On Hierarchical Power Scheduling for the Macrogrid and Cooperative Microgrids

Yu Wang, *Student Member, IEEE*, Shiwen Mao, *Senior Member, IEEE*, and R. M. Nelms, *Fellow, IEEE*

***Abstract*—Although considerable advances have been made in single microgrid (MG) systems, the problem of cooperation among MGs and the macrogrid has attracted considerable interest only recently. As in wireless communications systems, exploiting the temporal, spatial, and technological diversities in multiple cooperative MGs could bring about more efficient power generation and distribution. This paper investigates a hierarchical power scheduling approach to optimally manage power trading, storage, and distribution in a smart power grid with a macrogrid and cooperative MGs. We first formulate the problem as a convex optimization problem and then decompose it into a two-tier formulation. The first-tier problem jointly considers user utility, transmission cost, and grid load variance, while the second-tier problem minimizes the power generation and transmission cost, and exploits distributed storage in the MGs. We develop an effective online algorithm to solve the first-tier problem and prove its asymptotic optimality, as well as a distributed optimal algorithm for solving the second-tier problem. The proposed algorithms are evaluated with trace-driven simulations and are shown to outperform several existing schemes with considerable gains.**

***Index Terms*—Cooperation, demand response, diversity, microgrid (MG), online algorithm, smart grid (SG).**

I. Introduction

THE DECENTRALIZED generation at most renewable energy sources and the supporting technologies such as photovoltaics and microturbines have driven the demand for a new distributed power grid system, the *microgrid* (MG) [1]. Unlike traditional centralized power generation, the MG features distributed generation (DG) to support local users. DG is the basis of distributed energy resource (DER) systems, which is usually comprised of small power units, such as microturbines (25–100 kW) and small photovoltaic panels (1–10 kW). An MG can operate either in the *island* mode, where the local demand is supported with the MG's own DG and power storage, or the *grid-connected* mode, where the MG can acquire energy from, and/or contribute extra power to the macrogrid [2]. MG is regarded as an important paradigm for the next generation power grid, the smart grid (SG) [3]–[5]. SG technologies, such as smart metering, communications, and distributed control, will speed up the integration of MGs, and thus the penetration of DGs.

Over the past decade, MGs are built, experimented, and tested around the world [6]. In a single MG, research works cover several main topics, including interface or coupling between an MG and the macrogrid, DER dispatching and power support, and energy management. In [2], the MG control strategies and energy management are examined from several aspects. In [7], a detailed report is presented to test the building and management of a hydrogen MG in Spain in a simple and reliable way. In [8], Tsikalakis and Hatziargyriou present a control operation for a centralized controller for MGs, which maximizes its value by optimizing the production of local DGs and power exchanges with the main distribution grid during interconnected operation. In [9], the authors introduce an economic power dispatching scheme for stable operation of an MG, while a multiagent system is presented in [10] for DER energy management in an MG. Bahrani *et al*. [11] propose a multivariable digital control design methodology for the voltage regulation of an islanded single DG unit MG and its dedicated load.

Although more works are focused on the optimization and control of a single MG [12]–[15], the problem of *cooperation* among MGs and the macrogrid has attracted considerable interest recently. With such cooperation, MGs and the macrogrid will each gain tremendous benefits, such as reduced power loss, lower operational cost, and load peak reduction [16]–[22]. The obvious advantages stem from exploiting the *temporal, spatial*, and *technological diversities* in a multiple MG system. For instance, an MG supporting a business area will have a very different temporal demand profile from that of an MG supporting a residential area; the DGs in geographically distributed MGs can also have different generation levels at same time of the day; and different DGs are affected by weather differently: an MG with a photovoltaic array may suffer low generation during a storm, while a neighboring microturbine-based MG, caught in the same storm, may generate a large amount of power exceeding its own demand. As in wireless communication systems, exploiting such diversity through MG cooperation could bring about more efficient power generation and distribution.

The problem of cooperative MGs has been considered in several recent papers. In [18], the authors present a decentralized control strategy modeling the MGs as a team of cooperative agents to minimize the costs of energy storage and the power exchanged among the MGs. Wei *et al*. [19] propose a game theoretic coalition to optimally reduce the total power losses in a MGs power system with power storage devices, and demonstrate the overhead of communications. In [21], the authors

Manuscript received August 27, 2014; revised December 16, 2014; accepted March 22, 2015. Date of publication March 29, 2015; date of current version December 02, 2015. This work was supported in part by the U.S. National Science Foundation (NSF) under Grant CNS-0953513, and in part by the NSF Broadband Wireless Access and Application Center (BWAC) Site at Auburn University, Auburn, AL, USA. This work was presented in part as a poster at the IEEE International Conference on Mobile Ad hoc and Sensor Systems, Philadelphia, PA, USA. Paper no. TII-14-0903.

The authors are with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849-5201 USA (e-mail: yzw0032@tigermail.auburn.edu; smao@ieee.org; nelmsrm@auburn.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2015.2417496

formulate the optimal decision making problem in cooperative MG networks as a linear quadratic Gaussian problem. There have been some recent works that consider the power flow between the MGs and the macrogrid. For instance in [22], the authors present an optimal energy management framework for a cooperative network of heterogeneous MGs to achieve an efficient tradeoff between low operation cost and good energy service for customers.

The power grid is currently under a transition from traditional centralized distribution to decentralized distribution. In practice, the DG in MGs is usually not able to generate power stably and constantly. On the other hand, MGs can provide surplus power to the macrogrid. Therefore, it is important to incorporate all the key factors in a holistic manner, i.e., the generation cost, power generation and transmission losses, load smoothing, distributed storage, and the utility of power users. A control strategy would be highly desired that considers all the key factors for both the macrogrid and MGs.

In this paper, we consider a power grid consisting of the macrogrid and several cooperative MGs. The goal is to exploit *MG diversity gain* to optimize both the MG performance and user satisfaction. With cooperation, an MG is able to share its excess power with other MGs nearby or with the macrogrid. Due to limited storage capacity, the MG can sell its extra power to other MGs suffering power shortage. Alternatively, the MG could buy power from other MGs as well when its DG suffers low generation, such that both the power loss and the cost can be reduced compared to buying power directly from the macrogrid. On the other hand, the macrogrid could provide more storage capacity for the MGs, while the extra power from the MGs will in turn reduce the need of traditional power generation in the macrogrid. Grid load smoothness of the macrogrid could be achieved if the power flows from/to the MGs are optimally managed and scheduled.

In particular, under some mild assumptions, we formulate the cooperative MG problem as a convex optimization problem by capturing the key factors in a grid system, i.e., operation cost, power generation and transmission losses, user utility, distributed storage, and grid load smoothing. We then decompose the original problem into a two-tier power control problem. The first-tier control is for the macrogrid, aiming to maximize user utility, minimize power transmission cost from/to the macrogrid, and smooth the grid load of the macrogrid. The second-tier control is for each MG, aiming to minimize the cost of the MGs for power generation and transmission, while guaranteeing the power demand of MG users. It balances the power level with the macrogrid and makes energy trading and storage decisions within the MG network.

The power flow between MGs and the macrogrid is on one side, the power injected from outside of the MG network for MGs, and on the other side, a special load for the macrogrid, which is positive as usual if the power is transmitted to the MGs out from the macrogrid, and is negative reversely. This way, the two-tier controls are well integrated. For the first-tier problem, we develop an effective online algorithm that does not require any future information and is proven to be asymptotically optimal; for the second-tier problem, we develop a distributed algorithm for optimal solutions. The performance of the proposed hierarchical power scheduling scheme is validated with trace-driven simulations, where fast convergence and superior performance over several comparison schemes are observed.

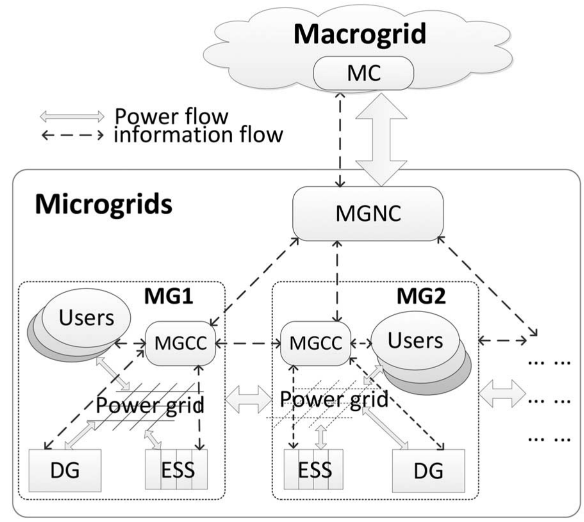

Fig. 1. Illustration of the power grid network.

This paper is organized as follows. We present the system model and problem formulation in Section II. We develop the asymptotically optimal online algorithm for the macrogrid in Section III, a distributed algorithm for cooperative MGs in Section IV, and present these algorithms in Section V. Performance evaluation is presented in Section VI. Section VII concludes this paper.

II. Problem Statement

A. System Model

We consider a power grid system with one macrogrid and many MGs as shown in Fig. 1. The macrogrid supports its own set of power users mostly with the traditional power generation sources. The macrogrid controller (MC) collects information from the smart meters at the macrogrid users to optimally distribute power to the users, and from the MG control center (MGCC) to trade power with the MGs.

As shown in Fig. 1, each MG consists of one or more DGs, an energy storage system (ESS), a smart infrastructure (such as smart meters and communication links), a set of users, and an MGCC. The energy users demand power from a designated MG. The independent DGs generate power to support the demand inside the MG. The ESS stores the extra power and discharges to satisfy the excess demand exceeding the DG generation. It usually consists of many batteries as well as some PHEVs. Furthermore, for excessive needs from its users, each MG will request and buy energy from other MGs or from the macrogrid. The MGCC in each MG controls the power distribution of the entire MG. Note that in an MG of the SG environment, information flow is essential for the control and cooperation. Both energy and information flows are enabled among the MGs and the macrogrid. The MGCC acquires user demand from the smart meters on the user side through a wired or wireless communication network. The MGCC also decides to sell or store excessive power based on the grid information. In this way, the energy generated among the MGs could be used efficiently to minimize the power from the macrogrid and to support the macrogrid needs when possible.

While MC and MGCC are the core of power scheduling inside the macrogrid and MGs, respectively, the MG network controller (MGNC) works both as a controller of the MG network and a bridge between the macrogrid and the MGs. It coordinates the information exchange and power transmissions between the macrogrid and the cooperative MGs.

B. Problem Formulation

We assume a time slotted system with $\mathbb{T} = \{1, 2, \ldots, T\}$ time slots. Let $\mathbb{N} = \{1, 2, \ldots, N\}$ denote the set of independent power users in the macrogrid. Each user $i \in \mathbb{N}$ demands power $d_i(t)$ at time t. Let $U(d_i(t), \omega_i(t))$ be the utility function for user i, which indicates the users' overall satisfactory level, and is a concave and strictly increasing function of $d_i(t)$. The parameter $\omega_i(t) \in (0, 1)$ denotes user i's level of flexibility, while a larger number closer to 1 (0) indicating a higher (lower) level of flexibility. The function $G(\cdot)$ indicates the generation cost in the macrogrid, which is strictly convex and increasing. In practice, a quadratic function is used; see our previous work [23] for more details on the utility function and generation cost function.

We use $\mathbb{M} = \{1, 2, \ldots, M\}$ to denote the set of all MGs. Each MG m has N_m users with total demand $d_m(t)$ at time t. Unlike the macrogrid, the power supply in MGs may be unstable in many cases, and thus, it provides more flexibility to consider users' demand as a whole. Let $p_m(t)$ be the power load in the macrogrid, transmitted to or received from MG $m \in \mathbb{M}$ at time t. It is positive when power is transmitted from the macrogrid to MG m, and is negative for the reverse direction. We define $\sigma_{0m} \in (0, 1)$ [19] as the ratio of transmission loss from the macrogrid to MG m, and $\sigma_{m0} \in (0, 1)$ as from the MG m to the macrogrid. And to simplify the expression, we define $\sigma_m = \sigma_{0m}$ if $p_m(t) > 0$, and $\sigma_m = \sigma_{m0}$ otherwise. Thus, when $p_m(t)$ is positive, the power received in MG m is $p_m(t)\sigma_m$; when $p_m(t)$ is negative, the generation in MG m is $p_m(t)/\sigma_m$. Let $p_{km}(t)$ denote the power received in MG m from MG k at time t, and $p_{mm}(t)$ be the power generated and used in MG M by itself. Similarly, $\sigma_{km} \in (0, 1)$ denotes the ratio of transmission loss between MG k and MG m.[1] Therefore, $\sum_{m\in\mathbb{M}} p_{km}(t)/\sigma_{km}$ is the total power generated in MG k for MGs, and $\sum_{k\in\mathbb{M}} p_{km}(t)$ is the total power in MG m received from all the MGs.

We use a general convex function $C_m(\cdot)$ to represent the transmission cost between the macrogrid and MG m because it costs more for the same amount of loss of power as the total power loss increases to a higher level. Similarly, we use convex functions $G_k(\cdot)$ and $C_{km}(\cdot)$ to denote the power generation cost in MG k and the power transmission cost between MG k and MG m, respectively. Without loss of generality, we assume the utility functions, the transmission cost functions, and the generation cost functions all have the same unit (e.g., dollar).

Jointly considering user utility, power transmission cost, and the load variance in the macrogrid, as well as power generation cost, power transmission cost in the MG network, we formulate the power scheduling problem Prob-MAMG as follows:

$$\max \sum_{t=1}^{T} \sum_{i\in\mathbb{N}} U(d_i(t), \omega_i(t)) - \frac{\alpha T}{2} \mathrm{Var}(\vec{l}_T) - \sum_{t=1}^{T} \left(\sum_{m\in\mathbb{M}} C_m(p_m(t)) + \sum_{k\in\mathbb{M}} G_k \left(\sum_{m\in\mathbb{M}} \frac{p_{km}(t)}{\sigma_{km}} - p_k(t) I(\sigma_k) \right) + \sum_{m\in\mathbb{M}} \sum_{k\in\mathbb{M}} C_{km}(p_{km}(t)) \right) \tag{1}$$

$$\text{s.t. } d_i(t) \geq d_{i,\min}(t) \quad \forall\, i \in \mathbb{N},\, t \in \mathbb{T} \tag{2}$$

$$G(l(t)) \leq b_{\max}(t) \quad \forall\, t \in \mathbb{T} \tag{3}$$

$$|p_m(t)| \leq p_{\mathrm{m},\max}(t) \quad \forall\, m \in \mathbb{M},\, t \in \mathbb{T} \tag{4}$$

$$\sum_{k\in\mathbb{M}} p_{km}(t) + p_m(t) R(\sigma_m) - s'_m(t) = d_m(t) \quad \forall\, m \in \mathbb{M},\, t \in \mathbb{T} \tag{5}$$

$$\sum_{m\in\mathbb{M}} \frac{p_{km}(t)}{\sigma_{km}} - p_k(t) I(\sigma_k) \leq g_{k,\max}(t) \quad \forall\, k \in \mathbb{M},\, t \in \mathbb{T} \tag{6}$$

where α is the weight to tradeoff the dual objectives, and the variance function $\mathrm{Var}(\cdot)$ is defined as

$$\mathrm{Var}(\vec{l}_T) = \frac{1}{T} \sum_{t=1}^{T} \left(l(t) - \frac{1}{T} \sum_{k=1}^{T} l(k) \right)^2 \tag{7}$$

where each element of vector $\vec{l}_T$ is the load of the macrogrid at time t computed as $l(t) = \sum_{i\in\mathbb{N}} d_i(t) + \sum_{m\in\mathbb{M}} p_m(t)$. The indicator function $I(\sigma_k)$ is defined as $I(\sigma_k) = 1/\sigma_k$ if $p_k(t) < 0$, and $I(\sigma_k) = 0$ otherwise; the indicator function $R(\sigma_m)$ is defined as $R(\sigma_m) = \sigma_m$ if $p_k(t) > 0$, and $R(\sigma_m) = 1/\sigma_m$ otherwise. Moreover, $d_{i,\min}(t)$ in constraint (2) is the minimum demand of user i, $b_{\max}(t)$ in constraint (3) is the generation cost limit for the energy provider, and $p_{m,\max}(t)$ in (4) is the maximum amount of transmission allowed in one time slot. In constraint (6), $g_{k,\max}(t)$ is the maximum generation in MG k at time t, and $s_m(t)$ is the power storage level of MG m at time t and $s'_m(t)$ is the amount of power to be stored in the time slot, computed as

$$s'_m(t) = s_m(t) - \xi \cdot s_m(t-1) \tag{8}$$

$$s_{m,\min} \leq s_m(t) \leq s_{m,\max} \quad \forall\, m \in \mathbb{M} \tag{9}$$

where ξ is the storage loss ratio in the ESS, and $s_{m,\min}$ ($s_{m,\max}$) is the lower (upper) bound on the storage capacity.

In Prob-MAMG, all the functions and the constraints are convex, which means it is a convex optimization problem. However, it cannot be solved unless all the constraints from the macrogrid and MGs are known a priori for the entire time period $\mathbb{T}$. Even with all these necessary information, it is very

[1] Note that although σ_{0m} can be same as σ_{m0}, and σ_{km} can be same as σ_{mk} in some cases. The reciprocity of transmission loss ratios is not an essential requirement in our model.

difficult to solve such a complex problem in practice. Note that in Prob-MAMG, the macrogrid and MG m are coupled by the power flow $p_m(t)$. As discussed, the macrogrid usually generates much more power than the MGs, and thus $p_m(t)$ can be seen as a special load in the macrogrid. Therefore, we can decompose the Prob-MAMG into two tiers. The first tier related to the macrogrid solves for power distribution for users, i.e., $d_i(t)$ and power exchanged with the MGs, i.e., $p_m(t)$. The second tier for the MGs matches $p_m(t)$ in MG m and solves for $p_{km}(t)$, i.e., the power transmissions among the cooperative MGs.

The first-tier problem for the macrogrid Prob-MA1 is formulated as follows:

$$\begin{aligned} \max \sum_{t=1}^{T} & \left(\sum_{i\in\mathbb{N}} U(d_i(t), \omega_i(t)) - \sum_{m\in\mathbb{M}} C_m(p_m(t)) \right) \\ & - \frac{\alpha T}{2} \mathrm{Var}(\vec{l}_T) \\ \text{s.t.} \ & (2) - (4). \end{aligned} \tag{10}$$

The second-tier problem for the MGs Prob-MG1 is as follows:

$$\begin{aligned} \min \sum_{k\in\mathbb{M}} & \left(G_k \left(\sum_{m\in\mathbb{M}} \frac{p_{km}(t)}{\sigma_{km}} - p_k^*(t) I(\sigma_k) \right) \right. \\ & \left. + \sum_{m\in\mathbb{M}} C_{km}(p_{km}(t)) \right) \\ \text{s.t.} \ & (5), (6), \ \text{and} \ (9) \end{aligned} \tag{11}$$

where $p_k^*(t)$ is part of the solution to Prob-MA1 (see Section III). Now, it is clear that in Prob-MA1, user utility, load variance, and power transmission cost to/from the macrogrid are optimized, and Prob-MG1 aims to minimize the power generation cost and transmission cost among the cooperative MGs. In Section III, we reformulate Prob-MA1 and develop an online algorithm that is asymptotically optimal. In Section IV, we solve problem Prob-MG1 with a distributed algorithm for optimal solutions.

III. Online Power Distribution in the Macrogrid

A. Reformulation and Optimal Offline Solution

In Prob-MA1, all the power users and MGs are independent. Thus, we can reformulate Prob-MA1 by replacing the grid load variance term with $\mathrm{Var}(\vec{l}_T) = \sum_{i\in\mathbb{N}} \mathrm{Var}(\vec{d}_{i,T}) + \sum_{m\in\mathbb{M}} \mathrm{Var}(\vec{p}_{m,T})$. We thus obtain Prob-MA2 as follows:

$$\begin{aligned} \max\colon \mathbf{F}(\mathbf{d},\mathbf{p}) = & \sum_{t=1}^{T} \left(\sum_{i\in\mathbb{N}} U(d_i(t), \omega_i(t)) - \sum_{m\in\mathbb{M}} C_m(p_m(t)) \right) \\ & - \frac{\alpha T}{2} \left(\sum_{i\in\mathbb{N}} \mathrm{Var}(\vec{d}_{i,T}) + \sum_{m\in\mathbb{M}} \mathrm{Var}(\vec{p}_{m,T}) \right) \\ \text{s.t.} \ & (2) - (4) \end{aligned} \tag{12}$$

where $\mathrm{Var}(\vec{d}_{i,T}) = \frac{1}{T}\sum_{t=1}^{T}(d_i(t) - \frac{1}{T}\sum_{k=1}^{T} d_i(k))^2$ and $\mathrm{Var}(\vec{p}_{m,T}) = \frac{1}{T}\sum_{t=1}^{T}(p_m(t) - \frac{1}{T}\sum_{k=1}^{T} p_m(k))^2$.

In Prob-MA2, the utility function $U(\cdot)$ is concave, and both the transmission cost function $C_m(\cdot)$ and the variance function $\mathrm{Var}(\cdot)$ are convex. Therefore, Prob-MA2 is a convex optimization problem with a convex set of the constraints. Furthermore, Prob-MA2 has a unique solution since $U(\cdot)$ is strictly increasing. Thus, we can select the constraints $d_{i,\min}(t)$ and $p_{m,\max}(t)$, so that Prob-MA2 is feasible and the Slater's condition is satisfied, and obtain the optimal solution by solving the KKT conditions [25] as

$$\begin{cases} I_{\mathbb{N}} \left(U'(\tilde{d}_i(t), \omega_i(t)) - \alpha(\tilde{d}_i(t) - \bar{\tilde{d}}_{i,T}) + \tilde{\nu}_i(t) \right) \\ \quad + I_{\mathbb{M}} \left(-C'_m(\tilde{p}_m(t)) - \alpha(\tilde{p}_m(t) - \bar{\tilde{p}}_{m,T}) + \tilde{\gamma}_m(t) \right. \\ \quad \left. - \tilde{\rho}_m(t) \right) - \tilde{\mu}(t) G'(\tilde{l}(t))/b_{\max}(t) = 0 \\ \tilde{\mu}(t) \left(G(\tilde{l}(t))/b_{\max}(t) - 1 \right) = 0 \\ \tilde{\nu}_i(t) \left(\tilde{d}_i(t) - d_{i,\min}(t) \right) = 0 \\ \tilde{\gamma}_m(t) \left(\tilde{p}_m(t) + p_{m,\max}(t) \right) = 0 \\ \tilde{\rho}_m(t) \left(\tilde{p}_m(t) - p_{m,\max}(t) \right) = 0 \\ \tilde{\mu}(t), \tilde{\nu}_i(t), \tilde{\gamma}_m(t), \tilde{\rho}_m(t) \geq 0 \quad \forall\, i \in \mathbb{N},\ m \in \mathbb{M},\ t \in \mathbb{T} \end{cases} \tag{13}$$

where $\tilde{d}_i(t)$ and $\tilde{p}_m(t)$ are the optimal points; the indicator $I_{\mathbb{N}} = 1$ for the users in $\mathbb{N}$, and $I_{\mathbb{N}} = 0$ otherwise; the indicator $I_{\mathbb{M}} = 1$ for the MGs in $\mathbb{M}$, and $I_{\mathbb{M}} = 0$ otherwise; and

$$\bar{\tilde{d}}_{i,T} = \frac{1}{T}\sum_{k=1}^{T} \tilde{d}_i(k), \quad \bar{\tilde{p}}_{m,T} = \frac{1}{T}\sum_{k=1}^{T} \tilde{p}_m(k). \tag{14}$$

From the gradient condition of the above KKT conditions, we derive the Lagrange multiplier $\mu(t)$ as

$$\begin{aligned} \tilde{\mu}(t) = & \left(I_{\mathbb{N}} \left(U'(\tilde{d}_i(t), \omega_i(t)) - \alpha(\tilde{d}_i(t) - \bar{\tilde{d}}_{i,T}) + \tilde{\nu}_i(t) \right) \right. \\ & \left. + I_{\mathbb{M}} \left(-C'_m(\tilde{p}_m(t)) - \alpha(\tilde{p}_m(t) - \bar{\tilde{p}}_{m,T}) + \tilde{\gamma}_m(t) - \tilde{\rho}_m(t) \right) \right) \\ & / (G'(\tilde{l}(t))/b_{\max}(t)). \end{aligned} \tag{15}$$

The optimal solution to Prob-MA2 can be found by solving its KKT conditions (13)–(15). However, solving the KKT conditions requires the information on $\bar{\tilde{d}}_{i,T}$ and $\bar{\tilde{p}}_{m,T}$, which are the average of the user demand $\tilde{d}_i(t)$ and the exchanged power with MG m $\tilde{p}_m(t)$ for the entire time period T, respectively. To derive the optimal solution to Prob-MA2, the constraints $d_{i,\min}(t)$, $p_{m,\max}(t)$, and $b_{\max}(t)$ over the entire time window $\mathbb{T}$ are also needed. This is an *offline optimal solution*, which may not be practical in some cases.

B. Online Power Distribution in the Macrogrid

We next present an online algorithm for Prob-MA2 in this section. It can be seen that in (15), $\bar{\tilde{d}}_{i,T}$ and $\bar{\tilde{p}}_{m,T}$ are the only two terms requiring future information [24], while these time averages can be approximated by properly defined updating equations. Motivated by this observation, we first present an approximation problem that can be solved without future information, and then prove that its solution is convergent to the optimal offline solution to the original problem Prob-MA2.

Specifically, we replace the average terms $\frac{1}{T}\sum_{k=1}^{T} d_i(k)$ and $\frac{1}{T}\sum_{k=1}^{T} p_m(k)$ in (12) by two new terms $\hat{d}_i(t)$ and $\hat{p}_m(t)$, respectively, and remove the time sum notation, so that the problem can be solved at time t. We thus obtain a new problem Prob-MA3 at time t as follows:

$$\max \sum_{i\in\mathbb{N}} U(d_i(t),\omega_i(t)) - \sum_{m\in\mathbb{M}} C_m(p_m(t)) - \frac{\alpha}{2}\sum_{i\in\mathbb{N}}(d_i(t) - \hat{d}_i(t-1))^2 - \frac{\alpha}{2}\sum_{m\in\mathbb{M}}(p_m(t)-\hat{p}_m(t-1))^2$$
$$\text{s.t. } (2)-(4) \tag{16}$$

where $\hat{d}_i(t)$ and $\hat{p}_m(t)$ are updated at each time slot t as

$$\begin{cases}\hat{d}_i(t) = \hat{d}_i(t-1) + \frac{\alpha}{t+\alpha}\cdot(d_i^*(t)-\hat{d}_i(t-1))\\ \hat{p}_m(t) = \hat{p}_m(t-1) + \frac{\alpha}{t+\alpha}\cdot(p_m^*(t)-\hat{p}_m(t-1))\end{cases} \tag{17}$$

where $d_i^*(t)$ and $p_m^*(t)$ denote the solutions to Prob-MA3. This way, we decompose the problem over a time window $\mathbb{T}$ into many problems to be solved by the MC at each time t without requiring any future information. Because the updating equations in (17) only use the solutions to Prob-MA3 in the previous time slot, we use $\hat{d}_i(t)$ and $\hat{p}_m(t)$ to approximate the average terms $\bar{d}_{i,T}^*$ and $\bar{p}_{m,T}^*$, respectively. The following lemma and theorem state that $\hat{d}_i(t)$ and $\hat{p}_m(t)$ are convergent, and the online solutions are convergent to the offline solutions. The complete proof is presented in the Appendix.

Lemma 1: The updating terms in Prob-MA3, i.e., $\hat{d}_i(t)$ and $\hat{p}_m(t)$, are convergent to the time averages of its solution $\bar{d}_{i,T}^*$ and $\bar{p}_{m,T}^*$, respectively, when T is sufficiently large. That is, for $i\in\mathbb{N}$ and $m\in\mathbb{M}$, we have

$$\lim_{T\to\infty}\hat{d}_i(T) = \lim_{T\to\infty}\bar{d}_{i,T}^* = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} d_i^*(t) \tag{18}$$

$$\lim_{T\to\infty}\hat{p}_m(T) = \lim_{T\to\infty}\bar{p}_{m,T}^* = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} p_m^*(t). \tag{19}$$

Theorem 1: The solution to Prob-MA3 converges asymptotically to the solution to Prob-MA2.

According to Theorem 1, we can solve Prob-MA3 in each time t using the information of the current time, while still achieving the optimal results in a certain amount of time slots.

IV. Distributed Cooperative Power Scheduling for MGs

In Section II-B, we formulate problem Prob-MG1, which can also be shown to be convex. We can solve Prob-MG1 with some convex optimization techniques [25], such as KKT conditions as in solving Prob-MA2 and Prob-MA3. In practical scenarios, a distributed algorithm is more appealing for reducing the computational complexity, reducing delay in realtime power scheduling, and enhancing scalability. Please see our previous work [26] for more discussions on the benefits of using distributed algorithms in the SG. In this section, we develop a distributed cooperative power scheduling algorithm for the MGs, by decomposing Prob-MG1 into multiple subproblems to be solved by the MGCC in each MG.

A. Problem Reformulation

The constraints on the power storage levels $s_m(t)$, i.e., (8) and (9), can be merged as

$$\frac{s_{m,\min}}{(1-\xi)} - s_{m,\max} \le s'_m(t) \le s_{m,\max} - s_{m,\min}. \tag{20}$$

Substituting (20) into (5), we have a new constraint for $\sum_{k\in\mathbb{M}} p_{km}(t)$. Then, Prob-MG1 (11) can be rewritten with the new constraint as

$$\min: \sum_{k\in\mathbb{M}}\left(G_k\left(\sum_{m\in\mathbb{M}}\frac{p_{km}(t)}{\sigma_{km}} - p_k^*(t)I(\sigma_k)\right) + \sum_{m\in\mathbb{M}} C_{km}(p_{km}(t))\right) \tag{21}$$

$$\text{s.t. } C_{m,\min}(t) \le \sum_{k\in\mathbb{M}} p_{km}(t) \le C_{m,\max}(t) \quad \forall\, m\in\mathbb{M},\, t\in\mathbb{T} \tag{22}$$

$$\sum_{m\in\mathbb{M}}\frac{p_{km}(t)}{\sigma_{km}} - p_k^*(t)I(\sigma_k) \le g_{k,\max}(t) \quad \forall k\in\mathbb{M},\, t\in\mathbb{T} \tag{23}$$

where

$$\begin{cases} C_{m,\min}(t) &= \frac{s_{m,\min}}{1-\xi} - s_{m,\max} + d_m(t) - p_m^*(t)R(\sigma_m)\\ C_{m,\max}(t) &= s_{m,\max} - s_{m,\min} + d_m(t) - p_m^*(t)R(\sigma_m).\end{cases} \tag{24}$$

In Prob-MG1, the variables are $p_{km}(t)$ for each MG pair k and m. We next decompose Prob-MG1 into subproblems using only local information with the *dual decomposition* technique [27].

B. Cooperative Distributed Power Scheduling for MGs

We first derive the Lagrangian of Prob-MG1 as follows:

$$\begin{aligned} L(\mathbf{P}(t), \vec{\lambda}_M(t), \vec{\beta}_M(t)) &= \sum_{k\in\mathbb{M}}\left(G_k\left(\sum_{m\in\mathbb{M}}\frac{p_{km}(t)}{\sigma_{km}} - p_k^*(t)I(\sigma_k)\right) + \sum_{m\in\mathbb{M}} C_{km}(p_{km}(t))\right)\\ &\quad + \sum_{m\in\mathbb{M}}\lambda_m(t)\left(\sum_{k\in\mathbb{M}} p_{km}(t) - C_{m,\max}(t)\right)\\ &\quad + \sum_{m\in\mathbb{M}}\beta_m(t)\left(C_{m,\min}(t) - \sum_{k\in\mathbb{M}} p_{km}(t)\right)\\ &= \sum_{k\in\mathbb{M}}\left(G_k\left(\sum_{m\in\mathbb{M}}\frac{p_{km}(t)}{\sigma_{km}} - p_k^*(t)I(\sigma_k)\right) + \sum_{m\in\mathbb{M}}(C_{km}(p_{km}(t)) + (\lambda_m(t)-\beta_m(t))p_{km}(t))\right)\\ &\quad + \sum_{m\in\mathbb{M}}(\beta_m(t)C_{m,\min}(t) - \lambda_m(t)C_{m,\max}(t)) \end{aligned} \tag{25}$$

where $\beta_m(t) \geq 0$ and $\lambda_m(t) \geq 0$ are the Lagrange multipliers associated with the two inequalities in constraint (22), respectively. We then decompose Prob-MG1 into M subproblems $S(\vec{\lambda}_M(t), \vec{\beta}_M(t))$ for the MGs, termed as Prob-MG2

$$\begin{aligned}\min\ & S_k(\vec{\lambda}_M(t), \vec{\beta}_M(t)) \\ &= G_k\left(\sum_{m\in\mathbb{M}} \frac{p_{km}(t)}{\sigma_{km}} - p_k^*(t) I(\sigma_k)\right) \\ &\quad + \sum_{m\in\mathbb{M}} \left(C_{km}(p_{km}(t)) + (\lambda_m(t) - \beta_m(t)) p_{km}(t)\right) \\ \text{s.t. } & (23). \end{aligned} \tag{26}$$

The dual problem of Prob-MG1 is as follows [27]:

$$\max D(\vec{\lambda}_M(t), \vec{\beta}_M(t)) \tag{27}$$

$$\text{s.t. } \lambda_m(t) \geq 0, \quad \beta_m(t) \geq 0 \quad \forall\, m \in \mathbb{M} \tag{28}$$

where

$$\begin{aligned} D(\vec{\lambda}_M(t), \vec{\beta}_M(t)) &= \min \left\{ L(\mathbf{P}(t), \vec{\lambda}_M(t), \vec{\beta}_M(t)) \right\} \\ &= \sum_{k\in\mathbb{M}} S_k(\lambda_M(t), \beta_M(t)) \\ &\quad + \sum_{m\in\mathbb{M}} (\beta_m(t) C_{m,\min}(t) \\ &\quad - \lambda_m(t) C_{m,\max}(t)). \end{aligned} \tag{29}$$

We thus decompose Prob-MG1 into M subproblems each of which can be solved by the MGCC in each MG. Furthermore, because the primal problem (21) is convex and has feasible solutions for proper selections of $s_{m,\min}$, $s_{m,\max}$, and $g_{k,\max}(t)$, *strong duality* holds [27], so that the optimal solution can be obtained from the dual problem (29).

For given $\vec{\lambda}_M(t)$ and $\vec{\beta}_M(t)$, the subproblem $S_k(\vec{\lambda}_M(t), \vec{\beta}_M(t))$ for MG k is convex because the generation cost function $G_k(\cdot)$, the transmission cost function $C_{km}(\cdot)$, and constraint (23) are all convex as discussed in Section II. The subproblems can be solved by commonly used methods such as KKT conditions and the interior point method (IPM) [25]. After solving the subproblems, the dual problem will be solved by the MGNC by gathering all the solutions to the subproblems from the MGCCs. Furthermore, function $S_k(\cdot)$ is differentiable because both $G(\cdot)$ and $C(\cdot)$ are differentiable. We can use the following *gradient method* to obtain the dual variables $\lambda_m(t)$ and $\beta_m(t)$

$$\begin{cases} \lambda_{m,t}(j+1) = \\ \quad \left[\lambda_{m,t}(j) - \delta\left(C_{m,\max}(t) - \sum_{k\in\mathbb{M}} p^*_{km,t}(j)\right)\right]^+ \\ \beta_{m,t}(j+1) = \\ \quad \left[\beta_{m,t}(j) + \tau\left(C_{m,\min}(t) - \sum_{k\in\mathbb{M}} p^*_{km,t}(j)\right)\right]^+ \end{cases} \tag{30}$$

where δ and τ are step sizes; $p^*_{km,t}(j)$ is the solution to Prob-MG2 (26) for given $\lambda_{m,t}(j)$ and $\beta_{m,t}(j)$; and $[\cdot]^+$ is the projection onto the nonnegative orthant [27]. The dual variable $\vec{\lambda}_M(t)$ and $\vec{\beta}_M(t)$ will converge to the dual optimal $\vec{\lambda}^*_M(t)$ and $\vec{\beta}^*_M(t)$, respectively, since *strong duality* holds [27]. The optimal solution $\mathbf{P}^*(t)$ to Prob-MG1 can be acquired by solving each Prob-MG2 for $\vec{\lambda}^*_M(t)$ and $\vec{\beta}^*_M(t)$.

V. Optimal Hierarchical Power Scheduling for the Entire System

In this section, we summarize the analyses in Sections III and IV, and present the hierarchical power scheduling algorithms, termed HPS, for the entire power grid system. As discussed, HPS consists of two tiers: 1) online power distribution in the macrogrid; and 2) cooperative distributed power scheduling in the MGs. Furthermore, the lower tier algorithm consists of two parts: one for the MGNC and the other for each MGCC. The proposed algorithms are presented in Algorithms 1–3, where $\epsilon_m > 0$ and $\varepsilon_m > 0$ are small tolerance values for termination conditions, for all $m \in \mathbb{M}$.

Algorithm 1. Online Power Distribution in the Macrogrid

```
Initialize d̂_i(0) and p̂_m(0), for all i ∈ ℕ and m ∈ 𝕄
For i = 1 : T do
    Receive constraint p_{m,max}(t) from the MGNC;
    Solve Prob-MA3 (16);
    Send solution p*_m(t) to the MGNC;
    Update d̂_i(t) and p̂_m(t) for all i ∈ ℕ and m ∈ 𝕄 as in (17);
    Exchange power p*_m(t) with MG m, for all m;
    Distribute power d*_i(t) to Macrogrid user i, for all i;
end
```

Algorithm 2. Distributed Cooperative Power Scheduling Algorithm for the MGNC

```
for i = 1 : T do
    Receive p_{m,max}(t) from MG m and forward it to the MC, for all m ∈ 𝕄;
    Receive p*_m(t) from the MC and forward it to MG m, for all m ∈ 𝕄;
    Initialize λ⃗_{M,t}(0) ≥ 0 and β⃗_{M,t}(0) ≥ 0, and broadcast them to all the MGs;
    repeat
        Receive p*_{km,t}(j) from the MGs;
        Update λ⃗_{M,t}(j) and β⃗_{M,t}(j) using (30);
        Broadcast them to all the MGs, for all k, m ∈ 𝕄;
    until(|λ_{m,t}(j+1) − λ_{m,t}(j)| < ϵ_m and |β_{m,t}(j+1) − β_{m,t}(j)| < ε_m);
    Broadcast λ⃗*_M(t) and β⃗*_M(t) to all the MGs;
end
```

Note that the MC in the macrogrid requires information on $p_{m,\max}(t)$ to solve Prob-MA3. Each MG m can estimate $p_{m,\max}(t)$ according to (31), where the first term refers to the maximum possible amount of power transmitted to the macrogrid from MG m, and the second term is the maximum possible amount of power that MG m can accept from the macrogrid.

Algorithm 3. Distributed Cooperative Power Scheduling Algorithm for Each MGCC

for $i = 1 : T$ **do**
- Estimate the maximum exchanged power with the Macrogrid $p_{m,\max}(t)$ and report it to the MGNC;
- Receive $p_m^*(t)$ and calculate constraints $C_{m,\min}(t)$ and $C_{m,\max}(t)$ using (24);
- **repeat**
  - Receive $\vec{\lambda}_{M,t}(j)$ and $\vec{\beta}_{M,t}(j)$ and solve Prob-MG2 (26);
  - Send solution $p_{km,t}^*(j)$ to the MGNC, for all $k, m \in \mathbb{M}$;
- **until** ($\vec{\lambda}_M^*(t)$ *and* $\vec{\beta}_M^*(t)$ *are received*);
- Calculate $s_m'(t)$ using (5) and charge or discharge the ESS accordingly, for all $m \in \mathbb{M}$;
- Transmit power $p_{km}^*(t)$ to MG m and exchange $p_m^*(t)$ with the Macrogrid, for all $k, m \in \mathbb{M}$;

end

The estimate of $p_{m,\max}(t)$ is based on the power dispatching information of the last time slot. This works well for a short operation cycle, e.g., 15 min, because between two adjacent short time cycles, major grid-related parameters such as generation and demand are usually closely correlated, while 15-min cycles are sufficient for power scheduling in a large macrogrid with several MGs under current technology of information processing and communications.

$$\min\Bigg\{|g_{m,\max}(t) + \sum_{k\in\mathbb{M}} p_{km}^*(t-1) + (1-\xi)s_m(t-1) - \frac{s_{m,\min}}{1-\xi} - d_m(t)|, |-\sum_{k\in\mathbb{M}} p_{mk}^*(t-1) - (s_{m,\max} - (1-\xi)s_m(t-1)) - d_m(t)|\Bigg\}. \quad (31)$$

The complexity of Algorithm 1 is related to the number of users and MGs, and the number of calculations solving the KKT equations of Prob-MA3 (16). According to [25], the complexity of Algorithm 1 is roughly $\mathcal{O}((N+M)^3)$. The complexity of Algorithms 2 and 3 is related to the product of the number of iterations of the dual variables and the number of calculations solving Prob-MG2 (26). And the complexity of Algorithms 2 and 3 is about $\mathcal{O}((2M)^3 \cdot M^3) = \mathcal{O}(M^6)$. The complexity analysis is quite conservative, and thus, the complexity of the proposed algorithm is polynomial of the number of macrogrid users and MGs, which can be processed easily within a 15-min cycle, based on the processing ability of current microcomputers.

The proposed cooperative distributed algorithm is also well suited for larger power systems due to the scalability. It is also worth noting that there is no information exchange directly between the MC and MGCCs. The MGNC connects the MC and MGCCs in the system, so that there is only one information connection point between the macrogrid and the MGNC, which increases the level of security and privacy protection.

VI. Performance Evaluation

In this section, we present a trace-driven simulation study to evaluate the efficacy of the proposed HPS scheme. The simulation data and parameters for the macrogrid are based on the power usage traces in the Southern California Edison (SCE) area recorded in 2011 [28]. The data for MGs are based on some statistical distributions, which are averaged over a large number of random runs.

We consider a power system as in Fig. 1 with a macrogrid and four MGs. The macrogrid supports 400 power users, while each MG supports 100 users. The user demand is based on the SCE trace and user utility function is defined as [23]

$$U(p_i(t), \omega_i(t)) = \begin{cases} \omega_i(t) \cdot p_i(t) - p_i(t)^2/8, & \text{if } 0 \le p_i(t) \le 4\omega_i(t) \\ 4 \cdot \omega_i(t), & \text{if } p_i(t) \ge 4\omega_i(t). \end{cases}$$

As in [29], the power generated in each MG is independently chosen from a uniform distribution in (10 kW, 450 kW). The generation cost function in the macrogrid is assumed to be $G(x) = 4x^2$ and the MG generation cost function is $G_k(x) = 1.5x^2$. The transmission cost between the macrogrid and MG m is assumed to be $C_m(x) = \theta_m x^2$; the transmission cost between MG k, and m is $C_{km}(x) = \frac{\theta_{km}}{2}x^2$, for $k \neq m$ and $C_{mm}(x) = 0$, for all m [17], [19]. θ_m and θ_{km} are transmission cost coefficients, which are defined as $\theta_m = 1 - \sigma_m$, if $x > 0$, and $\theta_m = \frac{1-\sigma_m}{\sigma_m}$ otherwise; $\theta_{km} = \frac{1-\sigma_{km}}{\sigma_{km}}$.

In practice, the transmission loss ratios σ_m and σ_{km} usually differ from MG to MG because different factors such as distances. For practical considerations, we assume that $\{\sigma_{01}, \sigma_{02}, \sigma_{03}, \sigma_{04}\} = \{0.2, 0.6, 0.5, 0.3\}$, $\{\sigma_{10}, \sigma_{20}, \sigma_{30}, \sigma_{40}\} = \{0.5, 0.4, 0.67, 0.59\}$, $\sigma_{12} = 0.56$, $\sigma_{13} = 0.71$, $\sigma_{14} = 0.67$, $\sigma_{21} = 0.56$, $\sigma_{23} = 0.56$, $\sigma_{24} = 0.63$, $\sigma_{31} = 0.71$, $\sigma_{32} = 0.56$, $\sigma_{34} = 0.56$, $\sigma_{41} = 0.67$, $\sigma_{42} = 0.63$, $\sigma_{43} = 0.56$, and $\sigma_{mm} = 1$, for all m. The corresponding transmission cost coefficients are $\theta_{12} = 0.8$, $\theta_{13} = 0.4$, $\theta_{14} = 0.5$, $\theta_{21} = 0.8$, $\theta_{23} = 0.8$, $\theta_{24} = 0.6$, $\theta_{31} = 0.4$, $\theta_{32} = 0.8$, $\theta_{34} = 0.8$, $\theta_{41} = 0.5$, $\theta_{42} = 0.6$, $\theta_{43} = 0.8$, and $\theta_{mm} = 1$, for all m. We also assume different storage capacities to the MGs as $(s_{1,\min}, s_{1,\max}) = (30, 300)$ kW, $(s_{2,\min}, s_{2,\max}) = (50, 100)$ kW, $(s_{3,\min}, s_{3,\max}) = (60, 500)$ kW, and $(s_{4,\min}, s_{4,\max}) = (100, 200)$ kW. The algorithms are executed on 15-min time slots.

A. HPS Performance

The HPS algorithms contain two iterative sequences: 1) $d_i(t)$ and $p_m(t)$ in the macrogrid control; and 2) $p_{km,t}(j)$, $\vec{\lambda}_{M,t}(j)$ and $\vec{\beta}_{M,t}(j)$ in the MG control. In the first tier, the convergence of $d_i(t)$ and $p_m(t)$ is over multiple time slots; whereas in the second tier, the convergence is achieved within every time slot. Because the second-tier control requires the solution $p_m^*(t)$ from the first-tier control as a constraint, the convergence of $p_m(t)$ is critical for HPS performance. The parameter α in the updating equations (17) should be carefully selected. For the second tier, the MGNC needs to exchange $\vec{\lambda}_{M,t}(j)$ and $\vec{\beta}_{M,t}(j)$ for several iterations until they converge. For given termination

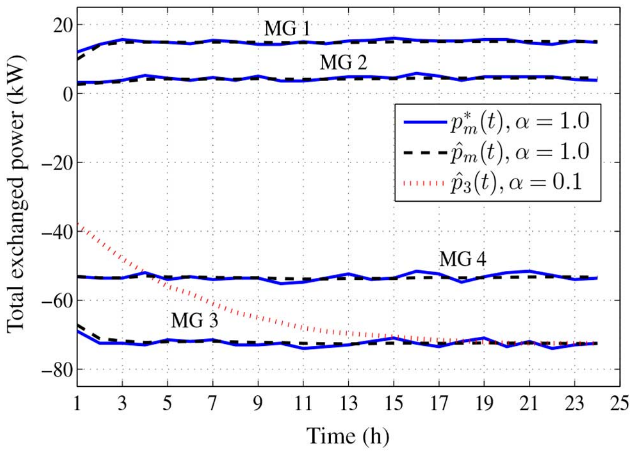

Fig. 2. Convergence of $\hat{p}_m(t)$.

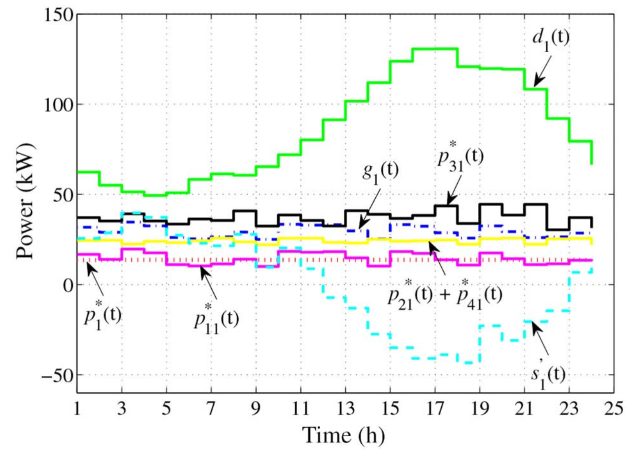

Fig. 4. Power scheduling in MG 1.

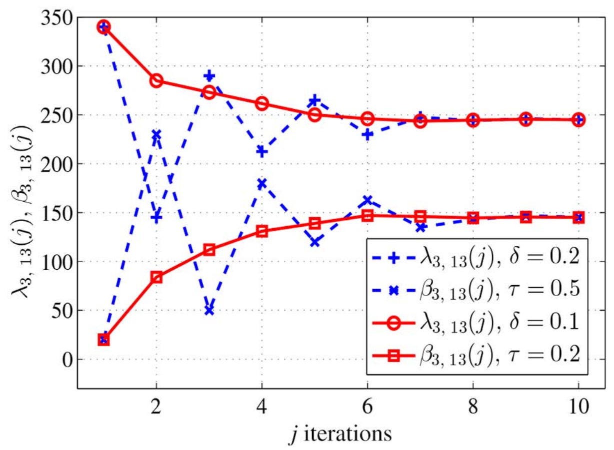

Fig. 3. Convergence of $\lambda_{3,13}(j)$ and $\beta_{3,13}(j)$ with different δ and τ.

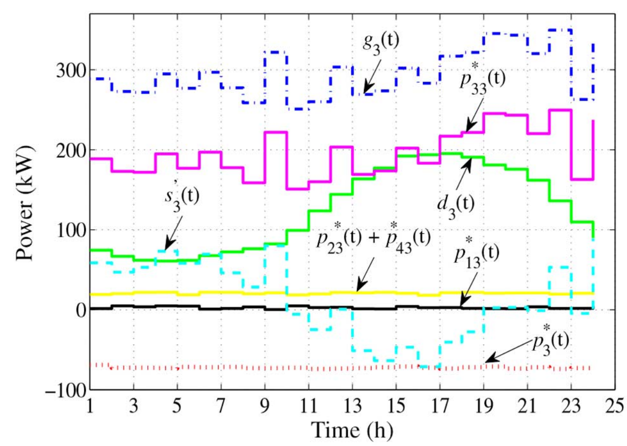

Fig. 5. Power scheduling in MG 3.

conditions, the step-size parameters δ and τ will affect the speed of convergence of $\vec{\lambda}_M(t)$ and $\vec{\beta}_M(t)$. Intuitively, small δ and τ guarantee the convergence, but may require more iterations. We illustrate the effect of the parameters on convergence in Figs. 2 and 3.

Fig. 2 shows that $\hat{p}_m(t)$ converges in every MG, and $p_m^*(t)$ fluctuates around $\hat{p}_m(t)$, as specified in Lemma 1. From the two sequences of $\hat{p}_3(t)$ with $\alpha = 1$ and $\alpha = 0.1$, it can be seen that a smaller α results in slower convergence. However, a large α may lead to larger variance in the transient phase. From a larger number of simulation runs, we set $\alpha = 1$ in our simulations. It is also worth noting that the MGs have different levels of $p_m^*(t)$. This is because different MGs have different generation levels, storage limits, and transmission cost coefficients. For example, $p_1^*(t)$ has a positive level of 13.5 kW, which means MG 1 requires a 13.5-kW load from the macrogrid. This may due to low generation, small storage, or large transmission cost with the macrogrid. On the other hand, $p_3^*(t)$ has a negative level of -70 kW, which means that MG 3 transfers 70 kW to the macrogrid. Furthermore, the sum of $p_m^*(t)$'s is negative, meaning that the macrogrid acquires power from the MG network in this time frame.

In Fig. 3, the evolutions of $\lambda_{3,13}(j)$ and $\beta_{3,13}(j)$ in a time slot are plotted with different step sizes δ and τ. The curves with larger step sizes have larger variances and slightly slower convergence speed. Both $\lambda_{3,13}(j)$ and $\beta_{3,13}(j)$ have a very fast convergence in six to eight iterations, which also indicates a fast convergence of $p_{3m}^*(t)$ as stated in Section IV-B. The fast convergence is due to the transformed constraints (24) of Prob-MG1, which further restricts the set of feasible $p_{km}^*(t)$'s. As a result, $\lambda_m(t)$ and $\beta_m(t)$ are forced to increase or decrease in a reverse direction, which reduces the number of iterations needed for convergence.

In addition to convergence performance, HPS can be further evaluated with respect to power scheduling. In Figs. 4 and 5, we present the power flows in MGs 1 and 3, respectively. We find in Fig. 4 that MG 1 has a very low power generation $g_1(t)$ from about 20 to 40 kW, such that it cannot support the power demand in the range of 50–130 kW with its own generation alone. However, it only requests less than 20 kW from the macrogrid, but accepts more than 40 kW from MG 3 and 20 kW from MGs 2 and 4. Note that the transmission cost coefficient between the macrogrid and MG 1 is $\theta_1 = 1 - \sigma_{01} = 0.8$, which is larger than the coefficients between MG 1 and the other MGs. To find a sufficient power flow to support its users, while keeping the cost low, MG 1 chooses to request more power from MG 3. MG 3 has a large generation and storage capacity, so that it can share much power with MG 1 and the macrogrid. However,

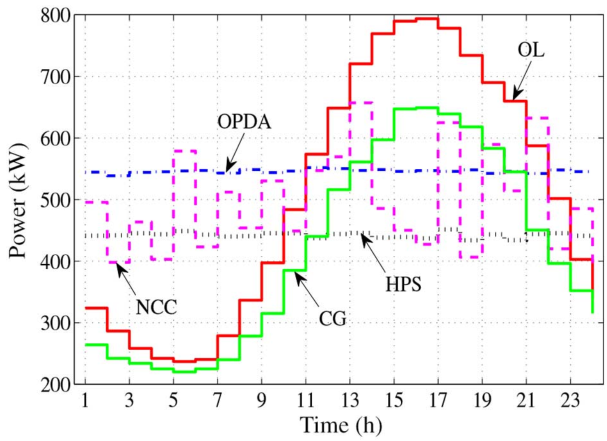

Fig. 6. Macrogrid load under different power scheduling schemes.

we observe a relatively low power flow around 20 kW between MG 3 and MG 2 and 4. This can also be explained by the objective to minimize the transmission cost. Actually, both $\frac{\theta_{32}}{2}$ and $\frac{\theta_{34}}{2}$ are 0.4, which is very close to $\theta_3 = \frac{1-\sigma_{30}}{\sigma_{30}} = 0.5$.

In real power systems, the transmission cost ratio σ and coefficient θ can be affected by many different factors, such as distance, the power gap, and the complexity of the system. In a real system, MG 1 may have shorter distance and smaller gap of power level with MG 3 compared to that with MG 2 and 4. The storage also plays an important role as a power buffer to enhance system stability and capacity. As a result, with all the key factors considered, HPS is able to achieve a balance in the power system, maximize the macrogrid user utilities, smooth the load of the macrogrid, and minimize the cost in the cooperative MGs.

B. Comparison With Existing Schemes

We next provide a comparison study of HPS versus several existing schemes [17], [23]. In Fig. 6, we show the load of the macrogrid under five different power scheduling schemes. The original load (OL) is based on the SCE trace of a 1-day period in September, 2011. The online power distribution algorithm (OPDA) proposed in [23] has considered many factors including user utility and grid load variance in a macrogrid, but no MG is involved in the model and algorithm. Thus, the OPDA curve in Fig. 6 is obtained by running OPDA in a macrogrid with 800 users. A coalition game (CG) is used in [17] to minimize the power loss in an MG network, where power flows between MGs and macrogrid are allowed. However, it does not consider smoothing the macrogrid load. For comparison purpose, we also develop another scheme [termed as no cooperation control scheme (NCC)], which only allows power flow between each MG and the macrogrid, while power trading among the MGs is not allowed.

In Fig. 6, the OL curve has the largest peak. OPDA achieves an expected smooth load in the macrogrid, because the variance term is explicitly minimized. OPDA also distributes power well such that the user demands can be satisfied. However, without the cooperative MGs, the macrogrid generates 7.8% more power in total to achieve 31.7% of peak reduction. With NCC, the macrogrid can exchange power with the MGs, which usually have random power generations. As a result, power generation in the macrogrid under NCC is almost the same as that under OL. NCC achieves 17.2% peak reduction while causing 21.4% increase in the macrogrid grid load variance. For CG, coalitions are formed among the MGs to minimize power loss. We find a 15.7% of generation reduction, 18.4% of peak reduction, and 40.9% of variance reduction. It achieves a fairly good result by exploiting the DG from the cooperative MGs. However, it does not explicitly consider variance reduction; the resulting variance is actually still large.

With HPS, we jointly consider all the above factors. As a result, under HPS, the macrogrid has a 97.1% variance reduction, 43.1% of peak reduction, and 13.1% of generation reduction. Compared to CG, HPS achieves considerably better results on peak reduction and variance reduction, and the macrogrid generation reduction is only slightly lower (i.e., 13.1% versus 15.7% with CG).

VII. Conclusion

In this paper, we developed a hierarchical power scheduling scheme to optimally manage the power distribution in the smart grid with one macrogrid and cooperative MGs. We first presented a formulation considering both the macrogrid, which jointly considers user utility, generation cost, transmission cost, and grid load smoothing, and the MGs, which aims to minimize the cost of power generation and transmission within the MGs. We then decompose the problem into a two-tier formulation and developed the corresponding online and distributed algorithms for solving both problems, which were proven to be asymptotically optimal. The proposed algorithms were validated with trace-driven simulations.

Appendix

A. Proof of Lemma 1

Proof: We prove (18) in this section, while (19) can be proved in the same way. Rewrite the first update equation in (17) and sum up from $t = 1$ to T. We have

$$\sum_{t=1}^{T}\left(\frac{t+\alpha}{\alpha}\right)(\hat{d}_i(t) - \hat{d}_i(t-1)) = \sum_{t=1}^{T}(d_i^*(t) - \hat{d}_i(t-1)).$$

Expanding the sum on the left-hand side (LHS), some terms can be canceled. Also, adding term $-\hat{d}_i(T) + \hat{d}_i(T)$ to the right-hand side (RHS), it follows that

$$\frac{1}{\alpha}\left(T\cdot\hat{d}_i(T) - \sum_{t=1}^{T}\hat{d}_i(t-1)\right) - (\hat{d}_i(T) - \hat{d}_i(1))$$
$$= \sum_{t=1}^{T}(d_i^*(t) - \hat{d}_i(T) + \hat{d}_i(T) - \hat{d}_i(t-1)).$$

Taking limit over T on both sides, we have

$$\lim_{T\to\infty}\frac{T\cdot\hat{d}_i(T) - \sum_{t=1}^{T}\hat{d}_i(t-1)}{\alpha\cdot T} - \lim_{T\to\infty}\frac{\hat{d}_i(T) - \hat{d}_i(1)}{T}$$
$$= \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}(d_i^*(t) - \hat{d}_i(T) + \hat{d}_i(T) - d_i^*(t-1)).$$

The second term of the LHS goes to zero. Rearranging the remaining terms, we have

$$\lim_{T\to\infty}\left(\frac{1-\alpha}{\alpha}\right)\left(\hat{d}_i(T)-\frac{1}{T}\sum_{t=1}^{T}\hat{d}_i(t-1)\right) = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}(d_i^*(t)-\hat{d}_i(T)). \quad (32)$$

Due to the first updating function in (17), $\hat{d}_i(t)$ is convergent as $t\to\infty$. Thus, the LHS of (32) is zero. It follows (32) that (18) holds true. ■

B. Proof of Theorem 1

Proof: The convergence of the online solution is equivalent to the convergence of the online objective value to that of the offline problem Prob-MA2. Thus, we next prove that $\lim_{T\to\infty}\frac{1}{T}(\mathbf{F}(\mathbf{d}^*,\mathbf{p}^*)-\mathbf{F}(\tilde{\mathbf{d}},\tilde{\mathbf{p}}))=0$, where $(\mathbf{d}^*,\mathbf{p}^*)$ is the solution to Prob-MA3 and $(\tilde{\mathbf{d}},\tilde{\mathbf{p}})$ is the solution to Prob-MA2.

It can be shown that Prob-MA3 is also a convex optimization problem, and the Slater's condition is also satisfied. We derive the KKT conditions of Prob-MA3 as follows:

$$\begin{cases} I_{\mathbb{N}}\left(U'(d_i^*(t),\omega_i(t))-\alpha(d_i^*(t)-\hat{d}_i(t-1))+\nu_i^*(t)\right) \\ \quad +I_{\mathbb{M}}\left(-C'(p_m^*(t))-\alpha(p_m^*(t)-\hat{p}_m(t-1))+\gamma_m^*(t)\right. \\ \quad \left.-\rho_m^*(t)\right)-\mu^*(t)G'(l^*(t))/b_{\max}(t)=0 \\ \mu^*(t)\left(G(l^*(t))/b_{\max}(t)-1\right)=0 \\ \nu_i^*(t)\left(d_i^*(t)-d_{i,\min}(t)\right)=0 \\ \gamma_m^*(t)\left(p_m^*(t)+p_{m,\max}(t)\right)=0 \\ \rho_m^*(t)\left(p_m^*(t)-p_{m,\max}(t)\right)=0 \\ \mu^*(t),\nu_i^*(t),\gamma_m^*(t),\rho_m^*(t)\geq 0 \quad \forall\, i\in\mathbb{N},\, m\in\mathbb{M},\, t\in\mathbb{T} \end{cases} \quad (33)$$

where the nonnegative Lagrangian multipliers $\mu^*(t)$, $\nu_i^*(t)$, $\gamma_m^*(t)$, and $\rho_m^*(t)$ are the dual points where the KKT conditions are satisfied and the optimal value is achieved.

To prove the theorem, we need another differentiable concave function $\mathbf{H}(\tilde{\mathbf{d}},\tilde{\mathbf{p}})$ as defined in (34) shown at the bottom of the page.

Recall that $G(\tilde{l}(t))$ is the generation cost and $b_{\max}(t)$ is the maximum generation cost in the macrogrid. Therefore the second term on the RHS of (34) is nonnegative. Following (33), both the last two terms on the RHS of (34) are nonnegative either. It follows that

$$\mathbf{F}(\tilde{\mathbf{d}},\tilde{\mathbf{p}})\leq\mathbf{H}(\tilde{\mathbf{d}},\tilde{\mathbf{p}}). \quad (35)$$

Due to the concavity and differentiability of $\mathbf{H}(\cdot)$, we have

$$\mathbf{H}(\tilde{\mathbf{d}},\tilde{\mathbf{p}})\leq\mathbf{H}(\mathbf{d}^*,\mathbf{p}^*)+\bigtriangledown\mathbf{H}(\mathbf{d}^*,\mathbf{p}^*)\bullet((\tilde{\mathbf{d}},\tilde{\mathbf{p}})-(\mathbf{d}^*,\mathbf{p}^*)) \quad (36)$$

where $\bullet$ denotes the inner product operation. According to (35) and (36), we can derive inequality (37) shown at the bottom of the page.

Substituting (33) into inequality (37), we have

$$\begin{aligned}\mathbf{F}(\tilde{\mathbf{d}},\tilde{\mathbf{p}})&\leq\mathbf{F}(\mathbf{d}^*,\mathbf{p}^*)\\ &+\sum_{t=1}^{T}\sum_{i\in\mathbb{N}}\alpha(\tilde{d}_i(t)-d_i^*(t))\left(\frac{1}{T}\sum_{k=1}^{T}d_i^*(k)-\hat{d}_i(t-1)\right)\\ &+\sum_{t=1}^{T}\sum_{m\in\mathbb{M}}\alpha(\tilde{p}_m(t)-p_m^*(t))\left(\frac{1}{T}\sum_{k=1}^{T}p_m^*(k)-\hat{p}_m(t-1)\right).\end{aligned}$$

Adding $-\hat{d}_i(T)+\hat{d}_i(T)$ and $-\hat{p}_m(T)+\hat{p}_m(T)$ to the last two terms on the RHS of the above inequality, respectively, taking limit over T on both sides, and applying Lemma 1, we have

$$\lim_{T\to\infty}\frac{\mathbf{F}(\tilde{\mathbf{d}},\tilde{\mathbf{p}})}{T}\leq\lim_{T\to\infty}\frac{\mathbf{F}(\mathbf{d}^*,\mathbf{p}^*)}{T}. \quad (38)$$

$$\begin{aligned}\mathbf{H}(\tilde{\mathbf{d}},\tilde{\mathbf{p}})=&\sum_{t=1}^{T}\left(\sum_{i\in\mathbb{N}}U(\tilde{d}_i(t),\omega_i(t))-\sum_{m\in\mathbb{M}}C_m(\tilde{p}_m(t))\right)-\frac{\alpha T}{2}\left(\sum_{i\in\mathbb{N}}\mathrm{Var}(\vec{\tilde{d}}_{i,T})+\sum_{m\in\mathbb{M}}\mathrm{Var}(\vec{\tilde{p}}_{m,T})\right)-\sum_{t=1}^{T}\mu^*(t)\left(\frac{G(\tilde{l}(t))}{b_{\max}(t)}-1\right)\\ &+\sum_{t=1}^{T}\sum_{i\in\mathbb{N}}\nu_i^*(t)(\tilde{d}_i(t)-d_{i,\min}(t))+\sum_{t=1}^{T}\sum_{m\in\mathbb{M}}(\gamma_m^*(t)(\tilde{p}_m(t)+p_{m,\max}(t))-\rho_m^*(t)(\tilde{p}_m(t)-p_{m,\max}(t)))\\ =&\,\mathbf{F}(\tilde{\mathbf{d}},\tilde{\mathbf{p}})+\sum_{t=1}^{T}\mu^*(t)\left(1-\frac{G(\tilde{l}(t))}{b_{\max}(t)}\right)+\sum_{t=1}^{T}\sum_{i\in\mathbb{N}}\nu_i^*(t)(\tilde{d}_i(t)-d_{i,\min}(t))+\sum_{t=1}^{T}\sum_{m\in\mathbb{M}}((\gamma_m^*(t)-\rho_m^*(t))\tilde{p}_m(t)\\ &+(\gamma_m^*(t)+\rho_m^*(t))p_{m,\max}(t)).\end{aligned} \quad (34)$$

$$\begin{aligned}\mathbf{F}(\tilde{\mathbf{d}},\tilde{\mathbf{p}})\leq&\,\mathbf{F}(\mathbf{d}^*,\mathbf{p}^*)+\sum_{t=1}^{T}\mu^*(t)\left(1-\frac{G(l^*(t))}{b_{\max}(t)}\right)+\sum_{t=1}^{T}\sum_{i\in\mathbb{N}}\nu_i^*(t)(d_i^*(t)-d_{i,\min}(t))+\sum_{t=1}^{T}\sum_{m\in\mathbb{M}}((\gamma_m^*(t)-\rho_m^*(t))p_m^*(t)\\ &+(\gamma_m^*(t)+\rho_m^*(t))p_{m,\max}(t))+\sum_{t=1}^{T}\sum_{i\in\mathbb{N}}(\tilde{d}_i(t)-d_i^*(t))\left(U'(d_i^*(t),\omega_i(t))+\nu_i^*(t)-\mu^*(t)\frac{G'(l^*(t))}{b_{\max}(t)}+\frac{\alpha}{T}\sum_{k=1}^{T}d_i^*(k)-\alpha d_i^*(t)\right)\\ &+\sum_{t=1}^{T}\sum_{m\in\mathbb{M}}(\tilde{p}_m(t)-p_m^*(t))\left(-C_m'(p_m^*(t))+\gamma_m^*(t)-\rho_m^*(t)-\mu^*(t)\frac{G'(l^*(t))}{b_{\max}(t)}+\frac{\alpha}{T}\sum_{k=1}^{T}p_m^*(k)-\alpha p_m^*(t)\right).\end{aligned} \quad (37)$$

On the other hand, $(\tilde{\mathbf{d}}, \tilde{\mathbf{p}})$ is the optimal solution to Prob-MA2 and thus $\mathbf{F}(\tilde{\mathbf{d}}, \tilde{\mathbf{p}})$ is the optimal objective value of Prob-MA2. Since it is a maximization problem, we have

$$\mathbf{F}(\tilde{\mathbf{d}}, \tilde{\mathbf{p}}) \geq \mathbf{F}(\mathbf{d}^*, \mathbf{p}^*). \tag{39}$$

Considering both (38) and (39), we conclude that Theorem 1 holds true. ■

References

[1] B. Lasseter, "Microgrids [distributed power generation]," in *Proc. IEEE Power Eng. Soc. Winter Meeting*, Columbus, OH, USA, Jan. 2001, pp. 146–149.

[2] F. Katiraei, R. Iravani, N. Hatziargyriou, and A. Dimeas, "Microgrids management," *IEEE Power Energy*, vol. 6, no. 3, pp. 54–65, May 2008.

[3] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid–The new and improved power grid: A survey," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 4, pp. 944–980, Apr. 2012.

[4] V. Gungor *et al.*, "Smart grid technologies: Communication technologies and standards," *IEEE Trans. Ind. Informat.*, vol. 7, no. 4, pp. 529–539, Nov. 2011.

[5] V. Gungor *et al.*, "A survey on smart grid potential applications and communication requirements," *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 28–42, Feb. 2013.

[6] N. Hatziargyriou, H. Asano, R. Iravani, and C. Marnay, "Microgrids: An overview of ongoing research, development, and demonstration projects," *IEEE Power Energy Mag.*, vol. 5, no. 4, pp. 78–94, Jul. 2007.

[7] L. Valverde, F. Rosa, and C. Bordons, "Design, planning and management of a hydrogen-based microgrid," *IEEE Trans. Ind. Informat.*, vol. 9, no. 3, pp. 1398–1404, Aug. 2013.

[8] A. Tsikalakis and N. Hatziargyriou, "Centralized control for optimizing microgrids operation," *IEEE Trans. Energy Convers.*, vol. 23, no. 1, pp. 241–248, Mar. 2008.

[9] S.-J. Ahn, S.-R. Nam, J.-H. Choi, and S.-I. Moon, "Power scheduling of distributed generators for economic and stable operation of a microgrid," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 398–405, Mar. 2013.

[10] T. Logenthiran, D. Srinivasan, and D. Wong, "Multi-agent coordination for DER in MicroGrid," in *Proc. IEEE Int. Conf. Sustain. Energy Technol. (ICSET'08)*, Singapore, Nov. 2008, pp. 77–82.

[11] B. Bahrani, M. Saeedifard, A. Karimi, and A. Rufer, "A multivariable design methodology for voltage control of a single-DG-unit microgrid," *IEEE Trans. Ind. Informat.*, vol. 9, no. 2, pp. 589–599, May 2013.

[12] J.-Y. Kim *et al.*, "Cooperative control strategy of energy storage system and microsources for stabilizing the microgrid during islanded operation," *IEEE Trans. Power Electron.*, vol. 25, no. 12, pp. 3037–3048, Dec. 2010.

[13] I. Balaguer, Q. Lei, S. Yang, U. Supatti, and F. Z. Peng, "Control for grid-connected and intentional islanding operations of distributed power generation," *IEEE Trans. Ind. Electron.*, vol. 58, no. 1, pp. 147–157, Jan. 2011.

[14] M. Di Silvestre, G. Graditi, and E. Sanseverino, "A generalized framework for optimal sizing of distributed energy resources in micro-grids using an indicator-based swarm approach," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 152–162, Feb. 2014.

[15] T. Vandoorn, B. Renders, L. Degroote, B. Meersman, and L. Vandevelde, "Active load control in islanded microgrids based on the grid voltage," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 139–151, Mar. 2011.

[16] J. Matamoros, D. Gregoratti, and M. Dohler, "Microgrids energy trading in islanding mode," in *Proc. IEEE Smart Grid Commun. (SmartGridComm'12)*, Tainan City, Taiwan, Nov. 2012, pp. 49–54.

[17] W. Saad, Z. Han, and H. Poor, "Coalitional game theory for cooperative micro-grid distribution networks," in *Proc. IEEE Int. Conf. Commun. Workshop (ICC'11)*, Kyoto, Japan, Jun. 2011, pp. 1–5.

[18] H. Dagdougui and R. Sacile, "Decentralized control of the power flows in a network of smart microgrids modeled as a team of cooperative agents," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 2, pp. 510–519, Mar. 2014.

[19] C. Wei, Z. Fadlullah, N. Kato, and I. Stojmenovic, "On optimally reducing power loss in micro-grids with power storage devices," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1361–1370, Jul. 2014.

[20] S. Zhang, J. Yang, X. Wu, and R. Zhu, "Dynamic power provisioning for cost minimization in islanding micro-grid with renewable energy," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. (ISGT)*, Washington, DC, USA, Feb. 2014, pp. 1–5.

[21] R. Minciardi and R. Sacile, "Optimal control in a cooperative network of smart power grids," *IEEE Syst. J.*, vol. 6, no. 1, pp. 126–133, Mar. 2012.

[22] D. Nguyen and L. B. Le, "Optimal energy management for cooperative microgrids with renewable energy resources," in *Proc. IEEE Smart Grid Commun. (SmartGridComm)*, Vancouver, BC, USA, Canada, Oct. 2013, pp. 678–683.

[23] Y. Wang, S. Mao, and R. Nelms, "Online algorithm for optimal real-time energy distribution in the smart grid," *IEEE Trans. Emerg. Topics Comput.*, vol. 1, no. 1, pp. 10–21, Jun. 2013.

[24] Y. Wang, G. Cao, S. Mao, and R. Nelms, "Analysis of solar generation and weather data in smart grid with simultaneous inference of nonlinear time series," in *Proc. Int. Workshop Smart Cities Urban Informat. (SmartCity'15)*, Hong Kong, China, Apr. 2015, pp. 1–6.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[26] Y. Wang, S. Mao, and R. Nelms, "Distributed online algorithm for optimal real-time energy distribution in the smart grid," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 70–80, Feb. 2014.

[27] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

[28] SCE. (2011). *Regulatory Information–SCE Load Profiles–2011 Static Load Profiles* [Online]. Available: http://www.sce.com/005_regul_info/eca/DOMSM11.DLP

[29] H. Li and W. Zhang, "QoS routing in smart grid," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM'10)*, Miami, FL, USA, Dec. 2010, pp. 1–6.

Yu Wang (S'13) received the M.E. degree in instrument science and technology, and the B.E. degree in measuring and control technology and instrumentation from Southeast University, Nanjing, China, in 2011 and 2008, respectively. Since 2011, he has been pursuing the Ph.D. degree from the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA.

His research interests include smart grid and optimization.

Shiwen Mao (S'99–M'04–SM'09) received the Ph.D. degree in electrical and computer engineering from Polytechnic University, Brooklyn, NY, USA.

Currently, he is the McWane Associate Professor with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA. His research interests include wireless networks and multimedia communications, such as cognitive radio, small cells, 60-GHz mmWave networks, free-space optical networks, and smart grid.

Dr. Mao is on the Editorial Board of the IEEE Transactions on Wireless Communications, the IEEE Internet of Things Journal, and the IEEE Communications Surveys and Tutorials. He is a Distinguished Lecturer of the IEEE Vehicular Technology Society. He received the 2013 IEEE Communications Society Multimedia Communications Technical Committee Outstanding Leadership Award and the National Science Foundation CAREER Award in 2010. He is a co-recipient of the 2015 IEEE Wireless Communications and Networking Conference Best Paper Award, the 2013 IEEE International Conference on Communications Best Paper Award, and the 2004 IEEE Communications Society Leonard G. Abraham Prize in the field of communication systems.

R. M. Nelms (F'04) received the B.E.E. and M.S. degrees in electrical engineering from Auburn University, Auburn, AL, USA, in 1980 and 1982, respectively, and the Ph.D. degree in electrical engineering from Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, in 1987.

Currently, he is a Professor and Chair of the Department of Electrical and Computer Engineering, Auburn University. His research interests include power electronics, power systems, and electric machinery.

Dr. Nelms is a Registered Professional Engineer in Alabama.