithm proposed above, however, can be rather slow due to the diminishing step-sizes. In order to improve the convergence speed, we further develop an accelerated distributed algorithm with a fixed step-size and show that the new algorithm exponentially solves the optimal DER coordination problem if the fixed step-size is less than a certain critical value and the cost functions satisfy some additional properties. Compared with other existing algorithms with diminishing step-sizes [17]–[19], the new proposed algorithm solves the optimal DER coordination problem faster. Our proposed algorithms are discrete-time by design and readily to be implemented, while the algorithm proposed in [21] is continuous-time and requires discretization for the implementation.

The remainder of the paper is organized as follows. Section II introduces some preliminaries on graph theory and convex analysis and notations. In Section III, we formulate the optimal DER coordination problem. In Section IV, a distributed algorithm with diminishing step-sizes and an accelerated distributed algorithm with a fixed step-size are developed. Section V presents case studies and simulation results. Concluding remarks are offered in Section VI.

## II. PRELIMINARIES

In this section, we first present some background on graph theory [24], which is needed to describe the communication network among DERs. Let  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  denote a directed graph (digraph) with the set of nodes (agents)  $\mathcal{V} = \{1, \dots, N\}$  and the set of edges  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ . A directed edge from node  $i$  to node  $j$  is denoted by  $(i, j) \in \mathcal{E}$ . For notational simplicity, we assume that the digraph does not have any self loop, i.e.,  $(i, i) \notin \mathcal{E}$  for all  $i \in \mathcal{V}$  although each node  $i$  has access to its own information. A directed path from node  $i_1$  to node  $i_k$  is a sequence of nodes  $\{i_1, \dots, i_k\}$  such that  $(i_j, i_{j+1}) \in \mathcal{E}$  for  $j = 1, \dots, k-1$ . If there exists a directed path from node  $i$  to node  $j$ , then node  $j$  is said to be reachable from node  $i$ . A digraph  $\mathcal{G}$  is said to be strongly connected if every node is reachable from every other node.

In this paper, an agent is assigned to each distributed generator and energy storage in the power system. These agents exchange information according to the topology of the communication network, which may be different from the physical network, and is modeled as a time-varying directed graph  $\mathcal{G}(k) = (\mathcal{V}, \mathcal{E}(k))$ , where the edge set changes over time due to unexpected loss of communication links. All nodes that can transmit information to node  $i$  directly at time  $k$  are said to be its in-neighbors and belong to the set  $\mathcal{N}_i^{\text{in}}(k) = \{j \in \mathcal{V} \mid (j, i) \in \mathcal{E}(k)\}$ . The nodes which receive information from agent  $i$  at time  $k$  belong to the set of its out-neighbors, denoted by  $\mathcal{N}_i^{\text{out}}(k) = \{j \in \mathcal{V} \mid (i, j) \in \mathcal{E}(k)\}$ . The cardinality of  $\mathcal{N}_i^{\text{out}}(k)$  is called its out-degree at time  $k$  and is denoted by  $d_i(k) = |\mathcal{N}_i^{\text{out}}(k)|$ . The joint graph of  $\mathcal{G}(k)$  in the time interval  $[k_1, k_2)$  with  $k_1 < k_2 \leq \infty$  is denoted as  $\mathcal{G}([k_1, k_2)) = \cup_{k \in [k_1, k_2)} \mathcal{G}(k) = (\mathcal{V}, \cup_{k \in [k_1, k_2)} \mathcal{E}(k))$ . A time-varying directed network  $\mathcal{G}(k)$  is said to be *uniformly jointly strongly connected* if there exists an integer  $B > 0$  such that  $\mathcal{G}([k_0, k_0 + B))$  is strongly connected for any  $k_0 \in \mathbb{Z}_+$ , where  $\mathbb{Z}_+$  is the set of non-negative integers.

We now provide some background on basic convex analysis [25]. A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is *convex* if  $f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$  for all  $x, y \in \mathbb{R}^n$  and for all  $\theta \in (0, 1)$ . If the inequality holds for all  $x \neq y$ , then the function is *strictly convex*. A continuously differentiable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is *strongly convex* if there exists a constant  $\exists \mu > 0$  such that  $f(y) \geq f(x) + \nabla f(x)'(y - x) + \frac{\mu}{2}\|y - x\|^2$  for all  $x, y \in \mathbb{R}^n$ . A continuously differentiable function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is *smooth* if it has a Lipschitz continuous gradient, i.e., there exists a constant  $L > 0$  such that  $\|\nabla f(y) - \nabla f(x)\| \leq L\|y - x\|$  for all  $x, y \in \mathbb{R}^n$ .

*Notations:* Given a matrix  $A$ ,  $A'$  denotes its transpose, and  $A_{ij}$  denotes its  $(i, j)$ -th entry. A column vector  $x \in \mathbb{R}^n$  will be denoted by  $x = (x_1, x_2, \dots, x_n)'$ .

## III. PROBLEM FORMULATION AND MOTIVATION

Consider a power network of  $N + M$  distributed energy resources, where the first  $N$  agents are distributed generators and the last  $M$  agents are energy storages. The objective of optimal coordination is to minimize the total cost on the premise that all DERs collectively meet a given demand profile

during a finite-time horizon  $\mathcal{T} = \{1, \dots, T\}$ , where  $T$  is the number of time periods.

The total cost is the sum of DERs' costs over a number of time periods:

$$\sum_{t=1}^T \sum_{i=1}^{N+M} C_i(p_{i,t}), \quad (1)$$

where  $C_i(p_{i,t})$  is the cost function of DER  $i$  during the period  $t$  and  $p_{i,t}$  is the power from generator or storage  $i$  during period  $t$ . The power from storage is measured at the grid coupling point, and is positive when injecting power into grid, i.e., using generator convention.

Compared to most existing studies [1], [17]–[19], where cost functions are assumed to be quadratic, this paper considers general convex cost functions that satisfy the following assumption.

**Assumption 1.** For each  $i = 1, \dots, N + M$  and each  $t \in \mathcal{T}$ , the cost function  $C_i(p_{i,t}) : \mathbb{R} \rightarrow \mathbb{R}_+$  is strictly convex and continuously differentiable.

The power from the DGs and ESs together need to meet a given demand over a period of  $T$ , i.e.,

$$\sum_{i=1}^{N+M} p_{i,t} - D_t = 0, \quad \forall t \in \mathcal{T}, \quad (2)$$

where  $D_t$  is the given total demand of period  $t$ .

For each DG  $i \in \mathcal{N} := \{1, \dots, N\}$ , there are two constraints due to physical limits. The first one is the capacity limit on how much power DG  $i$  can generate at each time period.

$$p_i^{\min} \leq p_{i,t} \leq p_i^{\max}, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N}, \quad (3)$$

where  $p_i^{\min}, p_i^{\max}$  for  $i \in \mathcal{N}$  are the lower and upper power limits of generator  $i$ , respectively. The second constraint is ramping up/down constraints

$$\Delta p_i \leq p_{i,t} - p_{i,t-1} \leq \Delta \bar{p}_i, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{N}, \quad (4)$$

where  $\Delta p_i, \Delta \bar{p}_i$  are the lower and upper bounds of ramping rates of generator  $i$ , respectively.

For each ES  $i \in \mathcal{M} := \{N + 1, \dots, N + M\}$ , since ES cannot be charged and discharged at the same time, we define

$$p_{i,t} = p_{i,t}^+ - p_{i,t}^-, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{M}, \quad (5)$$

where

$$0 \leq p_{i,t}^+ \leq p_i^{\max}, \quad 0 \leq p_{i,t}^- \leq p_i^{\max}, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{M}, \quad (6)$$

and  $p_i^{\max} > 0$  is the upper bound of the power limit of ES  $i$ .

Note that due to the charging/discharging efficiencies, the rate of change of energy stored in ES is given by

$$p_{i,t}^{\text{batt}} = \frac{1}{\eta_i^+} p_{i,t}^+ - \eta_i^- p_{i,t}^-, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{M}, \quad (7)$$

where  $p_{i,t}^{\text{batt}}$  is the rate of change of energy stored in ES  $i$  at the end of period  $t$ , and  $\eta_i^+, \eta_i^- \in (0, 1)$  are discharging and charging efficiency of ES  $i$ , respectively.

The energy stored in ES  $i$  evolves according to the following dynamics:

$$E_{i,t} = E_{i,t-1} - p_{i,t}^{\text{batt}} \Delta T \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{M}, \quad (8)$$

where  $E_{i,t}$  is the energy stored in ES  $i$  at the end of time period  $t$  and  $\Delta T$  is the size of period.

The energy stored in ES  $i$  needs to be within the storage capacity, i.e.,

$$0 \leq E_{i,t} \leq E_i^{\max} \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{M}, \quad (9)$$

where  $E_i^{\max}$  is the energy capacity of ES  $i$ .

The energy stored in ES  $i$  at the end of the scheduling period is set to be equal to the initial energy state as shown in (10)

$$E_{i,T} = E_{i,0} \quad \forall i \in \mathcal{M}, \quad (10)$$

but can also be set to other feasible values.

With the above model, the optimal DER coordination problem can be formulated as the following convex optimization problem:

$$\mathbf{PP} : \min_{p_{i,t}} \sum_{t=1}^T \sum_{i=1}^{N+M} C_i(p_{i,t}), \quad (11)$$

subject to (2)–(10). For the feasibility of **PP** and for reliable power system operation, we make the following assumption.

**Assumption 2.** Assume that the demand can be served solely by generators, i.e.,

$$\sum_{i=1}^N p_i^{\min} \leq D_t \leq \sum_{i=1}^N p_i^{\max}. \quad (12)$$

**Remark 1.** Note that since a physical storage device cannot be charged and discharged at the same time, we need to ensure that either  $p_{i,t}^+$  or  $p_{i,t}^-$  needs to be zero, i.e.,

$$p_{i,t}^+ p_{i,t}^- = 0, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{M}.$$

However, when  $\eta_i^+ \eta_i^- < 1$ , there is no need to add this nonconvex constraint into the optimization problem (11), as shown in [19, Theorem 1].

**Remark 2.** Compared with existing studies for optimal DG coordination [3]–[14], the optimal coordination problem of both DGs and ESs are more challenging. Since there is only limited energy that can be stored in a storage, the operations of storages in different periods are interdependent. It is thus indispensable to formulate the optimization problem over multiple periods concurrently as given in (11). A similar problem formulation has been studied in [17], [21]. In these problem formulations, the physical constraints such as transmission line loss and power flow and transmission line flow constraints have not been considered. As discussed in [21], the design of distributed algorithms for such a problem is already challenging, therefore we leave the extension to handle other physical constraints as a future research direction.

**Remark 3.** Compared to [17], [21], the charging and discharging losses are considered in our problem formulation. These losses introduce nonlinearity to the dynamics of energy storages, which makes the optimization problem nonconvex. By introducing two lifted variables  $p_{i,t}^+$  and  $p_{i,t}^-$  for ESs in (5), we convert the nonconvex optimization problem to its convex equivalency (11).

The existing studies on DER coordination focused on fixed communication networks [17]–[21]. However, in practice, the

communication network topology may vary due to unexpected loss of communication links [23]. This motivates the study in this paper. In particular, the communication topology for DERs is modeled as a time-varying directed graph  $\mathcal{G}(k) = (\mathcal{V}, \mathcal{E}(k))$ , where the agent set is  $\mathcal{V} = \{1, 2, \dots, N + M\}$ , in which the first  $N$  agents correspond to DGs, while the last  $M$  agents correspond to ESs, and the edge set  $\mathcal{E}(k)$  models communication links among DERs, which may change over time due to unexpected loss of communication links.

We make the following mild assumption regarding the communication network.

**Assumption 3.** *The time-varying directed communication network  $\mathcal{G}(k)$  is uniformly jointly strongly connected, i.e., the joint communication network  $\mathcal{G}([k_0, k_0 + B))$  is strongly connected for any  $k_0 \in \mathbb{Z}_+$  with some integer  $B > 0$ .*

Note that the above assumption is a mild condition on the connectivity of communication topologies, since the network can be disconnected at any time instant as long as the joint graph over a period of time is strongly connected.

#### IV. MAIN RESULTS

In this section, we first present the Lagrangian-based approach for the optimal DER coordination in Section IV-A. In Section IV-B, we propose a distributed algorithm with diminishing step-sizes and show that the proposed distributed algorithm asymptotically achieves the optimal DER coordination if the time-varying directed communication network is uniformly jointly strongly connected and the cost functions are strictly convex. To improve the slow convergence due to the diminishing step-sizes, in Section IV-C, we develop a distributed algorithm with a fixed step-size and show that the new proposed algorithm exponentially realizes the optimal DER coordination under an additional assumption on the cost functions.

##### A. Lagrangian-Based Approach

For problem **PP** (11), we denote  $\Omega_{\mathcal{N},i}$  as the set of all  $p_i \in \mathbb{R}^T$  for which (3) and (4) are satisfied, where  $i \in \mathcal{N}$  and  $p_i = (p_{i,1}, p_{i,2}, \dots, p_{i,T})'$ . We also denote  $\Omega_{\mathcal{M},i}$  as the set of all  $p_i^+, p_i^- \in \mathbb{R}^T$  for which (6)–(10) are satisfied, where  $i \in \mathcal{M}$ ,  $p_i^+ = (p_{i,1}^+, p_{i,2}^+, \dots, p_{i,T}^+)'$ , and  $p_i^- = (p_{i,1}^-, p_{i,2}^-, \dots, p_{i,T}^-)'$ . Note that both sets  $\Omega_{\mathcal{N},i}$  and  $\Omega_{\mathcal{M},i}$  are convex polyhedra since all constraints are affine. Moreover, the objective function in (11) is also strictly convex given Assumption 1. Therefore, the primal problem **PP** (11) has a unique minimizer, for which we denote as  $p_i^*$ . Moreover, there is zero duality gap and the dual optimal set is nonempty [26].

In order to solve the optimal DER coordination problem **PP** in (11), we can instead solve its dual problem with respect to constraint (2), which couples the operations of all DERs. Following the similar analysis as those in [1], [10], [27], [28], it is not hard to obtain the dual problem of **PP** as follows:

$$\mathbf{DP} : \max_{\lambda \in \mathbb{R}^T} \sum_{i=1}^{N+M} \Phi_i(\lambda), \quad (13)$$

where  $\lambda = (\lambda_1, \dots, \lambda_T)'$  and  $\lambda_i$  for  $i \in \mathcal{T}$  are dual variables associated with the power balance constraints (2),

$$\Phi_i(\lambda) = \min_{p_i \in \Omega_i} \sum_{t=1}^T C_i(p_{i,t}) - \lambda'(p_i - D^i), \quad i \in \mathcal{V}, \quad (14)$$

where  $\Omega_i = \Omega_{\mathcal{N},i}$  for  $i \in \mathcal{N}$ , and  $\Omega_i = \{p_i = p_i^+ - p_i^- \mid \{p_i^+, p_i^-\} \in \Omega_{\mathcal{M},i}\}$  for  $i \in \mathcal{M}$ .

Note that  $D^i \in \mathbb{R}^T$  are virtual local demands at DER  $i$  for all the periods such that  $\sum_{i=1}^{N+M} D^i = D$ , where  $D = (D_1, \dots, D_T)'$ . One choice is that  $D^i = \frac{1}{N+M}D$  for all  $i \in \mathcal{V}$ . Note that  $D_i$ 's has no physical meaning. The purpose of introducing these parameters is to convert the dual problem into the maximization problem of  $N + M$  local functions as shown in (13) such that distributed algorithms can be developed to solve the dual problem.

Under Assumption 1, for any given  $\lambda$ , the right-hand side of (14) has a unique minimizer given by

$$p_i(\lambda) = \arg \min_{p_i \in \Omega_i} \sum_{t=1}^T C_i(p_{i,t}) - \lambda' p_i, \quad i \in \mathcal{V}. \quad (15)$$

Furthermore, there exists at least one optimal solution to the dual problem (13), and the unique optimal solution of the primal problem (11) can be obtained by solving the local optimization problem (14), which is given by

$$p_i^* = p_i(\lambda^*) = \arg \min_{p_i \in \Omega_i} \sum_{t=1}^T C_i(p_{i,t}) - \lambda^{*'} p_i, \quad i \in \mathcal{V}, \quad (16)$$

where  $\lambda^* = (\lambda_1^*, \dots, \lambda_T^*)'$  is any dual optimal solution.

Note that  $C_i(\cdot)$  is strictly convex and the constrained set is a bounded polyhedron, it then follows from [29, pp. 669] that the function  $\Phi_i(\lambda)$  is continuously differentiable and the gradient is given by

$$\nabla \Phi_i(\lambda) = -(p_i(\lambda) - D^i). \quad (17)$$

In this paper, we aim to solve the DER coordination problem (11) by solving its dual problem (13) in a distributed manner. Compared with the existing literature [17]–[21], which focused on fixed communication networks, we aim to develop distributed algorithms for solving the DER coordination problem over time-varying directed networks.

##### B. Distributed Push-Sum and Gradient Based Algorithm

To address the challenges of time-varying directed networks, we propose a distributed algorithm based on the push-sum and (sub)gradient method developed recently in [23].

More specifically, at each iteration  $k \in \mathbb{Z}_+$ , each DER (agent)  $i \in \mathcal{V}$  maintains five vectors, namely,  $w_i(k)$ ,  $y_i(k)$ ,  $\lambda_i(k)$ ,  $p_i(k)$ ,  $v_i(k) \in \mathbb{R}^T$ , where  $w_i(k)$ ,  $y_i(k)$ , and  $v_i(k)$  are auxiliary vectors, and  $p_i(k)$  and  $\lambda_i(k)$  are agent  $i$ 's estimations of the primal solution (optimal generations of DGs and ESs) and the dual optimal solution (optimal incremental cost), respectively. For example,  $\lambda_i = (\lambda_{i,1}, \dots, \lambda_{i,T})'$ , where  $\lambda_{i,t}$  for  $i \in \mathcal{V}$  and  $t \in \mathcal{T}$  is agent  $i$ 's estimate of the optimal incremental cost (marginal price) for period  $t$ . At each iteration



$k \in \mathbb{Z}_+$ , each agent  $i \in \mathcal{V}$  updates its vectors according to (18):

$$w_i(k+1) = \sum_{j \in \mathcal{N}_i^{\text{in}}(k) \cup \{i\}} \frac{v_j(k)}{d_j(k)+1}, \quad (18a)$$

$$y_i(k+1) = \sum_{j \in \mathcal{N}_i^{\text{in}}(k) \cup \{i\}} \frac{y_j(k)}{d_j(k)+1}, \quad (18b)$$

$$\lambda_i(k+1) = \frac{w_i(k+1)}{y_i(k+1)}, \quad (18c)$$

$$p_i(k+1) = \arg \min_{p_i \in \Omega_i} \sum_{t=1}^T C_i(p_{i,t}) - \lambda_i(k+1)' p_i, \quad (18d)$$

$$v_i(k+1) = w_i(k+1) - \alpha(k+1)(p_i(k+1) - D^i) \quad (18e)$$

where the division in (18c) operates entry-wise. The algorithm is initialized with an arbitrarily assigned vector  $v_i(0) \in \mathbb{R}^T$  and  $y_i(0) = \mathbf{1}$  for all  $i \in \mathcal{V}$ .

The step-size  $\alpha(k+1)$  satisfies the following decaying (diminishing) conditions:

$$\sum_{k=1}^{\infty} \alpha(k) = \infty, \quad \sum_{k=1}^{\infty} \alpha^2(k) < \infty, \quad (19)$$

$$\alpha(k) \leq \alpha(s) \text{ for all } k > s \geq 1.$$

A typical choice for a sequence  $\alpha(k)$  satisfying (19) is  $\alpha(k) = \frac{a}{k+b}$ , where  $a > 0$  and  $b \geq 0$ .

The proposed algorithm (18) is inspired by the push-sum and (sub)gradient method developed in [23]. Note that it follows from (17) that, the term  $-(p_i(k+1) - D^i)$  in the update (18e), where  $p_i(k+1)$  is given by (18d), is the gradient of the function  $\Phi_i(\lambda_i(k+1))$ . The update (18) without the gradient term is called the push-sum algorithm [30]–[32], or the ratio consensus algorithm [33]–[35] in the literature. In these algorithms, all  $\lambda_i(k)$  converge to the average of initial values as  $k \rightarrow \infty$ . The inclusion of the gradient term in the update is to ensure that all  $\lambda_i(k)$  converge to a dual optimal solution  $\lambda^*$ .

**Remark 4.** In algorithm (18), at each iteration  $k \in \mathbb{Z}_+$ , each agent (DER)  $i$  sends the quantities  $\frac{v_i(k)}{d_i(k)+1}$  and  $\frac{y_i(k)}{d_i(k)+1}$  to all the agents  $j$  in its out-neighbors set and receives the corresponding messages from its in-neighbors. In order to implement the algorithm, each agent  $i$  needs to know its out-degree  $d_i(k)$ , which is necessary if agents are not aware of their unique identifier (ID) as shown in [36]. However, this may be impractical when agents use a broadcast-based communication [36]–[38]. To overcome the case that some agents are not aware of their out-degrees, the authors of [37] developed a strategy where each agent also transmits to its out-neighbors one extra running sum along with its unique identifier. By assuming that each agent knows its unique identifier, the authors of [38] proposed a distributed optimization algorithm with row stochastic matrices, which is easier to construct since each agent could assign edge weights to its in-neighbors. Note that the analyses in [37], [38] are applicable to fixed graphs. It is an interesting future research direction to extend their results to time-varying graphs.

The following theorem establishes the convergence result of algorithm (18).

**Theorem 1.** Suppose that Assumptions 1, 2, and 3 hold. The distributed algorithm (18) with the step-size  $\alpha(k)$  satisfying conditions in (19) asymptotically solves the optimal DER coordination problem (11), i.e.,  $p_i(k) \rightarrow p_i^*$  and  $\lambda_i(k) \rightarrow \lambda^*$  as  $k \rightarrow \infty$  for all  $i \in \mathcal{V}$ .

**Proof :** Note that the dual problem **DP** (13) has the same form as the distributed optimization problem considered in [23]. The difference is that the dual problem (13) is a maximization problem for the sum of concave functions while the problem in [23] is a minimization problem for the sum of convex functions. Therefore, in what follows we verify the sufficient conditions given in [23, Theorem 1] are indeed satisfied.

The first condition that the graph is uniformly jointly strongly connected holds due to Assumption 3. The second condition also holds since each function  $\Phi_i(\lambda)$  in the maximization problem (13) is concave and the optimal set is nonempty, due to Assumption 1. Finally, the gradient of each function  $\Phi_i(\lambda)$  given by (17) is uniformly bounded since the constrained set is bounded and the virtual local demand is also bounded. Hence, it follows from [23, Theorem 1], whose proof is somewhat involved, that  $\lambda_i(k) \rightarrow \lambda^*$  as  $k \rightarrow \infty$  for all  $i \in \mathcal{V}$ .

To be self-contained, here we briefly provide the essential idea of the proof, which contains two steps. In the first step, it can be shown that  $\lambda_i(k)$  for  $i \in \mathcal{V}$  converges to the average process  $\bar{v}(k)$  defined as:

$$\bar{v}(k) = \frac{1}{N+M} \sum_{i=1}^{N+M} v_i(k), \quad (20)$$

increasingly well as iteration goes on, i.e.,  $\lambda_i(k) \rightarrow \bar{v}(k)$  as  $k \rightarrow \infty$ . In the second step, it can be shown that the average process converges to a dual optimal solution  $\lambda^*$ , i.e.,  $\bar{v}(k) \rightarrow \lambda^*$  as  $k \rightarrow \infty$ .

Finally, it follows from the update equation (18d) and the zero duality between the primal problem (11) and the dual problem (13) that  $p_i(k) \rightarrow p_i^*$  as  $k \rightarrow \infty$  for all  $i \in \mathcal{V}$ . ■

### C. Accelerated Distributed Algorithm

In the previous section, we have developed a distributed algorithm with the diminishing step-sizes for solving the optimal DER coordination problem. However, the convergence of algorithm (18) is rather slow due to the diminishing step-sizes. In order to speed up the convergence, motivated by the recent advances in distributed optimization [39]–[43], in this section, we develop an accelerated distributed algorithm with a fixed step-size.

More specifically, at each iteration  $k \in \mathbb{Z}_+$ , each DER (agent)  $i \in \mathcal{V}$  maintains five vectors  $u_i(k)$ ,  $v_i(k)$ ,  $\lambda_i(k)$ ,  $p_i(k)$ ,  $y_i(k) \in \mathbb{R}^T$ , where  $u_i(k)$ ,  $v_i(k)$ , and  $y_i(k)$  are auxiliary vectors, and  $p_i(k)$  and  $\lambda_i(k)$  are agent  $i$ 's estimations of the primal solution (optimal generations of DGs and ESSs) and the

dual optimal solution (optimal incremental cost), respectively. These vectors are updated as follows:

$$u_i(k+1) = \sum_{j \in \mathcal{N}_i^{\text{in}}(k) \cup \{i\}} \frac{u_j(k) + \alpha y_j(k)}{d_j(k) + 1}, \quad (21a)$$

$$v_i(k+1) = \sum_{j \in \mathcal{N}_i^{\text{in}}(k) \cup \{i\}} \frac{v_j(k)}{d_j(k) + 1}, \quad (21b)$$

$$\lambda_i(k+1) = \frac{u_i(k+1)}{v_i(k+1)}, \quad (21c)$$

$$p_i(k+1) = \arg \min_{p_i \in \Omega_i} \sum_{t=1}^T C_i(p_{i,t}) - \lambda_i(k+1)' p_i, \quad (21d)$$

$$y_i(k+1) = \sum_{j \in \mathcal{N}_i^{\text{in}}(k) \cup \{i\}} \frac{y_j(k)}{d_j(k) + 1} - (p_i(k+1) - p_i(k)). \quad (21e)$$

The algorithm is initialized with any  $\lambda_i(0) = u_i(0) \in \mathbb{R}^T$ ,  $v_i(0) = \mathbf{1}$ ,  $p_i(0) = \arg \min_{p_i \in \Omega_i} \sum_{t=1}^T C_i(p_{i,t}) - \lambda_i(0)' p_i$ , and  $y_i(0) = -(p_i(0) - D^i)$ .

The proposed algorithm (21) is inspired by the algorithm recently proposed in [41], which uses a distributed inexact gradient method and employs the push-sum protocol for gradient tracking (termed as Push-DIGing). Intuitively, the update (21a) is a distributed inexact gradient method where the variable  $y_j(k)$  is used instead of the average gradient, and in the update (21e),  $y_j(k)$  tracks the average gradient by employing dynamic average consensus [44]. The updates (21b) and (21c) are based on the push-sum method to handle the time-varying directed graphs. Note that the implementation of the algorithm also requires each agent to know its out-degree.

**Remark 5.** Note that it follows from (17) that the term  $-(p_i(k+1) - p_i(k))$  in (21e) is equal to  $-(\nabla \Phi_i(\lambda_i(k+1)) - \nabla \Phi_i(\lambda_i(k)))$ , which is the new information contained in the most recent gradient evaluation. This idea has also been used in other recent studies to accelerate the convergence speed, among which [39]–[42] focused on undirected graphs and [41] studied directed graphs. Here, we propose algorithm (21) based on Push-DIGing developed [41] since it is applicable to time-varying directed graphs.

To establish convergence of algorithm (21), we make the following assumption on the cost functions.

**Assumption 4.** For each  $t \in \mathcal{T}$ , the cost function  $C_i(p_{i,t})$  or  $C_i(p_{i,t}^+, p_{i,t}^-)$  is increasing and twice continuously differentiable over  $p_i \in \Omega_{\mathcal{N},i}$  for  $i \in \mathcal{N}$  and  $\{p_i^+, p_i^-\} \in \Omega_{\mathcal{M},i}$  for  $i \in \mathcal{M}$ , respectively. Moreover, there exist two positive constants  $\alpha_i$  and  $\beta_i$ , such that  $\frac{1}{\alpha_i} \leq \nabla^2 C_i(p_{i,t}) \leq \frac{1}{\beta_i}$  for  $p_i \in \Omega_i$ .

It is easy to see that when Assumption 4 is satisfied, each cost function  $C_i(\cdot)$  is strongly convex for  $p_i \in \Omega_i$  since  $\nabla^2 C_i(p_{i,t}) \geq \frac{1}{\alpha_i}$  for  $p_i \in \Omega_i$ . Therefore,  $C_i(\cdot)$  is strictly convex over  $p_i \in \Omega_i$ , that is, Assumption 1 is satisfied for  $p_i \in \Omega_i$ . However, when the cost functions satisfy additional properties as given in Assumption 4, compared with algorithm (18) which asymptotically solves the optimal DER coordination problem,

the new algorithm (21) exponentially achieves the optimal DER coordination as shown in the following theorem.

**Theorem 2.** Suppose that Assumptions 2, 3, and 4 hold. Then there exists a constant  $\bar{\alpha} > 0$ , such that the distributed algorithm (21) with any fixed step-size  $0 < \alpha < \bar{\alpha}$ , exponentially solves the optimal DER coordination problem (11), i.e.,  $\lambda_i(k) \rightarrow \lambda^*$  and  $p_i(k) \rightarrow p_i^*$  as  $k \rightarrow \infty$  for all  $i \in \mathcal{V}$  with a linear convergence rate.

**Proof :** We first show that the each estimate  $\lambda_i(k)$  for  $i \in \mathcal{V}$  converges to the dual optimal solution  $\lambda^*$  exponentially (R-linearly in the language of optimization theory). We show this by verifying that all the sufficient conditions of [41, Theorem 5.9] are satisfied.

- Firstly, we show that each local dual function  $\Phi_i(\lambda)$  given in (14) is strongly concave and smooth. Note that it follows from (15) that

$$p_i(\lambda) = \text{proj}_{\Omega_i}(\nabla \tilde{C}_i^{-1}(\lambda)), \quad (22)$$

where  $\tilde{C}_i(p_i) = \sum_{t=1}^T C_i(p_{i,t})$ , and  $\nabla \tilde{C}_i^{-1}(\lambda)$  is the inverse function of  $\nabla \tilde{C}_i(p_i)$ , which exists over its argument domain since each function  $\nabla C_i(p_{i,t})$  is continuous and  $C_i(p_{i,t})$  is strictly increasing in its argument domain due to Assumption 4, and  $\text{proj}_{\Omega_i}(\cdot)$  denotes the projection to the set  $\Omega_i$ . It then follows from the standard property of the derivative of the inverse function that

$$\frac{\partial p_i}{\partial \lambda}(\lambda) = \begin{cases} \frac{1}{\tilde{C}_i''(p_i(\lambda))}, & \text{if } p_i \in \Omega_i, \\ 0 & \text{otherwise.} \end{cases} \quad (23)$$

Next from (17), we have

$$\nabla^2 \Phi_i(\lambda) = -\frac{\partial p_i}{\partial \lambda}(\lambda).$$

This together with (23) implies that

$$\nabla^2 \Phi_i(\lambda) = \begin{cases} -\frac{1}{\tilde{C}_i''(p_i(\lambda))}, & \text{if } p_i \in \Omega_i, \\ 0, & \text{otherwise.} \end{cases}$$

Finally, it follows from the fact that  $\frac{1}{\alpha_i} \leq \nabla^2 C_i(p_{i,t}) \leq \frac{1}{\beta_i}$  from Assumption 4 that when exists,

$$-\frac{\alpha_i}{T} \leq \nabla^2 \Phi_i(\lambda) \leq -\frac{\beta_i}{T}. \quad (24)$$

The second inequality of (24) implies that the local dual function  $\Phi_i(\lambda)$  given in (14) for any  $i \in \mathcal{V}$  is strongly concave. This also ensures that the dual problem (13) has a unique optimal solution  $\lambda^*$ .

Also note that it follows from the first inequality of (24) and [27, Lemma 3] that the local dual function  $\Phi_i(\lambda)$  has a Lipschitz continuous gradient. Thus the local dual function  $\Phi_i(\lambda)$  is smooth.

- Next, the time-varying directed graph is uniformly jointly strongly connected by Assumption 3.
- Finally, it is easy to see that the weight mixing matrix of algorithm (21) given by

$$A_{ij}(k) = \begin{cases} \frac{1}{d_j(k)+1}, & \text{if } j \in \mathcal{N}_i^{\text{in}}(k) \cup \{i\}, \\ 0, & \text{otherwise.} \end{cases}$$

is column stochastic for all  $k \in \mathbb{Z}_+$ .

Hence, it follows from [41, Theorem 5.9] that if  $\alpha < \bar{\alpha}$ , where  $\bar{\alpha}$  is a constant depending on the parameters of the communication network and cost functions, then each estimate  $\lambda_i(k)$  for  $i \in \mathcal{V}$  converges to the unique optimal solution  $\lambda^*$  exponentially (R-linearly in the language of optimization theory). That is, there exist some constants  $C > 0$  and  $0 < \rho < 1$  such that

$$\|\lambda_i(k) - \lambda^*\| \leq C\rho^k, \forall k \in \mathbb{Z}_+. \quad (25)$$

Next, we show that the estimate  $p_i(k)$  for each  $i \in \mathcal{V}$  converges to the optimal generation  $p_i^*$  exponentially.

Note that from (16) and (22), for all  $k \in \mathbb{Z}_+$ , we have

$$\begin{aligned} & \|p_i(k) - p_i^*\| \\ &= \left\| \text{proj}_{\Omega_i}(\nabla \tilde{C}_i^{-1}(\lambda_i(k))) - \text{proj}_{\Omega_i}(\nabla \tilde{C}_i^{-1}(\lambda^*)) \right\| \\ &\leq \left\| \nabla \tilde{C}_i^{-1}(\lambda_i(k+1)) - \nabla \tilde{C}_i^{-1}(\lambda^*) \right\| \\ &\leq \left\| \nabla^2 \tilde{C}_i^{-1}((1-c)\lambda_i(k+1) + c\lambda^*) \right\| \\ &\quad \times \|\lambda_i(k+1) - \lambda^*\| \\ &\leq \frac{\alpha_i C}{T} \rho^k, \end{aligned} \quad (26)$$

for some constant  $0 < c < 1$ , where the first inequality follows from the standard non-expansiveness property of the projection operator [25], the second inequality follows from the mean value theorem, and the last inequality follows from  $\nabla^2 C_i(p_{i,t}) \geq \frac{1}{\alpha_i}$  and (25). Therefore, (26) implies that  $p_i(k)$  converges to the optimal generation  $p_i^*$  with the same R-linear rate  $O(\rho^k)$  as the convergence of  $\lambda_i(k)$  to  $\lambda^*$ . Hence, algorithm (21) achieves the optimal DER coordination exponentially (R-linearly). ■

**Remark 6.** Note that the complicated closed form expressions for  $\bar{\alpha}$  and the convergence rate  $\rho$  can be derived by following the worst-case analysis given in [41, Theorem 5.9]. However, such an upper bound on fixed step-sizes is often conservative and larger values can be chosen in the numerical experiments. Nevertheless, it ensures the distributed algorithm to have provable convergence guarantees.

**Remark 7.** Compared to algorithm (18) with diminishing step-sizes, algorithm (21) uses a fixed step-size and solves to the optimal DER coordination problem faster. In order to establish faster convergence of (21), the local cost functions  $C_i(p_{i,t})$  for  $i \in \mathcal{V}$  and  $t \in \mathcal{T}$  need to satisfy more properties as given in Assumption 4. Therefore, depending on the properties of cost functions, we may choose different algorithms to solve the optimal DER coordination problem. If only Assumption 1 is satisfied, we use algorithm (18) to solve the optimal DER coordination problem. If Assumption 4 is satisfied, we can use algorithm (21) to solve the optimal DER coordination problem with a faster convergence rate.

## V. CASE STUDIES

In this section, we validate and evaluate the performance of algorithm (18) with diminishing step-sizes and algorithm (21) with a fixed step-size for optimal DER coordination using the IEEE 39-bus system modified from [45]–[47] shown in Fig. 1,

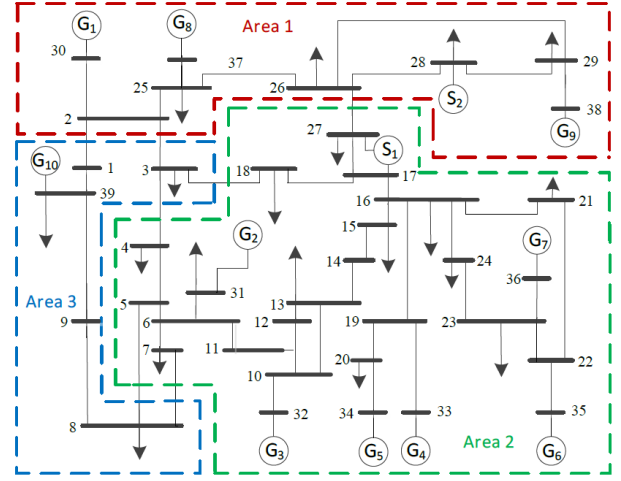


Fig. 1. IEEE 39-bus system.

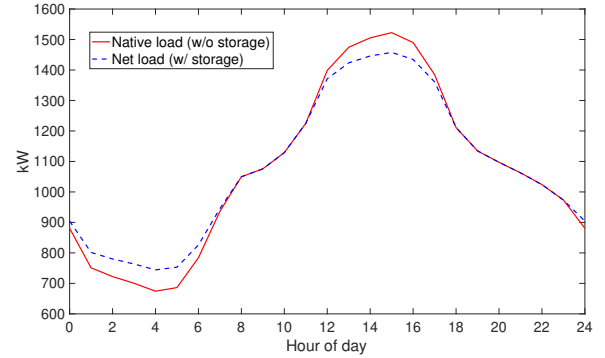


Fig. 2. Native load vs. Net load.

TABLE I  
GENERATOR PARAMETERS

Unit	$a_i$ (\$/kW <sup>2</sup> h)	$b_i$ (\$/kWh)	$c_i$ (\$/h)	Range (kW)
1	0.00024	0.0267	0.38	[30,60]
2	0.00052	0.0152	0.65	[20,60]
3	0.00042	0.0185	0.4	[50,200]
4	0.00031	0.0297	0.3	[20,140]
5	0.000248	0.0156	0.3312	[50,300]
6	0.000199	0.0116	0.4969	[110,500]
7	0.00028	0.0195	0.32	[40,250]
8	0.00028	0.0195	0.32	[40,250]
9	0.00042	0.0185	0.4	[50,200]
10	0.00031	0.0297	0.3	[20,140]

where Buses 30–39 are connected with DGs, and Buses 27 and 28 are connected to ESs. The demand during a 24-hour is plotted in red in Fig. 2.

The cost functions of DGs are given by  $C_i(p_{i,t}) = a_i^2 p_{i,t} + b_i p_{i,t} + c_i$  for all  $i \in \{30, 31, \dots, 39\}$  and for  $t \in \{1, 2, \dots, 24\}$ . The cost functions of ESs are given by  $C_i(p_{i,t}) = a_i^2 p_{i,t}$  for  $i \in \{27, 28\}$  and for  $t \in \{1, 2, \dots, 24\}$ . The parameters of DGs and ESs are given in Table I and Table II, respectively. Note that the cost functions satisfy both Assumptions 1 and 4.

The IEEE 39-bus system has been partitioned into three

TABLE II  
STORAGE PARAMETERS

Unit	$a_i$ (\$/kW <sup>2</sup> h)	$E_s$ (kWh)	$p_{\max}$ (kW)	$\eta_+$	$\eta_-$
1	0.0001	500	50	0.8	0.8
2	0.0001	400	40	0.88	0.88

areas as shown in Fig. 1. Within each area, the communication links among DERs are bidirectional. The communication links between different areas maybe unidirectional. More specifically, the communication links between Area 1 and Area 3 are (2, 1) and (2, 3), and the communication link between Area 1 and Area 2 is (27, 26). The communication network is modeled as a time-varying directed network switching among two fixed topologies. More specifically,

$$\mathcal{G}(k) = \begin{cases} \mathcal{G}_1, & \text{if } k \in [0, 1) \cup [2, 3) \cdots \cup [2s, 2s+1) \cdots, \\ \mathcal{G}_2, & \text{if } k \in [1, 2) \cup [3, 4) \cdots \cup [2s+1, 2s+2) \cdots, \end{cases}$$

where  $s \in \mathbb{Z}_+$ ,  $\mathcal{G}_1$  is the directed graph obtained by disconnecting Area 1 and Area 3, that is, by removing the links (2, 1) and (2, 3), and  $\mathcal{G}_2$  is the directed graph obtained by disconnecting Area 1 and Area 2, that is, by removing the link (27, 26). It is easy to check that each of the fixed topologies  $\mathcal{G}_1$  and  $\mathcal{G}_2$  is not connected. For example, in the directed graph  $\mathcal{G}_1$ , there is no directed path from agents in Area 1 to agents in Area 2 and Area 3. However, the time-varying directed graph  $\mathcal{G}(k)$  is uniformly jointly strongly connected since the joint graph  $\mathcal{G}([k_0, k_0 + B))$  is strongly connected for any  $k_0 \geq 0$  with  $B = 2$ . Thus, Assumption 3 is satisfied with  $B = 2$ . According to Theorems 1 and 2, algorithm (18) with properly chosen diminishing step-sizes and algorithm (21) with a fixed step-size less than a certain critical value, solve the optimal DER coordination problem, which will be verified in the following subsections, respectively.

#### A. The Algorithm with Diminishing Step-Sizes

We start by evaluating algorithm (18). We have verified that algorithm (18) with the diminishing step-size  $\alpha(k) = \frac{0.005}{k}$  which satisfies the conditions in (19) converges to the optimal generation roughly in 10000 steps and the computation time is 61.7079 seconds. The blue curve in Fig. 2 is the resulting net load (load minus storage). Fig. 2 shows how two ESs are coordinated to cut the peak and fill the valley. In particular, they are discharged during peak hours when the energy price is high and charged during off-peak hours when the energy price is low.

The obtained optimal generations for 10 DGs and 2 ESs over 24 hours are plotted in Fig. 3. The power output and state of charge (SOC) for both storages are provided in Fig. 4. For each storage, SOC is the same at the beginning and end of the scheduling period, but the total charging energy (area between the negative blue curve and x-axis) is more than the discharging energy (area between the positive blue curve and x-axis) due to charging and discharging losses. ESs are idle when the energy price is not high (or low) enough to make the discharging (or charging) profitable considering the round-trip efficiency. In particular, ES 1 is idle during

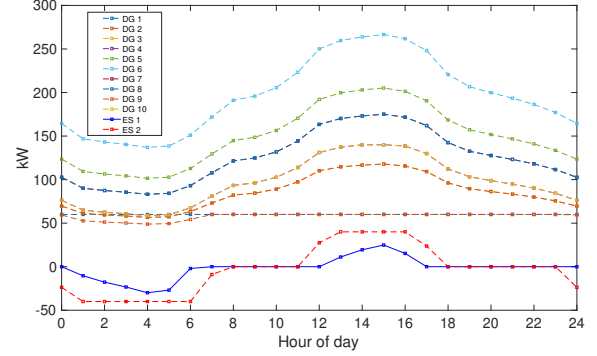


Fig. 3. Optimal power generations for DGs and ESs over 24 hours.

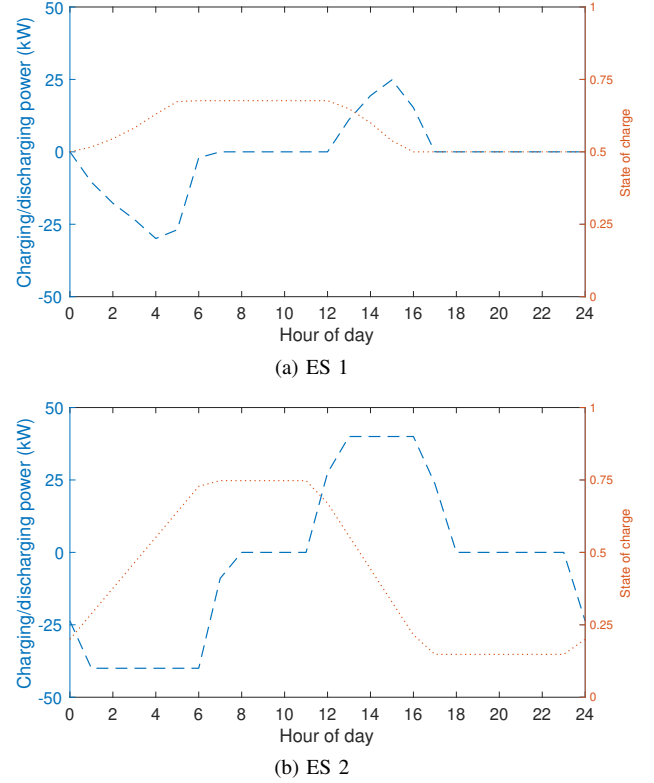


Fig. 4. Charging (negative) and discharging (positive) power and state of charge.

Hour 6-12 and 17-24 because of its low charging/discharging efficiency, while ES 2 is engaged more often due to its higher charging/discharging efficiency.

In order to show the convergence more clearly, the simulation results for Hour 1 are given in Fig. 5. Fig. 5(a) plots the evolutions of the estimates  $\lambda_{i,1}(k)$  all for  $i \in \mathcal{V}$ , which shows that all  $\lambda_{i,1}(k)$  converges to 0.0716 \$/kWh, which is roughly equal to the optimal incremental cost  $\lambda_1^* = 0.0699$  \$/kWh. In order to clearly show the convergence process for power generations, in Fig. 5(b), we have only included the evolutions for the power generation for DGs 1, 5, 6, 8, and 10, and ESs 1 and 2. They all converge to the corresponding optimal generations. Fig. 5(c) plots the evolution of the total generation in comparison with the total demand, which clearly shows that



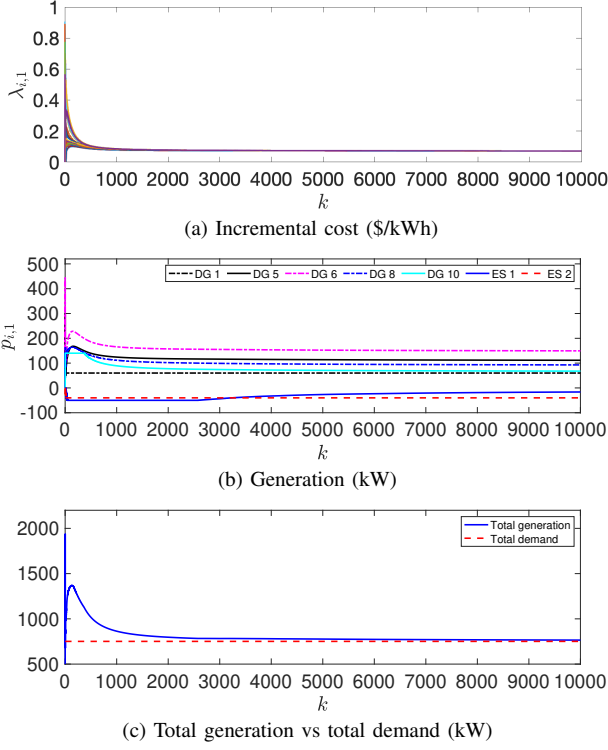


Fig. 5. Results of algorithm (18) with the diminishing step-size  $\alpha(k) = \frac{0.005}{k}$ .

the total generation meets the total demand for Hour 1, i.e.,  $D_1 = 750.9792$  kW.

### B. The Algorithm with a Fixed Step-Size

To accelerate the convergence, we next apply algorithm (21). We have checked that algorithm (21) with the fixed step-size  $\alpha = 5 \times 10^{-5}$  solves the optimal DER coordination problem roughly in 1000 steps and the computation time is 6.1681 seconds. The simulation results for Hour 1 are given in Fig. 6. Fig. 6(a) shows that all estimates  $\lambda_{i,1}(k)$  converge to the optimal incremental cost  $\lambda_1^* = 0.0699$  \$/kWh. The evolutions of power generations for DGs 1, 5, 6, 8, and 10, and ESs 1 and 2 are plotted in Fig. 6(b), which shows that they converge to the corresponding optimal generations. Fig. 6(c) plots the evolution of the total generation in comparison with the total demand, which clearly shows that the total generation meets the total demand for Hour 1, i.e.,  $D_1 = 750.9792$  kW.

Compared with the simulation results of algorithm (18) with the diminishing step-sizes shown in Fig. 5, the convergence of algorithm (21) with the fixed step-size is much faster. For example, at time step  $k = 1000$ , the estimates  $\lambda_{i,1}(k)$  are all roughly equal to optimal incremental cost  $\lambda_1^* = 0.0699$  \$/kWh, while using algorithm (18) with the diminishing step-sizes  $\alpha(k) = \frac{0.005}{k}$ , at the same time step, all estimates  $\lambda_{i,1}(k)$  are in the range of  $[0.0746, 0.0830]$ , which are in the neighborhood of the optimal incremental cost with 18.7% error.

To make the comparison more explicit, we have also plotted the total generations of these two algorithms together with the demand for Hour 1 in Fig. 7. As can be seen, the total generation by running algorithm (21) with  $\alpha = 5 \times 10^{-5}$  converges to

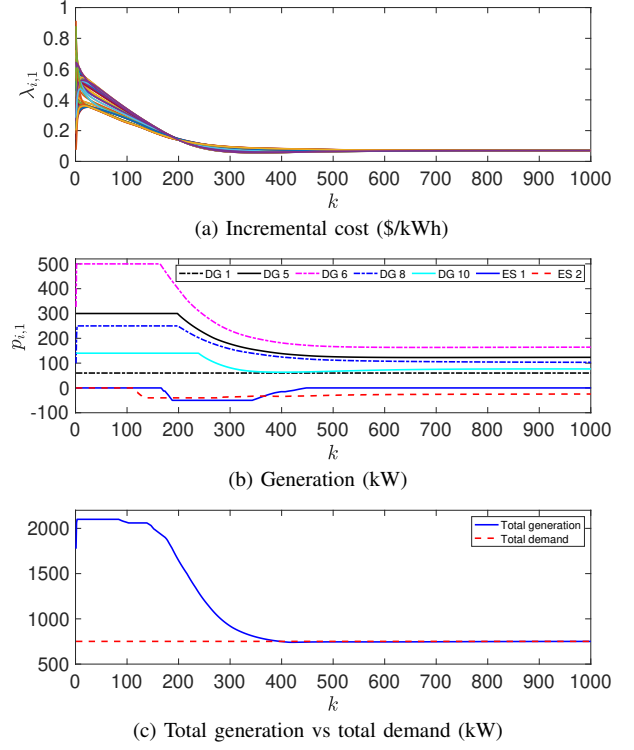


Fig. 6. Results of algorithm (21) with the fixed step-size  $\alpha = 5 \times 10^{-5}$ .

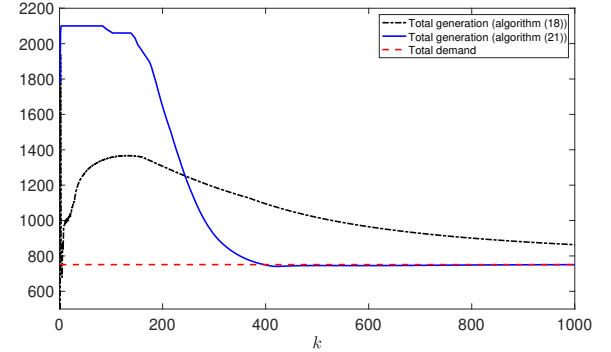


Fig. 7. Comparison of algorithm (18) and algorithm (21).

the total demand roughly at the time step  $k = 400$ , while the convergence of algorithm (18) with the diminishing step-sizes  $\alpha(k) = \frac{0.005}{k}$  is not achieved even at the time step  $k = 1000$ . Hence, the convergence of algorithm (21) with a fixed step-size is faster.

## VI. CONCLUSIONS

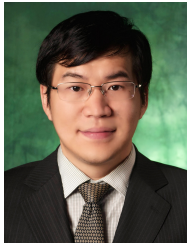
In this paper, we have considered the optimal coordination problem of DERs, including DGs and ESs. In the problem formulation, storage charging/discharging efficiencies were explicitly modeled. We first proposed a distributed algorithm with diminishing step-sizes and showed that the proposed algorithm with an appropriately chosen step-size asymptotically solves the optimal DER coordination problem over time-varying directed communication networks that are uniformly jointly strongly connected. To accelerate the convergence, we

next developed a distributed algorithm with a fixed step-size and showed that the new proposed algorithm exponentially solves the optimal DER coordination problem if the cost function satisfy additional properties. The performances of the proposed algorithms have been tested on the IEEE 39-bus system. One future research direction is to extend the proposed distributed algorithms to accommodate other communication effects, such as time delays and packet drops. Another interesting research direction is to extend the proposed distributed algorithms to accommodate additional physical constraints, such as those relevant to transmission line loss, power flow and transmission line flow.

## REFERENCES

- [1] T. Yang, D. Wu, W. Ren, H. Wang, Y. Hong, and K. H. Johansson, "Cooperative optimal coordination for distributed energy resources," in *Proc. IEEE Conf. Decision Control*, Dec. 2017, pp. 6334–6339.
- [2] R. Lasseter, A. Akhil, C. Marnay, J. Stephens, J. Dagle, R. Guttromson, A. S. Meliopoulos, R. Yinger, and J. Eto, "Integration of distributed energy resources: The CERTS Microgrid concept," *Lawrence Berkeley National Laboratory*, 2002.
- [3] Z. Zhang and M.-Y. Chow, "Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1761–1768, Nov. 2012.
- [4] S. Kar and G. Hug, "Distributed robust economic dispatch in power systems: A consensus + innovations approach," in *Proc. IEEE Power Energy Soc. Gene. Meet.*, Jul. 2012.
- [5] A. D. Domínguez-García, S. T. Cady, and C. N. Hadjicostis, "Decentralized optimal dispatch of distributed energy resources," in *Proc. IEEE Conf. Decision Control*, Dec. 2012, pp. 3688–3693.
- [6] S. Yang, S. Tan, and J.-X. Xu, "Consensus based approach for economic dispatch problem in a smart grid," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4416–4426, Nov. 2013.
- [7] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Trans. Autom. Control*, vol. 59, no. 5, pp. 1177–1189, May 2014.
- [8] H. Xing, Y. Mou, M. Fu, and Z. Lin, "Distributed bisection method for economic power dispatch in smart grid," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3024–3035, Nov. 2015.
- [9] T. Yang, D. Wu, Y. Sun, and J. Lian, "Minimum-time consensus based approach for power system applications," *IEEE Trans. Ind. Electron.*, vol. 63, no. 2, pp. 1318–1328, Feb. 2016.
- [10] T. Yang, J. Lu, D. Wu, J. Wu, G. Shi, Z. Meng, and K. H. Johansson, "A distributed algorithm for economic dispatch over time-varying directed networks with delays," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 5095–5106, Jun. 2017.
- [11] J. Wu, T. Yang, D. Wu, K. Kalsi, and K. H. Johansson, "Distributed optimal dispatch of distributed energy resources over lossy communication networks," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 3125–3137, Nov. 2017.
- [12] A. Cherukuri and J. Cortés, "Distributed generator coordination for initialization and anytime optimization in economic dispatch," *IEEE Trans. Control Netw. Syst.*, vol. 2, no. 3, pp. 226–237, Sep. 2015.
- [13] —, "Initialization-free distributed coordination for economic dispatch under varying loads and generator commitment," *Automatica*, vol. 74, pp. 183–193, 2016.
- [14] P. Yi, Y. Hong, and F. Liu, "Initialization-free distributed algorithms for optimal resource allocation with feasibility constraints and application to economic dispatch of power systems," *Automatica*, vol. 74, pp. 259–269, 2016.
- [15] Y. Xu, W. Zhang, G. Hug, S. Kar, and Z. Li, "Cooperative control of distributed energy storage systems in a microgrid," *IEEE Trans. Smart Grid*, vol. 1, no. 6, pp. 238–248, Jan. 2015.
- [16] H. Fang, D. Wu, and T. Yang, "Cooperative management of a Lithium-Ion battery energy storage network: a distributed MPC approach," in *Proc. IEEE Conf. Decision Control*, Dec. 2016, pp. 4226–4232.
- [17] G. Hug, S. Kar, and C. Wu, "Consensus + innovations approach for distributed multiagent coordination in a microgrid," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1893–1903, Jul. 2015.
- [18] T. Yang, D. Wu, A. A. Stoorvogel, and J. Stoustrup, "Distributed coordination of energy storage with distributed generators," in *Proc. IEEE Power Energy Soc. Gene. Meet.*, Jul. 2016.
- [19] D. Wu, T. Yang, A. A. Stoorvogel, and J. Stoustrup, "Distributed optimal coordination for distributed energy resources in power systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 2, pp. 414–424, Apr. 2017.
- [20] Z. Tang, D. J. Hill, and T. Liu, "A novel consensus-based economic dispatch for microgrids," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3920–3922, Jul. 2018.
- [21] A. Cherukuri and J. Cortés, "Distributed coordination of DERs with storage for dynamic economic dispatch," *IEEE Trans. Autom. Control*, vol. 63, no. 3, pp. 835–842, Mar. 2018.
- [22] D. Wu, C. Jin, P. Balducci, and M. Kintner-Meyer, "An energy storage assessment: Using optimal control strategies to capture multiple services," in *Proc. IEEE Power Energy Soc. Gene. Meet.*, Jul. 2015.
- [23] A. Nedić and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Trans. Autom. Control*, vol. 60, no. 3, pp. 601–615, Mar. 2015.
- [24] C. Godsi and G. F. Royle, *Algebraic Graph Theory*, ser. Graduate Texts in Mathematics. New York: Springer-Verlag, 2001, vol. 207.
- [25] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1999.
- [26] D. P. Bertsekas, A. Nedić, and A. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA: Athena Scientific, 2003.
- [27] S. H. Low and D. E. Lapsley, "Optimization flow control—I: Basic algorithm and convergence," *IEEE/ACM Trans. Networking*, vol. 7, no. 6, pp. 861–874, Dec. 1999.
- [28] C. Zhao, U. Topcu, and S. Low, "Optimal load control via frequency measurement and neighborhood area communication," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3576–3587, Nov. 2013.
- [29] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, 1989.
- [30] D. Kempe, A. Dobra, and J. Gehrke, "Gossip-based computation of aggregate information," in *44th Annual IEEE Symposium on Foundations of Computer Science*, Oct. 2003, pp. 482–491.
- [31] F. Bénézit, V. D. Blondel, P. Thiran, J. N. Tsitsiklis, and M. Vetterli, "Weighted gossip: Distributed averaging using non-doubly stochastic matrices," in *Proc. IEEE Info. Theory*, Jun. 2010, pp. 1753–1757.
- [32] K. I. Tsianos, S. Lawlor, and M. G. Rabbat, "Push-sum distributed dual averaging for convex optimization," in *Proc. IEEE Conf. Decision Control*, Dec. 2012, pp. 5453–5458.
- [33] A. D. Domínguez-García and C. N. Hadjicostis, "Distributed strategies for average consensus in directed graphs," in *Proc. IEEE Decision Control European Control Conf.*, Dec. 2011, pp. 2124–2129.
- [34] C. N. Hadjicostis and T. Charalambous, "Average consensus in the presence of delays in directed graph topologies," *IEEE Trans. Autom. Control*, vol. 59, no. 3, pp. 763–768, Mar. 2014.
- [35] T. Charalambous, Y. Yuan, T. Yang, W. Pan, C. N. Hadjicostis, and M. Johansson, "Distributed finite-time average consensus in digraphs in the presence of time-delays," *IEEE Trans. Control Netw. Syst.*, vol. 2, no. 4, pp. 370–381, Dec. 2015.
- [36] J. M. Hendrickx and J. N. Tsitsiklis, "Fundamental limitations for anonymous distributed systems with broadcast communications," in *53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2015, pp. 9–16.
- [37] C. N. Hadjicostis, N. Vaidya, and A. Dominguez-Garcia, "Robust distributed average consensus via exchange of running sums," *IEEE Trans. Autom. Control*, vol. 61, no. 6, pp. 1492–1507, Jun. 2016.
- [38] C. Xi, V. S. Mai, R. Xin, E. Abed, and U. A. Khan, "Linear convergence in optimization over directed graphs with row-stochastic matrices," *IEEE Trans. Autom. Control*, vol. 63, no. 10, pp. 3558–3565, Oct. 2018.
- [39] W. Shi, Q. Ling, G. Wu, and W. Yin, "EXTRA: An exact first-order algorithm for decentralized consensus optimization," *SIAM Journal on Optimization*, vol. 25, no. 2, pp. 944–966, 2015.
- [40] G. Qu and N. Li, "Harnessing smoothness to accelerate distributed optimization," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 3, pp. 1245–1260, Sep. 2018.
- [41] A. Nedić, A. Olshevsky, and W. Shi, "Achieving geometric convergence for distributed optimization over time-varying graphs," *SIAM Journal on Optimization*, vol. 27, no. 4, pp. 2597–2633, 2017.
- [42] J. Xu, S. Zhu, Y. C. Soh, and L. Xie, "Convergence of asynchronous distributed gradient methods over stochastic networks," *IEEE Trans. Autom. Control*, vol. 63, no. 2, pp. 434–448, Feb. 2018.
- [43] W. Du, L. Yao, D. Wu, X. Li, G. Liu, and T. Yang, "Accelerated distributed energy management for microgrids," in *Proc. IEEE Power Energy Soc. Gene. Meet.*, Aug. 2018.

- [44] M. Zhu and S. Martínez, "Discrete-time dynamic average consensus," *Automatica*, vol. 46, no. 2, pp. 322–329, 2010.
- [45] S. B. Yusof, G. J. Rogers, and R. T. H. Alden, "Slow coherency based network partitioning including load buses," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 1375–1382, Aug. 1993.
- [46] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.
- [47] E. Mallada, C. Zhao, and S. Low R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Optimal Load-Side Control for Frequency Regulation in Smart Grids," *IEEE Trans. Autom. Control*, vol. 62, no. 11, pp. 6294–6309, Dec. 2017.

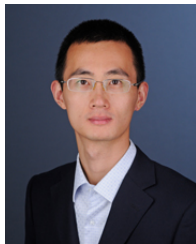


**Tao Yang** (M'12) is an Assistant Professor at the Department of Electrical Engineering, University of North Texas (UNT). He received the Ph.D. degree in electrical engineering from Washington State University in 2012. Between August 2012 and August 2014, he was an ACCESS Post-Doctoral Researcher with the ACCESS Linnaeus Centre, Royal Institute of Technology, Sweden. Prior to joining the UNT in 2016, he was a Scientist/Engineer with Energy & Environmental Directorate, Pacific Northwest National Laboratory. His research interests include distributed control and optimization with applications to power systems, cyber physical systems, networked control systems, and multi-agent systems. He received Ralph E. Powe Junior Faculty Enhancement Award (2019) from Oak Ridge Associated Universities and Best Student Paper award (as an advisor) at the 14th IEEE International Conference on Control & Automation.



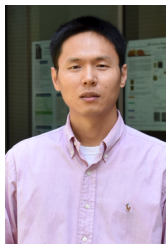
**Di Wu** (M'12) received the B.S. and M.S. degrees in electrical engineering from Shanghai Jiao Tong University in 2003 and 2006, respectively, and the Ph.D. degree in electrical and computer engineering from Iowa State University, Ames, in 2012. Dr. Wu is currently a staff engineer in Electricity Infrastructure and Buildings Division at Pacific Northwest National Laboratory. His research focuses on building-to-grid integration, energy storage for grid applications, and control and optimization of distributed energy resources. He serves as an Associate Editor for the

*IEEE Power and Energy Technology Systems Journal*.



**Huazhen Fang** (M'14) is an Assistant Professor of Mechanical Engineering at the University of Kansas. He received his Ph.D. degree in Mechanical Engineering from the Department of Mechanical & Aerospace Engineering at the University of California, San Diego (2014). Prior to that, he received M.Sc. in Mechanical Engineering from the University of Saskatchewan, Canada (2009), and B.Eng. in Computer Science & Technology from Northwestern Polytechnic University, China (2006). His research interests lie in control and estimation theory with ap-

plication to energy management, cooperative robotics and system prognostics. He has received the 2019 NSF CAREER Award and the awards of Outstanding Reviewer or Reviewers of the Year from *Automatica*, *IEEE Transactions on Cybernetics*, and *ASME Journal of Dynamic Systems, Measurement and Control*.



**Wei Ren** (M'04–SM'13–F'16) is currently a Professor with the Department of Electrical and Computer Engineering, University of California, Riverside. He received the Ph.D. degree in Electrical Engineering from Brigham Young University, Provo, UT, in 2004. Prior to joining UC Riverside, he was a faculty member at Utah State University and a postdoctoral research associate at the University of Maryland, College Park. His research focuses on distributed control of multi-agent systems and autonomous control of unmanned vehicles. Dr. Ren was a recipient

of the IEEE Control Systems Society Antonio Ruberti Young Researcher Prize in 2017 and the National Science Foundation CAREER Award in 2008. He is currently an Associate Editor for *Automatica* and *Systems and Control Letters*.



**Hong Wang** (M'97–SM'02) received the B.S. degree from the Huainan University of Mining Engineering, Huainan, China, in 1982, and the M.S. and Ph.D. degrees from the Huazhong University of Science and Technology, Wuhan, China, in 1984 and 1987, respectively. In 1992, he joined the University of Manchester Institute of Science and Technology, Manchester, U.K., where he was a Professor of process control between 2002 and 2016. He was a Research Fellow with Salford University, Salford, U.K., Brunel University, Uxbridge, U.K., and Southampton University, Southampton, U.K. Between 2016 - 2019, he was a Laboratory Fellow and Chief Scientist with Pacific Northwest National Laboratory, Richland, WA, USA, and is now with Oak Ridge National Laboratory, (USA) since Jan 2019 as a corporate fellow.

He has authored 300 papers and six books in these research areas. His current research interests include stochastic distribution control, fault detection and diagnosis, nonlinear control, and data-based modeling for complex systems. Dr. Wang was an Associate Editor of the IEEE Transactions on Automatic Control. He has served as an Associate Editor of the IEEE Transactions on Control Systems Technology and the IEEE Transactions on Automation Science and Engineering.



**Yigong Hong** (M'99–SM'02–F'17) received his B.S. and M.S. degrees from Dept of Mechanics of Peking University, China, and the Ph.D. degree from the Chinese Academy of Sciences (CAS), China. He is currently a professor in Academy of Mathematics and Systems Science, CAS, and serves as the Director of Key Lab of Systems and Control, CAS and the Director of the Information Technology Division, National Center for Mathematics and Interdisciplinary Sciences, CAS. Also, he is a Fellow of IEEE, a Fellow of Chinese Association for Artificial

Intelligence, and a Fellow of Chinese Association of Automation (CAA). Additionally, he is the chair of Technical Committee of Control Theory of CAA and was a board of governor of IEEE Control Systems Society. His current research interests include nonlinear control, multi-agent systems, distributed optimization and game, machine learning, and social networks. He serves as Editor-in-Chief of Control Theory and Technology. He also serves or served as Associate Editors for many journals including the IEEE Transactions on Automatic Control, IEEE Transactions on Control of Network Systems, and IEEE Control Systems Magazine. Moreover, he is a recipient of the Guan Zhaozhi Award at the Chinese Control Conference, Young Author Prize of the IFAC World Congress, Young Scientist Award of CAS, the Youth Award for Science and Technology of China, and the National Natural Science Prize of China.



**Karl Henrik Johansson** (SM'08–F'13) is Professor at the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology. He received MSc and PhD degrees from Lund University. He has held visiting positions at UC Berkeley, Caltech, NTU, HKUST Institute of Advanced Studies, and NTNU. His research interests are in networked control systems, cyber-physical systems, and applications in transportation, energy, and automation. He has served on the IEEE Control Systems Society Board of Governors, the IFAC Executive Board, and the European Control Association Council. He has received several best paper awards and other distinctions from IEEE and ACM. He has been awarded Distinguished Professor with the Swedish Research Council and Wallenberg Scholar with the Knut and Alice Wallenberg Foundation. He has received the Future Research Leader Award from the Swedish Foundation for Strategic Research and the triennial Young Author Prize from IFAC. He is Fellow of the IEEE and the Royal Swedish Academy of Engineering Sciences, and he is IEEE Distinguished Lecturer.