Distributed Energy Management in Smart Grid With Wind Power and Temporally Coupled Constraints

Wenchao Meng and Xiaoyu Wang, *Senior Member, IEEE*

***Abstract*—This paper investigates the problem of distributed energy management for both generation and demand sides in a smart grid by formulating the economic dispatch (ED) and demand response (DR) in a united framework. Our main contribution is to formulate a social welfare maximization problem for a more practical scenario by taking wind power, and temporally coupled constraints of the demands into account. The complexity lies in the non-quadratic cost function of wind power, the temporally coupled constraints of the demands, and the non-convexity of the optimization problem. Meanwhile, a smart grid has to guarantee privacy and accommodates plug-and-play features. Aiming at these challenges, we first relax an equality constraint to obtain a new convex optimization problem. Because of the coupling in the constraint, the Lagrange duality method is then adopted to decompose the problem into subproblems for generators and demands, which are regarded as agents. As a result, each agent solves its subproblem by exchanging information with only neighbor agents, and coordinates with others using the global information discovered by a distributed finite-time consensus algorithm. We also prove the convergence and optimality of the proposed distributed energy management algorithm (DEMA). Finally, simulations are performed on the IEEE 39-bus system to illustrate the performance of our DEMA.**

***Index Terms*—Distributed energy management, wind power, temporally coupled constraint, smart grid.**

NOMENCLATURE

κ	Shape factor of the Weibull distribution of wind.
$\boldsymbol{\lambda}, \boldsymbol{\mu}$	Lagrangian multiplier vectors.
$\mathcal{E}$	Set of edges between agents.
N_i	Neighbors of agent i.
$\mathcal{T}$	Set of time slots.
$\mathcal{V}$	Set of agents.
π_{it}^d	Profit of demand i at time slot t.
π_{it}^g	Profit of thermal generator (TG) i at time slot t.
π_{it}^w	Profit of wind turbine (WT) i at time slot t.
δ_i	Cost coefficient of the linear cost of wind turbine i.
A	Adjacency matrix of agents.
c	Scale factor of the Weibull distribution of wind.
C_{it}	Cost function of generator i at time slot t.
C_{pwi}	Cost coefficient of the underestimation of the availability of wind turbine i.
C_{rwi}	Cost coefficient of the overestimation of the availability of wind turbine i.
k	Index of iterations.
k'	Internal iteration index for the information discovery process.
P_{it}^L	Transmission losses induced by generator i at time slot t.
p_t	Electricity price at time slot t.
P_{it}	Scheduled power of TG i when $i \in V_g$ or demand i when $i \in V_d$ at time slot t.
r_i	Required power for demand i to complete a given task in the predefined time horizon.
S_{it}	Subproblems to be solved for agent i at time slot t.
U_{it}	Utility function of demand i at time slot t.
V_d	Set of demands.
V_g	Set of TGs.
V_w	Set of WTs.
v_{in}	Cut-in wind speed.
v_{out}	Cut-out wind speed.
v_r	Rated wind speed.
W_r	Rated wind power.
W_{it}	Scheduled wind power of WT i at time slot t.
W_{it}^{av}	Available generated power of WT i at time slot t.

Manuscript received June 14, 2016; revised September 22, 2016, November 24, 2016 and February 3, 2017; accepted February 17, 2017. This paper is supported by Visiting Scholarship of State Key Laboratory of Power Transmission Equipment & System Security and New Technology (Chongqing University)(2007DA10512716405).

W. Meng is with Department of Electronics, Carleton University, Ottawa, ON, Canada (e-mail:wmengzju@gmail.com).

X. Wang is with Department of Electronics, Carleton University, Ottawa, ON, Canada and State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University, Chongqing, China (email: xiaoyuw@ieee.org)

I. INTRODUCTION

THE smart grid, regarded as the next generation power grid, has better efficiency and reliability with possible integration of renewable energy sources. These new capabilities are made possible by integrating various advanced technologies such as information and communications technologies (ICT) and computer processing into the power grid [1]–[5]. In light of the smart grid, the future power system will be equipped with intelligent controllable electrical devices capable of exchanging information through the communication network. One of the fundamental optimization challenges is the energy management that plays a key role in optimization and scheduling of controllable devices in the smart grid.

A common practice in addressing the energy management problem is to utilize a centralized solution such as quadratic programming [6], Lagrangian relaxation technique [7] and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2017.2682001, IEEE Transactions on Industrial Electronics

particle swarm optimization [8]. However, these centralized algorithms are usually costly both in communication and computation since they require a complicate communication infrastructure to gather information globally, as well as a powerful central controller to process the large amount of data. Further, such a centralized control scheme is usually inflexible and thus unable to exhibit plug-and-play features [9].

To avoid such negative effects of centralized approaches, distributed algorithms have been proposed for the energy management. For example, the energy management problem is formulated as an incremental cost consensus problem in [10], and a Multi-agent system (MAS) based method has been proposed. It is not fully distributed since a leader agent has to be deployed to collect current power generated by each generator. To make the algorithm distributed, a novel consensus-based algorithm has been proposed in [11], where each generator can obtain all the information needed in a distributed fashion. Other distributed energy management approaches can be referred to [12], [13]. Notice that only conventional thermal generators with quadratic cost function are concerned about in the above mentioned literature. More recently, the authors in [14] have taken both conventional thermal generators and wind turbines into account in the energy management problem.

The aforementioned studies only focus on distributed generation side management,i.e., ED problem. In addition, many efforts have been made to demand response (see [15], [16] and reference therein, to just name a few). Traditionally, the power scheduling on the generation and demand sides is implemented separately. In the future power grid with integration of distributed energy resources, it is more practical and meaningful to consider both sides simultaneously. Because of this, some preliminary attempts have been made in [17], [18], where only conventional thermal generators are considered and the demands are temporally independent. However, in many occasions, the demand may be required to finish a given task by a deadline, that is to say, the cumulative power consumption must exceed a threshold by the given deadline. For example, an electric vehicle is required to charge at least 16 kWh before the next day for a 40-mile drive [19]. It implies that the power consumption of the demand at one time slot is temporally coupled with those at other time slots. Meanwhile, the smart grid will integrate wind power into the power system due to environmental friendliness of wind power and maturity of turbine techniques [20]–[22]. Thus, it is believed that the distributed energy management of the smart grid with wind power and temporal coupled demands needs to be further investigated.

Therefore, in this paper, a distributed energy management algorithm (DEMA) is proposed for the smart grid with a more practical scenario by taking wind power and the temporally coupled demands into account. We formulate the distributed energy management problem to maximize the social welfare by formulating the ED and DR in a united framework. The main challenges include the non-quadratic cost function of wind power, the temporally coupled demands, and the non-convexity of the optimization problem. To address these challenges, the original problem is first transformed into a convex optimization problem and then decoupled into simple subproblems that can be solved locally based on the Lagrange duality techniques. The global information needed for each agent can be discovered in finite steps in a distributed manner. Hence, the proposed method is fully distributed, and thus accommodates plug-and-play features. Meanwhile, The generators and demands are not required to reveal their private information such as incremental cost. The convergence and optimality of the proposed DEMA are guaranteed by theoretical analysis and validated by simulation studies on the IEEE 39-bus system.

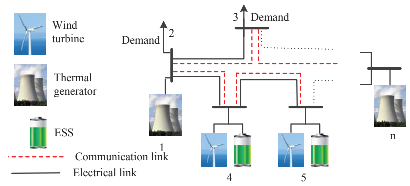

Fig. 1. Proposed smart grid structure.

II. Proposed Smart grid Structure

A smart grid system consisting of conventional thermal generators (TGs), wind turbines (WTs) and demands as shown in Fig. 1 is considered. Compared with [23], in order to enhance the controllability of the intermittent wind power, ESSs are usually utilized for wind power integration support, which is considered in this study. In this case, the cost of the wind turbine is a deterministic model considering the overestimation and underestimation costs of wind power [14], [24].

Each generator and demand are equipped with an agent. The communication among the agents can be modeled by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ [25], [26], where $\mathcal{V} = \{v_1, v_2, \cdots, v_n\}$ is the set of agents, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the set of edges. The edge $e_{ij} = (i, j) \in \mathcal{E}$ if and only if there is information exchange between agents i and j, and $e_{ij} \in \mathcal{E} \Leftrightarrow e_{ji} \in \mathcal{E}$. If there exists an edge between agents i and j, i and j are called adjacent. The adjacency matrix $A = (a_{ij}) \in R^{n\times n}$ is defined element-wise with $a_{ij} = 1$ if i and j are adjacent and $a_{ij} = 0$ otherwise. Throughout this paper, it is assumed that $a_{ii} = 0$. The neighbor set of agent i is denoted by $N_i = \{j \in \mathcal{V} | e_{ji} \in \mathcal{E}\}$.

III. Problem Formulation

The cycle of energy management time horizon is divided into T time slots and the length of each time slot is constant. At each time slot $t \in \mathcal{T} = \{1, 2, \cdots, T\}$, the electricity price is denoted by p_t.

A. Profit of Wind Turbine

By virtue of the integration of ESSs into the WTs and rapid development of control techniques [22], [27], the power output of the WT can be guaranteed to be equal to the scheduled wind power W_{it}. The overall cost for the i th WT can be expressed as [14], [28]

$$C_{it}(W_{it}) = \delta_i W_{it} + C_{pwi} E(Y_{it}^{ue}) + C_{rwi} E(Y_{it}^{oe}) \quad (1)$$

where $C_{pwi}E(Y_{it}^{ue})$ denotes the underestimation cost, namely the penalty cost for not using all the available wind power. The expression of $E(Y_{it}^{ue})$ is given by [28]

$$
\begin{aligned}
E\left(Y_{it}^{ue}\right) = &(W_r - W_{it})\left[\exp\left(-\frac{v_r^\kappa}{c^\kappa}\right) - \exp\left(-\frac{v_{out}^\kappa}{c^\kappa}\right)\right] \\
&+\left(\frac{W_r v_{in}}{v_r - v_{in}} + W_{it}\right)\left[\exp\left(-\frac{v_r^\kappa}{c^\kappa}\right)\right. \\
&\left.-\exp\left(-\frac{v_1^\kappa}{c^\kappa}\right)\right] + \frac{W_r c}{v_r - v_{in}}\left\{\Gamma\left[1+\frac{1}{\kappa},\left(\frac{v_1}{c}\right)^\kappa\right]\right. \\
&\left.-\Gamma\left[1+\frac{1}{\kappa},\left(\frac{v_r}{c}\right)^\kappa\right]\right\} \qquad (2)
\end{aligned}
$$

where $\exp(\cdot)$ stands for the exponential function, $\Gamma(\cdot)$ denotes the standard incomplete gamma function, and v_1 is an intermediary parameter defined as $v_1 = v_{in} + (v_r - v_{in}) W_{it}/W_r$. Likewise, $C_{rwi}E(Y_{it}^{oe})$ denotes the overestimation cost, namely the cost for purchasing the power from ESSs because of the deficit in wind power, The expression of $E\left(Y_{it}^{oe}\right)$ can be given as

$$
\begin{aligned}
E\left(Y_{it}^{oe}\right) = &W_{it}\left[1 - \exp\left(-\frac{v_{in}^\kappa}{c^\kappa}\right) + \exp\left(-\frac{v_{out}^\kappa}{c^\kappa}\right)\right] \\
&+\left(\frac{W_r v_{in}}{v_r - v_{in}} + W_{it}\right)\left[\exp\left(-\frac{v_{in}^\kappa}{c^\kappa}\right)\right. \\
&\left.-\exp\left(-\frac{v_1^\kappa}{c^\kappa}\right)\right] + \frac{W_r c}{v_r - v_{in}}\left\{\Gamma\left[1+\frac{1}{\kappa},\left(\frac{v_1}{c}\right)^\kappa\right]\right. \\
&\left.-\Gamma\left[1+\frac{1}{\kappa},\left(\frac{v_{in}}{c}\right)^\kappa\right]\right\} \qquad (3)
\end{aligned}
$$

Note that the subscript i in the above parameters such as $v_{in}, v_{out}, v_r, c, \kappa$ and W_r has been dropped for simplifying the notation. In power grids, transmission losses are inevitable and thus need to be considered. According to [29], the amount of transmission losses P_{it}^L induced by the i th WT at time slot t can be expressed as $P_{it}^L = B_i W_{it}^2$. Hence, the profit of wind turbine i at time slot t can be written as

$$\pi_{it}^w(W_{it}) = p_t\left(W_{it} - B_i W_{it}^2\right) - C_{it}(W_{it}) \qquad (4)$$

subject to

$$0 \le W_{it} \le W_r \qquad (5)$$

The grid-code constraints [30] need to be further investigated.

B. Profit of Thermal Generator

Referring to [11], the cost $C_{it}(P_{it})$ for each thermal generator i at time slot t can be expressed in the following quadratic form

$$C_{it}\left(P_{it}\right) = \alpha_i P_{it}^2 + \beta_i P_{it} + \gamma_i \qquad (6)$$

where α_i, β_i and γ_i are the cost coefficients. The amount of transmission losses P_{it}^L can be written as $P_{it}^L = B_i P_{it}^2$. Thereafter, the profit of thermal generator i at time slot t can be written as

$$\pi_{it}^g(P_{it}) = p_t\left(P_{it} - B_i P_{it}^2\right) - C_{it}(P_{it}) \qquad (7)$$

subject to

$$\underline{P}_{it} \le P_{it} \le \bar{P}_{it} \qquad (8)$$

where $\underline{P}_{it}$ and $\bar{P}_{it}$ are the lower and upper bounds of the thermal generator capability.

C. Profit of Demand

For each demand $i \in V_d$ at time slot $t \in \mathcal{T}$, we associate a non-decreasing, concave, and differentiable utility function $U_{it}(P_{it})$ of the form [31]

$$U_{it}\left(P_{it}\right) = \begin{cases} \omega_i P_{it} - \mu_i P_{it}^2, & P_{it} \le \frac{\omega_i}{2\mu_i} \\ \frac{\omega_i}{4\mu_i^2}, & P_{it} > \frac{\omega_i}{2\mu_i} \end{cases} \qquad (9)$$

where P_{it} is the power consumption level of the demand i at time slot t, and ω_i, μ_i are design parameters for the demand i. The demand is bounded such that

$$\underline{P}_{it} \le P_{it} \le \bar{P}_{it} \qquad (10)$$

where $\underline{P}_{it}$ and $\bar{P}_{it}$ are the lower and upper energy consumption bounds. Meanwhile, we consider the temporally-coupled constraint which couples the demand across the time horizon as

$$\sum_{t\in\mathcal{T}} P_{it} \ge r_i \qquad (11)$$

Thence, the profit of each demand i at time slot t is

$$\pi_{it}^d(P_{it}) = U_{it}(P_{it}) - p_t P_{it} \qquad (12)$$

D. Optimization Problem

The social welfare maximization problem is subject to the power balance constraint, i.e.,

$$\sum_{i\in V_w}\left(W_{it} - B_i W_{it}^2\right) + \sum_{i\in V_g}\left(P_{it} - B_i P_{it}^2\right) = \sum_{i\in V_d} P_{it} \qquad (13)$$

By recalling equations (4), (7), (12) and using (13), the sum profit of the generators and the demands across the time horizon can be obtained as

$$
\begin{aligned}
&\sum_{t\in\mathcal{T}}\left\{\sum_{i\in V_w}\pi_{it}^w(W_{it}) + \sum_{i\in V_g}\pi_{it}^g(P_{it}) + \sum_{i\in V_d}\pi_{it}^d(P_{it})\right\} \\
&=\sum_{t=1}^{T}\left\{-\sum_{i\in V_w}C_{it}(W_{it}) - \sum_{i\in V_g}C_{it}(P_{it}) + \sum_{i\in V_d}U_{it}(P_{it})\right\} \qquad (14)
\end{aligned}
$$

Hence, we formulate the following optimization problem

$$\min \sum_{t\in\mathcal{T}}\left\{\sum_{i\in V_w}C_{it}(W_{it}) + \sum_{i\in V_g}C_{it}(P_{it}) - \sum_{i\in V_d}U_{it}(P_{it})\right\}$$

$$s.t. \sum_{i\in V_w}\left(W_{it} - B_i W_{it}^2\right) + \sum_{i\in V_g}\left(P_{it} - B_i P_{it}^2\right) = \sum_{i\in V_d} P_{it} \qquad (15a)$$

$$0 \le W_{it} \le W_r \quad i \in V_w, \quad t \in \mathcal{T} \qquad (15b)$$

$$\underline{P}_{it} \le P_{it} \le \bar{P}_{it}, \quad i \in V_g \cup V_d, \quad t \in \mathcal{T} \qquad (15c)$$

$$\sum_{t\in\mathcal{T}} P_{it} \ge r_i \quad i \in V_d \qquad (15d)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2017.2682001, IEEE Transactions on Industrial Electronics

It has been proved in [28] that the wind power cost (1) is convex and $\partial C_{it}(W_{it})/\partial W_{it} > 0, \partial^2 C_{it}(W_{it})/\partial^2 W_{it} > 0$. Obviously, the thermal generator cost (6) is convex and the demand utility function (9) is concave. Hence, the objection function in (15) is a convex function. However, the optimization problem (15) is non-convex due to the non-convex feasible set caused by the equality quadratic constraint (15a).

IV. Distributed Algorithm Design

A. Problem Transformation

A sufficient condition, which can be satisfied in practical scenarios, is required for the problem transformation. Specifically, the maximal total demand should be larger than or equal to the total minimal generation subtracting power losses at each time slot $t \in \mathcal{T}$, i.e.,

$$\sum_{i \in V_d} \bar{P}_{it} \geq \sum_{i \in V_g} \left(\underline{P}_{it} - B_i \underline{P}_{it}^2 \right) \tag{16}$$

From (5), we have that the total minimal generation subtracting power losses of the wind turbines is equal to zero, and thus omitted in (16). In practical scenarios, the total minimal generation power of the thermal generators can also be chosen as zero, which implies that the sufficient condition (16) can always be satisfied. Thereafter, problem (15) can be converted to

$$\min \sum_{t \in \mathcal{T}} \left\{ \sum_{i \in V_w} C_{it}(W_{it}) + \sum_{i \in V_g} C_{it}(P_{it}) - \sum_{i \in V_d} U_{it}(P_{it}) \right\}$$

$$s.t. \sum_{i \in V_w} \left(W_{it} - B_i W_{it}^2 \right) + \sum_{i \in V_g} \left(P_{it} - B_i P_{it}^2 \right) \geq \sum_{i \in V_d} P_{it} \tag{17a}$$

$$(15b), (15c), \text{ and } (15d)$$

The feasible region of transformed problem (17) now belongs to a convex set, and thus problem (17) is a strictly convex problem. By analyzing the Karush-Kuhn-Tucker (KKT) conditions of problem (17) and problem (15) similar to [17], we can easily obtain that the optimal solution is always achieved for the converted problem (17) when (17a) takes the equal sign. It implies that problem (15) and problem (17) are equivalent. However, the transformed problem (17) still has the spatially coupled constraint (17a) and the temporally coupled constraint (15d), which make the problem difficult to tackle directly.

B. Subproblem Decomposition

For analysis convenience, we define $W \in \Re^{|V_w| \times T}$ with entries w_{it}, and $P \in \Re^{|V_g \cup V_d| \times T}$ with entries p_{it}. We also define $X = [W; P]$ denoting the decision variable matrix for the optimization problem (15), with $\underline{X}$ and $\bar{X}$ being the lower and upper bound matrixes of X, respectively. The subproblem decomposition is based on means of dual decomposition [32]. Specifically, the partial Lagrangian of the primal problem (17) is firstly derived as

$$\begin{aligned} \mathcal{L}(Z) = & \sum_{t \in \mathcal{T}} \left\{ \sum_{i \in V_w} C_{it}(W_{it}) + \sum_{i \in V_g} C_{it}(P_{it}) - \sum_{i \in V_d} U_{it}(P_{it}) \right\} \\ & + \sum_{t \in \mathcal{T}} \lambda_t \left(\sum_{i \in V_d} P_{it} - \sum_{i \in V_w} (W_{it} - B_i W_{it}^2) - \sum_{i \in V_g} (P_{it} \right. \\ & \left. - B_i P_{it}^2) \right) + \sum_{i \in V_d} \mu_i \left(r_i - \sum_{t \in \mathcal{T}} P_{it} \right) \end{aligned} \tag{18}$$

where $\lambda_t \geq 0$ is the Lagrangian multiplier at time slot t aimed at relaxing the supply demand constraint (17a), $\mu_i \geq 0$ is the Lagrangian multiplier for each demand aimed at relaxing the temporally coupled constraint (15d), while $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_T]$, and $\boldsymbol{\mu} = [\mu_1, \cdots, \mu_{|V_d|}]$ are the Lagrangian multiplier vectors. The dual function is the minimum value of the Lagrangian function over matrix X

$$\begin{aligned} \mathcal{D}(\boldsymbol{\lambda}, \boldsymbol{\mu}) = & \inf_{\underline{X} \leq X \leq \bar{X}} \mathcal{L}(X, \boldsymbol{\mu}, \boldsymbol{\lambda}) = \sum_{i \in \mathcal{V}_w \cup \mathcal{V}_g, t \in \mathcal{T}} S_{it}(\lambda_t) \\ & + \sum_{i \in \mathcal{V}_d, t \in \mathcal{T}} S_{it}(\lambda_t, \mu_i) + \sum_{i \in V_d} \mu_i r_i \end{aligned} \tag{19}$$

where S_{it} is defined as the (i,t)th subproblem to be solved by agent i at time slot t. From (18) and (19), the explicit form of subsystem S_{it} can be summarized as follows

a) For each wind turbine $i \in V_w$ at time slot $t \in \mathcal{T}$, its subproblem is

$$\begin{aligned} S_{it}(\lambda_t) \triangleq & \min_{W_{it}} C_{it}(W_{it}) - \lambda_t \left(W_{it} - B_i W_{it}^2 \right) \\ & \text{s.t.} \quad 0 \leq W_{it} \leq W_r \end{aligned} \tag{20}$$

b) For each thermal generator $i \in V_g$ at time slot $t \in \mathcal{T}$, its subproblem is

$$\begin{aligned} S_{it}(\lambda_t) \triangleq & \min_{P_{it}} C_{it}(P_{it}) - \lambda_t \left(P_{it} - B_i P_{it}^2 \right) \\ & \text{s.t.} \quad \underline{P}_{it} \leq P_{it} \leq \bar{P}_{it} \end{aligned} \tag{21}$$

c) For each demand $i \in V_d$ at time slot $t \in \mathcal{T}$, its subproblem is

$$\begin{aligned} S_{it}(\lambda_t, \mu_i) \triangleq & \min_{P_{it}} -U_{it}(P_{it}) + \lambda_t P_{it} - \mu_i P_{it} \\ & \text{s.t.} \quad \underline{P}_{it} \leq P_{it} \leq \bar{P}_{it} \end{aligned} \tag{22}$$

Each subproblem is independent without any coupled constraint as opposed to the problem (17). For each local optimization subproblem, given λ_t, μ_i, its optimal solution is

a) For each subproblem (20) for the wind turbine, its optimal solution is

$$\tilde{W}_{it} = \begin{cases} W_r, & \lambda_t \geq \frac{C'_{it}(W_r)}{1 - 2B_i W_r} \\ \left[\frac{C'_{it}(\tilde{W}_{it})}{1 - 2B_i \tilde{W}_{it}} \right]^{-1} (\lambda_t), & C'_{it}(0) < \lambda_t < \frac{C'_{it}(W_r)}{1 - 2B_i W_r} \\ 0, & \lambda_t \leq C'_{it}(0) \end{cases} \tag{23}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2017.2682001, IEEE Transactions on Industrial Electronics

b) For each subproblem (21) for the thermal generator, its optimal solution is

$$\tilde{P}_{it}=\begin{cases}\bar{P}_{it}, & \lambda_t \geq \frac{C'_{it}(\bar{P}_{it})}{1-2B_i\bar{P}_{it}}\\ \left[\frac{C'_{it}(\tilde{P}_{it})}{1-2B_i\tilde{P}_{it}}\right]^{-1}(\lambda_t), & \frac{C'_{it}(\underline{P}_{it})}{1-2B_i\underline{P}_{it}}<\lambda_t<\frac{C'_{it}(\bar{P}_{it})}{1-2B_i\bar{P}_{it}}\\ \underline{P}_{it}, & \lambda_t \leq \frac{C'_{it}(\underline{P}_{it})}{1-2B_i\underline{P}_{it}}\end{cases} \tag{24}$$

c) For each subproblem (22) for the demand, its optimal solution is

$$\tilde{P}_{it}=\begin{cases}\underline{P}_{it}, & \lambda_t-\mu_i \geq U'_{it}(\underline{P}_{it})\\ U'_{it}(\tilde{P}_{it})^{-1}(\lambda_t-\mu_i), & U'_{it}(\bar{P}_{it})<\lambda_t-\mu_i<U'_{it}(\underline{P}_{it})\\ \bar{P}_{it}, & \lambda_t-\mu_i \leq U'_{it}(\bar{P}_{it})\end{cases} \tag{25}$$

where $f(x)^{-1}$ denotes the inverse function of $f(x)$.

Theorem 1: Consider the subproblems (20), (21) and (22), their optimal solutions can be given by (23), (24) and (25), respectively.

Proof: See the Appendix A. ■

C. Update Laws of Lagrangian Multipliers

From the above analysis, we have that solving the dual problem (19) is sufficient to solve the primal problem (17) (or problem (15) due to the equivalence). For the differentiable dual function $\mathcal{D}(\boldsymbol{\lambda},\boldsymbol{\mu})$ in (19), we employ the gradient method to iteratively calculate the optimal solution. Specifically, the Lagrangian multipliers λ_t,μ_i are updated in the direction of a negative gradient of the form

$$\begin{aligned}\lambda_t(k+1)&=\left[\lambda_t(k)-\epsilon_\lambda\frac{\partial\mathcal{D}(\lambda,\mu)}{\partial\lambda_t(k)}\right]^+\\ \mu_i(k+1)&=\left[\mu_i(k)-\epsilon_\mu\frac{\partial\mathcal{D}(\lambda,\mu)}{\partial\mu_i(k)}\right]^+\end{aligned} \tag{26}$$

where $\epsilon_\lambda,\epsilon_\mu>0$ are the corresponding step sizes. It has been proved that the Lagrangian multipliers will converge to their optimal values $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ for a sufficiently small step, i.e., $\boldsymbol{\lambda}\rightarrow\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}\rightarrow\boldsymbol{\mu}^*$ when $k\rightarrow\infty$. The primal problem (15) and its dual problem (19) are strictly equivalent, due to its strong duality. It thus follows that the primal variables P and W calculated by (23), (24) and (25) will converge to their optimal values P^* and W^*, i.e., $P\rightarrow P^*$ and $W\rightarrow W^*$ when $k\rightarrow\infty$. Substituting the solutions (23), (24) and (25) into the subproblems (20), (21) and (22), we can obtain the gradient of the dual function from (18) and (19). Using the calculated gradient, the Lagrangian multiplier update rule can then be rewritten as

$$\lambda_t(k+1)=\left[\lambda_t(k)-\epsilon_\lambda\left\{\sum_{i\in V_d}\tilde{P}_{it}-\sum_{i\in V_w}(\tilde{W}_{it}-B_i\tilde{W}_{it}^2)-\sum_{i\in V_g}(\tilde{P}_{it}-B_i\tilde{P}_{it}^2)\right\}\right]^+ \tag{27}$$

$$\mu_i(k+1)=\left[\mu_i(k)-\epsilon_\mu\left(r_i-\sum_{t=1}^{T}\tilde{P}_{it}\right)\right]^+ \tag{28}$$

The update of Lagrangian multiplier λ_t requires the global smart grid information including the total demands $\sum_{i\in V_d}\tilde{P}_{it}$ and the total power generations subtracting power losses $\sum_{i\in V_w}(\tilde{W}_{it}-B_i\tilde{W}_{it}^2)+\sum_{i\in V_g}(\tilde{P}_{it}-B_i\tilde{P}_{it}^2)$. Hence, a distributed information discovery algorithm must be properly designed for each agent to obtain the global information.

D. Finite-time Global Information Discovery

The information discovery process of agent i can be represented as

$$s_i(k'+1)=\sum_{j=1}^{n}q_{ij}s_i(k') \tag{29}$$

where $s_i(k')$ denotes the local information discovered by agent i at iteration k'. q_{ij} is the communication coefficient between agents i and j. The communication coefficient is chosen as

$$q_{ij}=\begin{cases}\epsilon_s a_{ij}, & i\neq j\\ 1-\sum_{j=1,j\neq i}^{n}q_{ij}, & i=j\end{cases} \tag{30}$$

where $0<\epsilon_s<(1/\Delta)$ is a parameter with $\Delta=\max_i(\sum_{j\neq i}a_{ij})$. Referring to [25], we have that all s_i will converge to the the average value of initial states $s_i(0), i=1,\cdots,n$, i.e., $\lim_{k'\rightarrow\infty}s_i(k')=(1/n)\sum_{i=1}^{n}s_i(0)$. Consequently, by initializing s_i with local demand and local power generations subtracting power losses, and with the number of agents n discovered, we can obtain the required global information.

Algorithm 1: Global information discovery of agent i

1. Calculate the coefficients of the minimal polynomial of Q;
2. Obtain the vector η using (32);
3. Execute (29) until $k'=D$ and store $s_i(0),\cdots,s_i(D)$;
4. Calculate the final average value using (31);

The final value of (29) can be calculated as [33]

$$\lim_{k'\rightarrow\infty}s_i(k')=\frac{\mathbf{s}_D^T\eta}{\mathbf{1}^T\eta} \tag{31}$$

where $\mathbf{s}_D=[s_i(0),\cdots,s_i(D)]^T$ with $D+1$ being the degree of the minimal polynomial associated with the matrix Q. $\mathbf{1}\in R^D$ denotes the all-ones vector, and η is a vector given by

$$\eta=\begin{pmatrix}1\\ 1+\phi_D\\ 1+\phi_{D-1}+\phi_D\\ \vdots\\ 1+\sum_{j=1}^{D}\phi_j\end{pmatrix} \tag{32}$$

where $\phi_j, j=1,\cdots,D$ are the coefficients of the minimal polynomial of Q. From (31), we have that only D iterations (finite-time) are needed to determine the final value. Each agent is supposed to know the the number of agents and the structure of the network, which can be obtained using the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2017.2682001, IEEE Transactions on Industrial Electronics

distributed graph discovery algorithm in [14]. The distributed global information discovery procedure is summarized in Algorithm 1.

E. Distributed Energy Management Algorithm

For clarity, we summarize the distributed energy management algorithm (DEMA) in Algorithm 2. The convergence of Algorithm 2 is formally stated below.

Algorithm 2: Distributed EMA (DEMA)

repeat
- Implement Algorithm 1 to acquire $\sum_{i\in V_d} \tilde{P}_{it}$ and $\sum_{i\in V_w} (\tilde{W}_{it} - B_i \tilde{W}_{it}^2) + \sum_{i\in V_g} (\tilde{P}_{it} - B_i \tilde{P}_{it}^2)$;
- Update $\boldsymbol{\lambda}(k+1)$ according to (27);
- Update $\boldsymbol{\mu}(k+1)$ according to (28) ;
- Update $W_{it}(k+1)$ according to (23) when $i \in V_w$; Update $P_{it}(k+1)$ according to (24) when $i \in V_g$, according to (25) when $i \in V_d$;

until $\boldsymbol{\lambda}$ *converges as* $|\lambda_t(k+1) - \lambda_t(k)| \le \epsilon, \forall t$ *and* μ_i *converges as* $|\mu_i(k+1) - \mu_i(k)| \le \epsilon$;

Theorem 2: Consider the energy management problem given by (15), Algorithm 2 will converge to the optimal solution when $k \to \infty$.

Proof: See the Appendix B. ■

The parameter ϵ plays an important role in the stopping criteria of Algorithm 2. A smaller ϵ results in a more accurate optimal solution and thus needs more communication times. Note that the communication requirement determines the required time, which can be always satisfied since the proposed DEMA operates at a slow timescale [34] (See Remark 2 and 3 for more details). We can initially choose an initial value of ϵ, and then compare the solution with the optimal solution obtained by a centralized solution (e.g., using fmincon function in Matlab). If the deviation between the obtained solution and the optimal solution is sufficiently small, the value of ϵ is determined. Otherwise, choose a smaller ϵ and then evaluate the deviation between the new obtained solution and the optimal solution until the obtained solution satisfies the accuracy requirement. The determination of ϵ is given by Fig. 2, where the reducing method of ϵ is not unique.

Remark 1: Given a smart grid, the system parameters are supposed to be within a set, in which the ϵ determined from one group of system parameter is also applicable to other groups of system parameter. Note that the ϵ given in Fig. 2 is determined before implementing the distributed algorithm. At that stage, we can just collect information of generators together with loads and solve the problem using a centralized solution. After determining the value of ϵ, the distributed algorithm can be implemented using the same value ϵ to obtain the new solution when the system parameter changes. The assumption of system parameters within the aforementioned set is not strict since such a set can be made as large as desired by decreasing ϵ [32], [35]. It means that the chosen ϵ can work on a sufficiently large number of samples in the problem

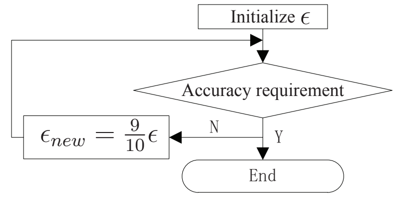

Fig. 2. Determination of the parameter ϵ.

space. Specifically, $\epsilon = 0$ is applicable to any sample in the problem space. This property can be analyzed as follows. The ϵ determines how many iterations are needed. This is because of the fact $\lambda_t(k+1) - \lambda_t(k) = 0, \mu_i(k+1) = \mu_i(k) = 0$ when the proposed DEMA converges. However, $\epsilon = 0$ is not preferred in practice since it results in unnecessary iterations and conservative use of communication [32].

Remark 2: Like most studies [11], [14], [15], [24] focused on the energy management problem, the proposed algorithm belongs to the tertiary control level that operates at a slow timescale (intervals of 5 minutes or longer) [34], and focuses on improving efficiency of power grids. By contrast, the voltage and frequency regulations belong to the primary and secondary control levels that operate at a fast timescale (from 30 seconds to a few minutes), and focus on stability. Hence, similar to [11], [14], [15], [24], the frequency and voltage regulations are not formulated in the energy management problem.

Remark 3: The time required for solving the optimization problem mainly depends on the frequency at which each agent is sending its messages, and the total communication times as shown below

$$\text{Time required} = \frac{1}{\text{Frequency of agents}} \times \text{Total communication times}$$

The frequency at which each agent is sending its messages is limited by the data rate of communications and the message size.

Under communication link failure condition, the global information discovery can be obtained by replacing q_{ij} with $q_{ij}(k')$ in (29) as [25]

$$s_i(k'+1) = \sum_{j=1}^{n} q_{ij}(k') s_i(k') \tag{33}$$

where

$$q_{ij}(k') = \begin{cases} \epsilon_s a_{ij}(k'), & i \neq j \\ 1 - \sum\limits_{j=1, j\neq i}^{n} q_{ij}(k'), & i = j \end{cases} \tag{34}$$

Under the time delay communication, the global information discovery can be designed as $s_i(k') = x_i(k')/y_i(k')$ with [36]

$$\begin{aligned} x_i(k'+1) =& q_{ii} x_i(k') \\ &+ \sum_{j\in N_i} \sum_{\tau=0}^{\bar{\tau}} q_{ij} x_j(k'-\tau) I_{k'-\tau,ij}(\tau) \end{aligned} \tag{35}$$

$$\begin{aligned} y_i(k'+1) =& q_{ii} y_i(k') \\ &+ \sum_{j\in N_i} \sum_{\tau=0}^{\bar{\tau}} q_{ij} y_j(k'-\tau) I_{k'-\tau,ij}(\tau) \end{aligned} \tag{36}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2017.2682001, IEEE Transactions on Industrial Electronics

where $q_{ij} = \frac{1}{|N_j|+1}$ for $i \in N_j \cup \{i\}$ with $|N_j|$ being the number of agent j's neighbors, and $I_{k'-\tau,ij}(\tau)$ is given by

$$I_{k'-\tau,ij}(\tau) = \begin{cases} 1, \text{ if } \tau_{ij}(k'-\tau) = \tau \\ 0, \text{ otherwise} \end{cases} \tag{37}$$

where τ_{ij} denotes the time delay in the communication link between agent i and agent j, and $\bar{\tau}$ denotes the upper bound of time delays. x_i is initialized with local demand (local power generations subtracting power losses) and y_i is initialized with 1.

V. VALIDATION STUDIES

Numerical analysis has also been carried out on the IEEE-39 bus system shown in Fig. 3 which has 3 wind turbines, 7 thermal generators and 18 demands. The characteristics of the demands and the generators are tabulated in Tables I,II, and III by referring to [14], [17]. Here, we set $\epsilon_\lambda = 0.001, \epsilon_\mu = 0.01$, and $\epsilon_s = 1/6.01$. The initial values $\lambda_t(0) = 5$, and $\mu_i(0) = 3$. It is assumed that $T = 6$, and the nodes $14, 16, 18$ and 19 are temporally-coupled constrained to complete a given task in the predefined time horizon. The required demands for the nodes $14, 16, 18$ and 19 are set as 190 kW, 250 kW, 310 kW and 310 kW, respectively.

TABLE I
PARAMETERS OF DEMANDS

Node	ω (¢/kWh)	μ (¢/kW^{2}h)	$\underline{P}$ (kW)	$\bar{P}$ (kW)
11	18.43	0.0545	50	100.43
12	13.17	0.0877	100	159.13
13	15.46	0.0547	40	80.56
14	10.03	0.1041	30	123.98
15	8.45	0.0870	80	109.55
16	15.38	0.0984	40	76.34
17	19.16	0.1564	80	137.93
18	16.85	0.0564	50	84.19
19	15.63	0.0950	50	104.06
20	6.75	0.0470	78	119.36
21	14.95	0.0970	103	176.19
22	5.87	0.0349	33	88.65
23	18.18	0.0879	99	175.31
24	15.08	0.0653	89	129.03
25	14.90	0.0897	123	167.42
26	19.45	0.1345	5	36.51
27	18.05	0.0924	43	79.38
28	15.83	0.1026	67	147.26

TABLE II
PARAMETERS OF THERMAL GENERATORS

Node	α (¢/kW^{2}h)	β (¢/kWh)	γ (¢/h)	$\underline{P}$ (kW)	$\bar{P}$ (kW)	B
1	0.0024	5.56	30	60	300.39	0.00021
2	0.0056	4.32	25	25	479.10	0.00031
3	0.0072	6.6	25	28	292.4	0.00011
4	0.0047	3.14	16	40	306.34	0.00022
5	0.0091	7.54	6	35	593.80	0.00041
7	0.0053	7.31	23	45	595.40	0.00045
10	0.0046	4.76	12	30	443.41	0.00012

TABLE III
PARAMETERS OF WIND TURBINES

Node	v_{in}	v_{out}	v_r	c	k
6, 8, 9	5	45	15	8	2
	d_j	C_{pwj}	C_{rwj}	W_r	B
	6	3.1	3.1	160	0.00033

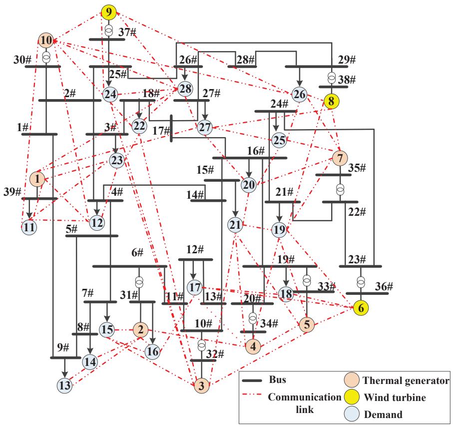

Fig. 3. IEEE 39-BUS System.

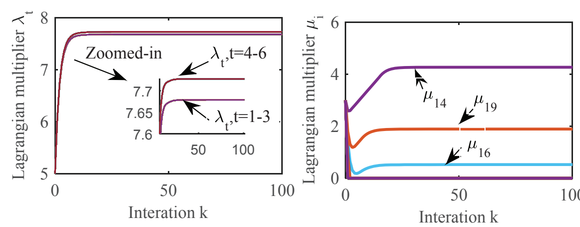

Fig. 4. Convergence of Lagrangian multipliers.

A. Convergence and Optimality

We first carry out numerical analysis to illustrate its convergence and optimality. Under this test, we assume that the lower bound $\underline{P}_{15}$ of node 15 is changed to 100 kW from $t = 4$. The convergence of the Lagrangian multipliers λ_t and μ_i is illustrated in Fig. 4. Fig. 5 shows the generator power, the demand power distribution, and the social welfare. Notice that the generator power and demand power distribution in Fig. 5 are for time slots $t = 1 - 3$ for demonstration purposes. Each demand power is shown as the negative one to distinguish from the generator units. The optimality of the proposed DEMA is verified by comparing our result with the optimal solution using global information. It can be observed that our proposed DEMA can converge to the optimal value in a distributed manner. Meanwhile, The power mismatch, defined as the total power generation minus the total demand will converge to zero with iterations as shown in Fig. 6. It implies that the power balance is maintained. The total demand over all time slots $\sum_{t=1}^{T} P_{it}$, the required demand r_i and the coordination parameter μ_i for the temporally-coupled constrained nodes are shown in Fig. 7. We can observe that the temporally coupled constraints are satisfied, and the coordination parameter μ_i is equal to zero when the required demand is overestimated. For example, the total demand is larger than the required for the node 18, and thus $\mu_{18} = 0$.

Note that the local load information at $k = 0$ for each demand agent can be obtained by solving (25) with initial values $\lambda_t(0)$ and $\mu_i(0)$, and the local load for each generator agent is zero. Hence, it can be easily verified that the average value of loads at $k = 0, t = 1$ is 51.6167 kW. Fig. 8 shows the performance of the global load information discovery at

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2017.2682001, IEEE Transactions on Industrial Electronics

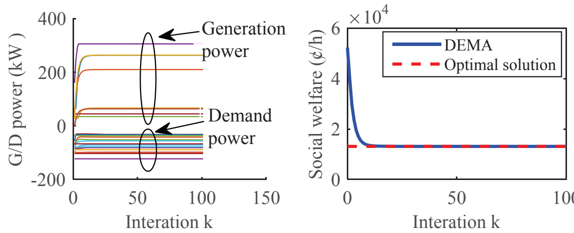

Fig. 5. Optimality of DEMA.

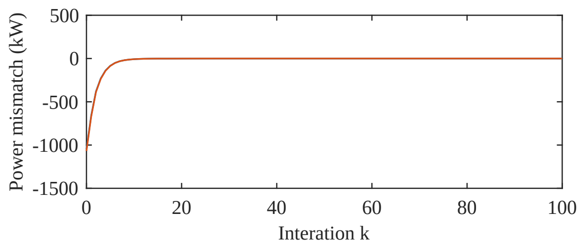

Fig. 6. Power mismatch under the DEMA.

iteration $k = 0$ and time slot t. It can be observed that all the local loads of agents can converge to the average value.

B. Plug-and-Play Capability

Plug and play adaptability is one of the most important features of smart grids. Without loss of generality, the node 28 plugs out at time step $k = 20$ due to faults. The other nodes detect this change and update their information. Then, the remaining agents ensure the convergence of the Lagrangian multipliers to new optimized values as shown in Fig. 9. The social welfare can also converge to the new optimal value as shown in Fig. 10. At the time step $k = 100$, the node 28 is plugged in again due to recovery and all the values are back to the previous values before the node 28 is disconnected.

C. Different Communication Conditions

In this test, different communication conditions including the reliable communication using (29), link failure communication using (33), and time delay communication using (35) (36) are simulated. The time delay in each communication link is chosen randomly from a uniform distribution on $[0, 3]$. It can be observed from Fig. 11 that the proposed DEMA can converge to the optimal value under different communication conditions. The main difference lies in the required communication times. More specifically, under the reliable communication using (29), the proposed DEMA needs 28 communication times at each step k since the degree of the minimal polynomial associated with the matrix Q for the communication structure given in Fig. 3 is 29. Under the link failure communication and time delay communication, the algorithm respectively needs 200 and 400 communication times at each step.

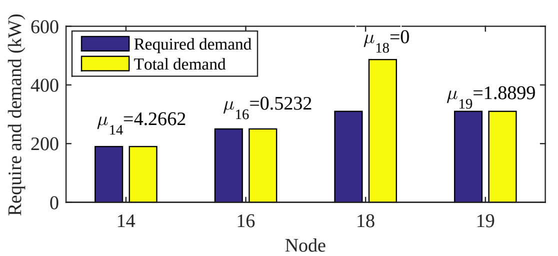

Fig. 7. Required demand, the total demand over all time slots, and the coordination parameter for temporally-coupled constrained nodes

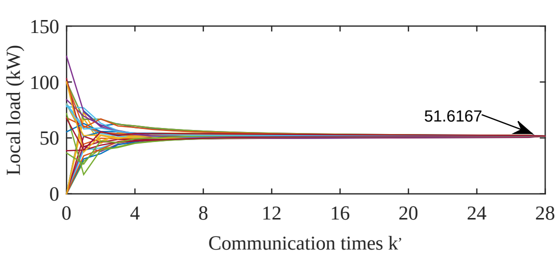

Fig. 8. Global information discovery at $k = 0, t = 1$.

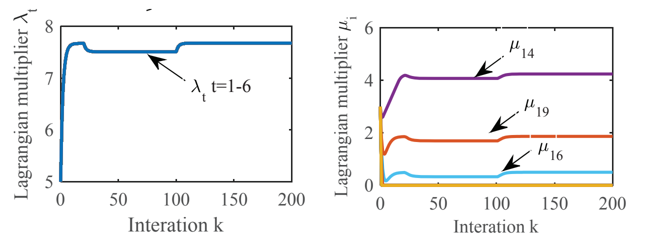

Fig. 9. Convergence of Lagrangian multipliers under the plug and play test.

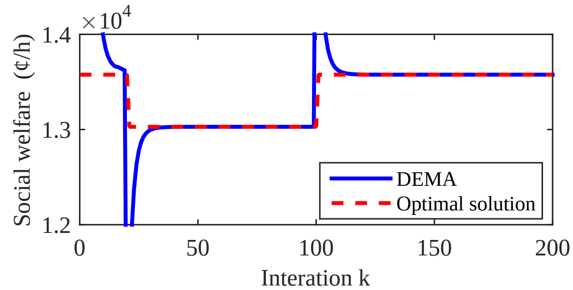

Fig. 10. Optimality of DEMA under the plug and play test.

D. A Four DG System With Detailed Electrical Models

In order to verify the practicalness of the proposed DEMA, numerical analysis has also been carried out on a four distributed generator (DG) system shown in Fig. 12 with more practical models, whose characteristics are tabulated in Table IV. The circuit breaker (CB) is supposed to be open at 0.3 s, implying that the microgrid operates in islanded mode from $t = 0.3$ s. The communication topology is chosen as the same to the power line. The parameters in the DEMA are given by Table V-Table VII. It is assumed that $T = 6$, and each time slot has 5 minutes. The nodes $5, 6, 7$ and 8 are

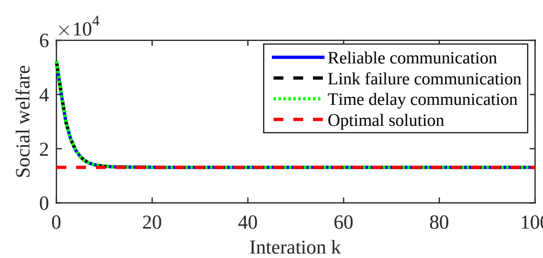

Fig. 11. Optimality of DEMA under the different communication condition test.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2017.2682001, IEEE Transactions on Industrial Electronics

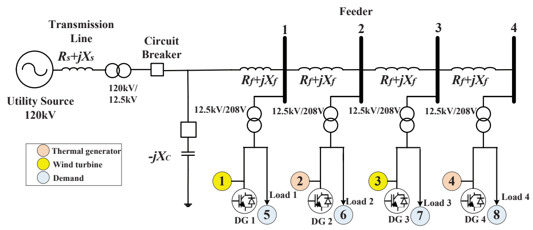

Fig. 12. A four DG system.

temporally-coupled constrained to complete a given task in the predefined time horizon. The required demands for the nodes $5, 6, 7$ and 8 are set as 9000 kW, 11400 kW, 9000 kW and 3000 kW, respectively. Fig. 13 and Fig. 14 show that both the social welfare maximization and the stability of frequency and voltage can be guaranteed. For clarity, we only show the response of frequency and voltage before 20s.

TABLE IV
DISTRIBUTION SYSTEM PARAMETERS

Parameters	Value	Parameters	Value
S_{base}	10 (MVA)	V_{base1}	$120\sqrt{2}/\sqrt{3}$ (kV)
V_{base2}	$12.5\sqrt{2}/\sqrt{3}$ (kV)	V_{base3}	$208\sqrt{2}/\sqrt{3}$ (V)
R_S	1.73×10^{-6} (p.u.)	X_S	3.47×10^{-5} (p.u.)
R_f	0.0029 (p.u.)	X_f	0.0041 (p.u.)

TABLE V
PARAMETERS OF THERMAL GENERATORS

Node	α (¢/kW^{2}h)	β (¢/kWh)	γ (¢/h)	$\underline{P}$ (kW)	$\bar{P}$ (kW)	B
2	0.0032	4.32	25	25	1600	0.000023
4	0.0024	5.56	30	60	1600	0.000013

TABLE VI
PARAMETERS OF WIND TURBINES

Node					
Node	v_{in}	v_{out}	v_r	c	k
$1, 3$	5	45	15	8	2
	d_j	C_{pwj}	C_{rwj}	W_r	B
	6	3.1	3.1	1350	0.00023

TABLE VII
PARAMETERS OF DEMANDS

Node	ω (¢/kWh)	μ (¢/kW^{2}h)	$\underline{P}$ (kW)	$\bar{P}$ (kW)
5	15.63	0.0950	50	2000
6	6.75	0.0470	78	2000
7	14.95	0.0970	103	2000
8	5.87	0.0349	33	500

VI. CONCLUSION

By taking wind power and the temporally coupled demands into consideration, we presented a distributed energy management algorithm for the smart grid. In our formulation, we consider a more practical scenario where both spatially and temporally coupled constraints are concerned. Specifically, the spatially coupled constraint means that the generation units and the demands have to maintain power balance for safe operation. The temporally coupled constraint means that the cumulative power consumption of a demand must exceed a threshold by a deadline in order to finish one task. The propose DEMA belongs to the distributed methods and thus accommodates plug-and-play features. We have proved the convergence and optimality of the proposed DEMA and verified it through

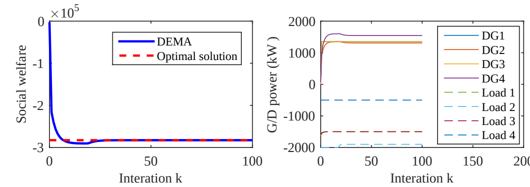

Fig. 13. Power dispatch result and social welfare of the four DG system.

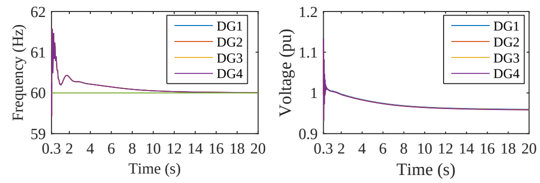

Fig. 14. Frequency and voltage response of the four DG system.

numerical analysis on the IEEE 39-bus system. More practical modeling and convergence improvement will be our future work.

APPENDIX A
PROOF OF THEOREM 1

Proof: Take the wind turbine subsystem (20) for an example, its Lagrangian function is

$$\mathcal{L} = C_{it}(W_{it}) - \lambda_t \left(W_{it} - B_i W_{it}^2\right) + \nu_1 \left(W_{it} - W_r\right) + \nu_2 \left(0 - W_{it}\right) \quad (38)$$

where $\nu_1, \nu_2 \geq 0$ are the dual variables. According to the KKT conditions of (20), we have

$$C_i^{'}(\tilde{W}_{it}) - \lambda_t + 2B_i\lambda_t\tilde{W}_{it} + \nu_1^* - \nu_2^* = 0 \quad (39)$$

$$\nu_1^* \left(\tilde{W}_{it} - W_r\right) = 0, \nu_1^* \geq 0 \quad (40)$$

$$\nu_2^* \left(0 - \tilde{W}_{it}\right) = 0, \nu_2^* \geq 0 \quad (41)$$

where $\tilde{W}_{it}, \nu_1^*, \nu_2^*$ are the optimal value of W_{it}, ν_1, ν_2, respectively. Notice that we use $\tilde{W}_{it}$ for the optimal value of W_{it} in the subsystem (20) while using W_{it}^* in the total problem (15).

1) If $0 < \tilde{W}_{it} < W_r$, we can infer from (40) (41) that $\nu_1^* = \nu_2^* = 0$. The equation (39) can then be written as $C_i^{'}(\tilde{W}_{it}) - \lambda_t + 2B_i\lambda_t\tilde{W}_{it} = 0$, from which we can easily obtain that $\frac{C_i^{'}(0)}{1-2B_i\times 0} < \lambda_t = \frac{C_i^{'}(\tilde{W}_{it})}{1-2B_i\tilde{W}_{it}} < \frac{C_i^{'}(W_r)}{1-2B_iW_r}$. Hence, in this case, we have $\tilde{W}_{it} = \left[\frac{C_{it}^{'}(\tilde{W}_{it})}{1-2B_i\tilde{W}_{it}}\right]^{-1}(\lambda_t)$.
2) If $\tilde{W}_{it} = W_r$, we can infer from (40) (41) that $\nu_1^* \geq 0, \nu_2^* = 0$. The equation (39) can then be written as $C_i^{'}(W_r) - \lambda_t + 2B_i\lambda_t W_r + \nu_1^* = 0$. It thus follows that $\lambda_t = \frac{C_i^{'}(W_r)}{1-2B_iW_r} + \frac{\nu_1^*}{1-2B_iW_r} > \frac{C_i^{'}(W_r)}{1-2B_iW_r}$.
3) Similar to case 2, if $\tilde{W}_{it} = 0$, we have $\lambda_t < \frac{C_i^{'}(0)}{1-2B_i\times 0}$.

The analysis for the thermal generator and the demand subsystems are similar to the above analysis and thus are omitted here. ■

APPENDIX B
PROOF OF THEOREM 2

Proof: By referring to [32], we have that the gradient update laws (27) and (28) converge to the optimal solutions

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2017.2682001, IEEE Transactions on Industrial Electronics

$\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ of the dual problem (19) for a sufficiently small step size. Due to the strong duality property, we have that W_{it} and P_{it} converge to the optimal solutions W_{it}^* and P_{it}^* of the problem (17), which has the same solution with the problem (15). ■

REFERENCES

[1] H. H. Abdeltawab and Y. A. R. I. Mohamed, "Market-oriented energy management of a hybrid wind-battery energy storage system via model predictive control with constraint optimizer," *IEEE Trans. Ind. Electron.*, vol. 62, no. 11, pp. 6658–6670, Nov. 2015.

[2] G. Zhabelova, V. Vyatkin, and V. N. Dubinin, "Toward industrially usable agent technology for smart grid automation," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2629–2641, Apr. 2015.

[3] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal dos attack scheduling in wireless networked control system," *IEEE Trans. Contr. Syst. Technol.*, vol. 24, no. 3, pp. 843–852, May. 2016.

[4] S. Liu, X. Wang, and P. X. Liu, "Impact of communication delays on secondary frequency control in an islanded microgrid," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2021–2031, Apr. 2015.

[5] S. He, J. Chen, X. Li, X. . Shen, and Y. Sun, "Mobility and intruder prior information improving the barrier coverage of sparse sensor networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 6, pp. 1268–1282, Jun. 2014.

[6] J.-Y. Fan and L. Zhang, "Real-time economic dispatch with line flow and emission constraints using quadratic programming," *IEEE Trans. Power Syst.*, vol. 13, no. 2, pp. 320–325, May. 1998.

[7] T. Guo, M. I. Henwood, and M. van Ooijen, "An algorithm for combined heat and power economic dispatch," *IEEE Trans. Power Syst.*, vol. 11, no. 4, pp. 1778–1784, Nov. 1996.

[8] K. T. Chaturvedi, M. Pandit, and L. Srivastava, "Self-organizing hierarchical particle swarm optimization for nonconvex economic dispatch," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1079–1087, Aug. 2008.

[9] W. Meng, X. Wang, and S. Liu, "Distributed load sharing of an inverter-based microgrid with reduced communication," *IEEE Trans. Smart Grid*, DOI: 10.1109/TSG.2016.2587685.

[10] 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.

[11] 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.

[12] G. Chen, F. Lewis, E. Feng, and Y. Song, "Distributed optimal active power control of multiple generation systems," *IEEE Trans. Ind. Electron.*, vol. 62, no. 11, pp. 7079–7090, Nov. 2015.

[13] V. Loia and A. Vaccaro, "Decentralized economic dispatch in smart grids by self-organizing dynamic agents," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 44, no. 4, pp. 397–408, Apr. 2014.

[14] F. Guo, C. Wen, J. Mao, and Y. D. Song, "Distributed economic dispatch for smart grids with random wind power," *IEEE Trans. Smart Grid*, vol. 7, no. 3, pp. 1572 – 1583, May. 2016.

[15] R. Deng, G. Xiao, R. Lu, and J. Chen, "Fast distributed demand response with spatially and temporally coupled constraints in smart grid," *IEEE Trans. Ind. Inf.*, vol. 11, no. 6, pp. 1597–1606, Dec. 2015.

[16] B. Chai, J. Chen, Z. Yang, and Y. Zhang, "Demand response management with multiple utility companies: A two-level game approach," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 722–731, Mar. 2014.

[17] C. Zhao, J. He, P. Cheng, and J. Chen, "Consensus-based energy management in smart grid with transmission losses and directed communication," *IEEE Trans. Smart Grid*, DOI:10.1109/TSG.2015.2513772.

[18] N. Rahbari-Asr, U. Ojha, Z. Zhang, and M. Y. Chow, "Incremental welfare consensus algorithm for cooperative distributed generation/demand response in smart grid," *IEEE Trans. Smart Grid*, vol. 5, no. 6, pp. 2836–2845, Nov. 2014.

[19] A. Ipakchi and F. Albuyeh, "Grid of the future," *IEEE Power Energy Mag.*, vol. 7, no. 2, pp. 52–62, 2009.

[20] W. Meng, Q. Yang, Y. Ying, Y. Sun, Z. Yang, and Y. Sun, "Adaptive power capture control of variable-speed wind energy conversion systems with guaranteed transient and steady-state performance," *IEEE Trans. Energy Convers*, vol. 28, no. 3, pp. 716–725, Sep. 2013.

[21] H. Jafarnejadsani and J. Pieper, "Gain-scheduled -optimal control of variable-speed-variable-pitch wind turbines," *IEEE Trans. Contr. Syst. Technol.*, vol. 23, no. 1, pp. 372–379, Jan. 2015.

[22] W. Meng, Q. Yang, and Y. Sun, "Guaranteed performance control of DFIG variable-speed wind turbines," *IEEE Trans. Contr. Syst. Technol.*, vol. 24, no. 6, pp. 2215–2223, Nov. 2016.

[23] X. Liu and W. Xu, "Economic load dispatch constrained by wind power availability: A here-and-now approach," *IEEE Trans. Sustain. Energy*, vol. 1, no. 1, pp. 2–9, Apr. 2010.

[24] J. Hetzer, D. C. Yu, and K. Bhattarai, "An economic dispatch model incorporating wind power," *IEEE Trans. Energy Convers.*,, vol. 23, no. 2, pp. 603–611, Jun. 2008.

[25] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[26] J. Qin, C. Yu, and H. Gao, "Coordination for linear multiagent systems with dynamic interaction topology in the leader-following framework," *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2412–2422, May. 2014.

[27] W. Meng, Q. Yang, S. Jagannathan, and Y. Sun, "Adaptive neural control of high-order uncertain nonaffine systems: A transformation to affine systems approach," *Automatica*, vol. 50, no. 5, pp. 1473–1480, May. 2014.

[28] X. Liu and W. Xu, "Minimum emission dispatch constrained by stochastic wind power availability and cost," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1705–1713, Aug. 2010.

[29] S. A.-H. Soliman and A.-A. H. Mantawy, *Modern optimization techniques with applications in electric power systems*, Berlin, Germany: Springer, Dec. 2011.

[30] H. H. Abdeltawab and Y. A. R. I. Mohamed, "Robust energy management of a hybrid wind and flywheel energy storage system considering flywheel power losses minimization and grid-code constraints," *IEEE Trans. Ind. Electron.*, vol. 63, no. 7, pp. 4242–4254, Jul. 2016.

[31] P. Samadi, A.-H. Mohsenian-Rad, R. Schober, V. W. Wong, and J. Jatskevich, "Optimal real-time pricing algorithm based on utility maximization for smart grid," in *Proc. SmartGridCom*, pp. 415–420. Gaithersburg, MD, USA, 2010.

[32] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[33] S. Sundaram and C. N. Hadjicostis, "Finite-time distributed consensus in graphs with time-invariant topologies," in *American Control Conference, 2007. ACC '07*, pp. 711–716, Jul. 2007.

[34] A. R. Bergen, *Power systems analysis*. Pearson Education India, 2009.

[35] 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.

[36] 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.

Wenchao Meng received the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2015.

He is currently a Postdoctoral Scholar with Carleton University, Ottawa, ON, Canada. His research interests include intelligent control and smart grids.

Xiaoyu Wang (M'08-SM'13) received the B.Sc. and M.Sc. degrees from Tsinghua University, Beijing, China, in 2000 and 2003, respectively, and the Ph.D. degree from the University of Alberta, Edmonton, AB, Canada, in 2008.

He is currently an Associate Professor with the Department of Electronics, Faculty of Engineering and Design, Carleton University, Ottawa, ON, Canada. His research interests include the integration of distributed energy resources and power quality.