An Autonomous Demand Response Algorithm based on Online Convex Optimization

Shahab Bahrami, Yu Christine Chen, and Vincent W.S. Wong
Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, Canada
email: {bahramis, chen, vincentw}@ece.ubc.ca

***Abstract*—A price-based demand response program is a viable solution for distribution network operators (DNOs) to motivate electricity consumers toward scheduling their load demand during off-peak periods. This paper addresses the problem of load scheduling in a demand response program, while accounting for load demand uncertainty and the distribution network operational constraints. The centralized load control is a nonconvex optimization problem due to the ac power flow equations. We use convex relaxation techniques to transform the problem into a semidefinite program (SDP), which is solved using online convex optimization techniques to address the load demand uncertainty. To tackle the issue of computational complexity, we use proximal Jacobian alternating direction method of multipliers (PJ-ADMM) to decompose the centralized problem into the customers' load scheduling subproblems. The decentralized algorithm is executed by each customer to schedule its load demand in real-time. Via simulations on the IEEE 37-bus test feeder, we show that the proposed algorithm enables customers to approximate the optimal load profile in the benchmark scenario without load uncertainty, and the approximation is tight. Furthermore, we show a negligible gap of 2.3% between the customers' cost with the proposed algorithm and the cost in the benchmark scenario.**

I. Introduction

Rapidly increasing electricity demand renders peak load management a critical and challenging issue for distribution network operators (DNOs). Demand response solutions can provide DNOs with a variety of peak load management strategies, such as direct load control to reduce the operational cost of the power grid [1]. A limitation of direct load control lies in customer privacy concerns, as the customer's overall energy consumption profile must be exposed to the DNO. Moreover, a centralized control scheme may incur significant computation and communication overhead. Recent advancements in smart metering have motivated DNOs to implement demand response programs with real-time pricing (RTP) [1]. These enable customers to fully exploit the flexibility in their consumption habits without invasion of their privacy. Furthermore, household energy consumption controllers (ECCs) can make demand response decisions on behalf of their corresponding customers in a distributed and parallel fashion resulting in lower computation requirements for the DNO.

There are various challenges in implementing a demand response program with RTP in a distribution feeder. First, the DNO must properly set energy price signals in real-time to motivate customers toward load scheduling, while maintaining grid-wide operational constraints (e.g., power balance equations, and bus voltage and branch flows limits). Second, an ECC may be uncertain about the customer's load demand and preferences (e.g., desirable load, demand variation flexibility). Third, a fast coordinating mechanism is required for real-time information exchange among the DNO and ECCs, thereby meeting the supply-demand balance continuously.

There have been some efforts in the literature to tackle the aforementioned challenges. Mechanisms such as stochastic approximation [2], robust optimization [3], reinforcement learning [4], Bayesian game [5], stochastic game [6], and online convex optimization [7] address the load uncertainties in a demand response program with RTP. However, these approaches do not consider the nonconvex constraints imposed by the topology and operation of the distribution network. A few works (e.g., [8]–[10]) have considered the distribution network operational constraints to study the coordination among demand response participants. However, these studies do not consider the uncertainties in the customers' power consumption and preferences.

In this paper, we propose a demand response framework with RTP that incorporates (i) operational constraints imposed by the distribution network, (ii) customer's load demand and preferences uncertainty, and (iii) fast decentralized optimization algorithm. The ECCs make real-time load scheduling decisions based on their past experiences and the control signals from the DNO to satisfy the network's operational constraints. The main contributions of this paper are as follows:

- *Distribution Network Constraints:* We consider the full ac power flow model and formulate the deterministic centralized load control problem as a nonconvex optimization problem. Then, we apply convex relaxation techniques to transform the problem into a semidefinite program (SDP).
- *Addressing Load Uncertainty:* We study the stochastic centralized load control problem. The stochastic process for variations in the customers' load demand and preferences may not be available. To address this challenge, we apply online convex optimization techniques [11] that do not require any knowledege on the stochatic process for the uncertain parameters. In our demand response framework, the DNO observes the actual values for the uncertain parameters and performs a real-time demand adjustment, which is formulated as an SDP. The DNO also uses projected gradient method [12] to determine the scheduled load demand for the next time interval.
- *Decentralized Algorithm Design:* To mitigiate the computation complexity of the proposed centralized online convex optimization model, we apply the alternating direction method of multipliers (ADMM) approach to decompose the problem into subproblems for the substation and customers. In particular, we use proximal

Jacobian (PJ)-ADMM [13, Algorithm 4] with prox-linear method [14]. We show that the proposed PJ-ADMM-based algorithm converges to a global optimal solution of the stochastic centralized load control problem.

- *Performance Evaluation:* We perform simulations on the IEEE 37-bus test feeder. We consider a demand response program with RTP scheme combined with inclining block rate (IBR). Simulation results show that the load demand during peak hours is reduced by 14% on average in the demand response program. Results also show that the scheduled load profile using the proposed algorithm with uncertainty approximates the load profile in the benchmark scenario without uncertainty, and the approximation is tight. Moreover, the gap between the total cost with the proposed algorithm and the cost in the benchmark scenario is 2.3%, which is negligible. Our algorithm converges to the solution of the centralized problem in 15 iterations on average. When compared with the centralized approach and the variable splitting (VS)-ADMM-based decentralized algorithm, our algorithm based on PJ-ADMM has a lower running time per time slot.

II. System Model

Consider a distribution feeder consisting of N buses. Let $\mathcal{N} = \{1, \ldots, N\}$ denote the set of buses, where bus N corresponds to the substation bus and bus $n \in \mathcal{N}^-$ corresponds to customer n, and $\mathcal{N}^- = \{1, \ldots, N-1\}$. Let $\mathcal{L} \subseteq \mathcal{N} \times \mathcal{N}$ denote the set of transmission lines. Each bus is equipped with an ECC. The ECC of each customer is responsible for scheduling the energy consumption of that customer. The ECC at the substation bus is responsible for controlling the active and reactive power flow into the distribution feeder. Each ECC is connected to the DNO via a two-way communication network, which enables information exchange of the price and energy demand for the corresponding bus. A demand response program is established for a predetermined period of time (e.g., one day), which is divided into a set $\mathcal{T} = \{1, \ldots, T\}$ of T time slots, each with equal duration, e.g., 15 minutes per time slot.

Let Y denote the network admittance matrix. For bus $n \in \mathcal{N}$, let e_n denote the n^{th} basis column vector in $\mathbb{R}^N$ and $Y_n = e_n e_n^{\mathrm{T}} Y$. Row n of matrix Y_n is equal to row n of the admittance matrix Y, and other entries of Y_n are zero. We use the lumped-element Π model for transmission lines. Let y_{nm} and $\overline{y}_{nm}$ denote the series and shunt admittance values at bus n for the line $(n,m) \in \mathcal{L}$, respectively. We define $Y_{nm} = (\overline{y}_{nm} + y_{nm}) e_n e_n^{\mathrm{T}} - y_{nm} e_n e_m^{\mathrm{T}}$, so that the entries (n,n) and (n,m) of Y_{nm} are $\overline{y}_{nm} + y_{nm}$ and $-y_{nm}$, respectively. Other entries of Y_{nm} are zero. We further define matrices $\mathbf{Y}_n$, $\overline{\mathbf{Y}}_n$, $\mathbf{Y}_{nm}$, $\overline{\mathbf{Y}}_{nm}$, and $\mathbf{M}_n$ as follows:

$$\mathbf{Y}_n = \frac{1}{2} \begin{bmatrix} \mathrm{Re}\{Y_n + Y_n^{\mathrm{T}}\} & \mathrm{Im}\{Y_n^{\mathrm{T}} - Y_n\} \\ \mathrm{Im}\{Y_n - Y_n^{\mathrm{T}}\} & \mathrm{Re}\{Y_n + Y_n^{\mathrm{T}}\} \end{bmatrix},$$

$$\overline{\mathbf{Y}}_n = -\frac{1}{2} \begin{bmatrix} \mathrm{Im}\{Y_n + Y_n^{\mathrm{T}}\} & \mathrm{Re}\{Y_n - Y_n^{\mathrm{T}}\} \\ \mathrm{Re}\{Y_n^{\mathrm{T}} - Y_n\} & \mathrm{Im}\{Y_n + Y_n^{\mathrm{T}}\} \end{bmatrix},$$

$$\mathbf{Y}_{nm} = \frac{1}{2} \begin{bmatrix} \mathrm{Re}\{Y_{nm} + Y_{nm}^{\mathrm{T}}\} & \mathrm{Im}\{Y_{nm}^{\mathrm{T}} - Y_{nm}\} \\ \mathrm{Im}\{Y_{nm} - Y_{nm}^{\mathrm{T}}\} & \mathrm{Re}\{Y_{nm} + Y_{nm}^{\mathrm{T}}\} \end{bmatrix},$$

$$\overline{\mathbf{Y}}_{nm} = -\frac{1}{2} \begin{bmatrix} \mathrm{Im}\{Y_{nm} + Y_{nm}^{\mathrm{T}}\} & \mathrm{Re}\{Y_{nm} - Y_{nm}^{\mathrm{T}}\} \\ \mathrm{Re}\{Y_{nm}^{\mathrm{T}} - Y_{nm}\} & \mathrm{Im}\{Y_{nm} + Y_{nm}^{\mathrm{T}}\} \end{bmatrix},$$

$$\mathbf{M}_n = \begin{bmatrix} e_n e_n^{\mathrm{T}} & 0 \\ 0 & e_n e_n^{\mathrm{T}} \end{bmatrix}.$$

Let $V_n(t)$ denote the voltage phasor of bus n in time slot t. Let $\mathbf{v}(t) = (V_n(t),\, n \in \mathcal{N})$ denote the vector of bus voltages. Let $\mathcal{N}_n \subseteq \mathcal{N}$ denote the set of buses connected to bus n. We construct the vector of bus voltages $\mathbf{v}_n(t)$ for bus n from vector $\mathbf{v}(t)$ such that the elements $m \in \mathcal{N}_n \cup \{n\}$ of vectors $\mathbf{v}_n(t)$ and $\mathbf{v}(t)$ are equal, and other elements of vector $\mathbf{v}_n(t)$ are zero. For bus n, we define variable vector $\mathbf{x}_n(t) = \left[(\mathrm{Re}\{\mathbf{v}_n(t)\})^{\mathrm{T}} \;\; (\mathrm{Im}\{\mathbf{v}_n(t)\})^{\mathrm{T}} \right]^{\mathrm{T}}$ consisting of the real and imaginary parts of $\mathbf{v}_n(t)$ in time slot t. We define variable matrix $\mathbf{W}_n(t) = \mathbf{x}_n(t)(\mathbf{x}_n(t))^{\mathrm{T}}$ for bus n in time slot t. We denote the active and reactive loads of customer $n \in \mathcal{N}^-$ in time slot t by $P_n(t)$ and $Q_n(t)$, respectively. We also denote the injected active power and reactive power into the substation bus N in time slot t by $P_N(t)$ and $Q_N(t)$, respectively. Let $|V_n(t)|$ denote the voltage magnitude of bus $n \in \mathcal{N}$ in time slot t. Let $S_{nm}(t)$ denote the apparent power flow through line $(n,m) \in \mathcal{L}$ in time slot t. With the notations established above, for time slot $t \in \mathcal{T}$, we can show that [15]

$$P_n(t) = -\mathrm{Tr}\{\mathbf{Y}_n \mathbf{W}_n(t)\}, \qquad n \in \mathcal{N}^- \tag{1a}$$

$$Q_n(t) = -\mathrm{Tr}\{\overline{\mathbf{Y}}_n \mathbf{W}_n(t)\}, \qquad n \in \mathcal{N}^- \tag{1b}$$

$$P_N(t) = \mathrm{Tr}\{\mathbf{Y}_N \mathbf{W}_N(t)\}, \tag{1c}$$

$$Q_N(t) = \mathrm{Tr}\{\overline{\mathbf{Y}}_N \mathbf{W}_N(t)\}, \tag{1d}$$

$$|V_n(t)|^2 = \mathrm{Tr}\{\mathbf{M}_n \mathbf{W}_n(t)\}, \qquad n \in \mathcal{N} \tag{1e}$$

$$|S_{nm}(t)|^2 = \mathrm{Tr}\{\mathbf{Y}_{nm} \mathbf{W}_n(t)\}^2 + \mathrm{Tr}\{\overline{\mathbf{Y}}_{nm} \mathbf{W}_n(t)\}^2, (n,m) \in \mathcal{L}. \tag{1f}$$

We use (1a)–(1f) to obtain the operational constraints imposed by the customers and distribution network. For customer n, we assume that the load demand in time slot $t \in \mathcal{T}$ consists of the uncontrollable load (i.e., base load) $P_n^{\mathrm{b}}(t)$ and controllable load demand $P_n^{\mathrm{c}}(t)$, i.e., we have $P_n(t) = P_n^{\mathrm{b}}(t) + P_n^{\mathrm{c}}(t)$. Let $P_n^{\mathrm{des}}(t)$ denote the *desirable* active power load demand for customer n in time slot t, had the customer not participated in the demand response program. By participating in the demand response program, customer n incurs a discomfort cost $d_n(P_n(t))$ in time slot t. It is used as a metric in monetary unit to express the customer's dissatisfaction with changing its load profile from the desirable pattern to the scheduled pattern. Similar to [16], we model the discomfort cost function as a weighted Euclidean distance between the scheduled and the desirable load demands. Let $\omega_n(t)$, $n \in \mathcal{N}^-$, $t \in \mathcal{T}$ denote a positive weighting coefficient measured in cents/kW^2 to reflect the customer's discomfort caused by changing the load demand from its desirable amount. The discomfort cost for customer $n \in \mathcal{N}^-$ in time slot t is obtained as

$$d_n(\mathbf{W}_n(t)) = \omega_n(t) \left(P_n^{\mathrm{des}}(t) + \mathrm{Tr}\{\mathbf{Y}_n \mathbf{W}_n(t)\} \right)^2. \tag{2}$$

The controllable load demand for customer n in time slot t is within the minimum and maximum limits $P_n^{\mathrm{c,min}}(t)$ and

$P_n^{\rm c,max}(t)$, respectively. Thus, the total active load demand of customer n in time slot t is within the limits $P_n^{\min}(t) = P_n^{\rm b}(t) + P_n^{\rm c,min}(t)$ and $P_n^{\max}(t) = P_n^{\rm b}(t) + P_n^{\rm c,max}(t)$. We use (1a) and obtain the following constraint for $t \in \mathcal{T}$, $n \in \mathcal{N}^-$:

$$P_n^{\min}(t) \leq -{\rm Tr}\{\mathbf{Y}_n \mathbf{W}_n(t)\} \leq P_n^{\max}(t). \tag{3}$$

Assume that ECC n can control the reactive power consumption at bus n in time slot t such that it is within the limits $Q_n^{\min}(t)$ and $Q_n^{\max}(t)$. We use (1b) to obtain the following constraint for $t \in \mathcal{T}$, $n \in \mathcal{N}^-$:

$$Q_n^{\min}(t) \leq -{\rm Tr}\{\overline{\mathbf{Y}}_n \mathbf{W}_n(t)\} \leq Q_n^{\max}(t). \tag{4}$$

The injected active power $P_N(t)$ and reactive power $Q_N(t)$ into the substation bus in time slot t are bounded by the *predetermined* limits $P_N^{\max}$, $Q_N^{\min}$, and $Q_N^{\max}$, respectively. For $t \in \mathcal{T}$, we use (1c) and (1d) to obtain

$$0 \leq {\rm Tr}\{\mathbf{Y}_N \mathbf{W}_N(t)\} \leq P_N^{\max}, \tag{5a}$$
$$Q_N^{\min} \leq {\rm Tr}\{\overline{\mathbf{Y}}_N \mathbf{W}_N(t)\} \leq Q_N^{\max}. \tag{5b}$$

Let $V_n^{\min}$ and $V_n^{\max}$ denote the lower and upper bounds on the voltage magnitude at bus n, respectively. For $t \in \mathcal{T}$, we use (1e) to obtain

$$(V_n^{\min})^2 \leq {\rm Tr}\{\mathbf{M}_n \mathbf{W}_n(t)\} \leq (V_n^{\max})^2, \;\; n \in \mathcal{N}. \tag{6}$$

Let $S_{nm}^{\max}$ denote the maximum apparent power flow through the line $(n,m) \in \mathcal{L}$. We substitute (1f) into $|S_{nm}(t)| \leq S_{nm}^{\max}$, and by constructing its matrix form, for $(n,m) \in \mathcal{L}$, we have

$$\begin{bmatrix} (S_{nm}^{\max})^2 & {\rm Tr}\{\mathbf{Y}_{nm}\mathbf{W}_n(t)\} & {\rm Tr}\{\overline{\mathbf{Y}}_{nm}\mathbf{W}_n(t)\} \\ {\rm Tr}\{\mathbf{Y}_{nm}\mathbf{W}_n(t)\} & 1 & 0 \\ {\rm Tr}\{\overline{\mathbf{Y}}_{nm}\mathbf{W}_n(t)\} & 0 & 1 \end{bmatrix} \succeq 0. \tag{7}$$

Let $\mathbf{W}_n^{k,k'}(t)$ denote the entry (k,k') of matrix $\mathbf{W}_n(t)$. For a neighboring bus $m \in \mathcal{N}_n$, the values of $\mathbf{W}_m^{n,n}(t)$ and $\mathbf{W}_m^{n+N,n+N}(t)$ are, respectively, equal to the square of real and imaginary parts of complex-valued voltage phasor at bus n in time slot t. Thus, for $t \in \mathcal{T}$, we have

$$\mathbf{W}_m^{n,n}(t) = \mathbf{W}_n^{n,n}(t), \qquad m \in \mathcal{N}_n,\, n \in \mathcal{N} \tag{8a}$$
$$\mathbf{W}_m^{n+N,n+N}(t) = \mathbf{W}_n^{n+N,n+N}(t), \qquad m \in \mathcal{N}_n,\, n \in \mathcal{N}. \tag{8b}$$

III. Centralized Load Control Problem

In this section, we study the load control problem in the demand response program. The reactive power control is usually performed for the sake of voltage regulation. Without loss of generality, we assume that the limits $Q_n^{\min}(t)$ and $Q_n^{\max}(t)$, $t \in \mathcal{T}$ are known *a priori* by ECC n. However, we take into account the uncertainty about parameters $P_n^{\min}(t)$, $P_n^{\max}(t)$, $\omega_n(t)$, and $P_n^{\rm des}(t)$ for the active power consumption of customer n in time slot t. Below, we formulate the *centralized* load control with and without the DNO's uncertainty about the active load demand and preferences of customers.

A. Deterministic Centralized Load Control

We assume that parameters $P_n^{\min}(t)$, $P_n^{\max}(t)$, $\omega_n(t)$, and $P_n^{\rm des}(t)$ for customer $n \in \mathcal{N}^-$ in time slot $t \in \mathcal{T}$ are known *a priori* by the DNO. The objective function of the DNO's load control problem in time slot t is the total cost of all customers in the feeder, i.e., the *social cost*. The cost of customer n in time slot t includes the discomfort cost in (2) and the electricity bill payment $\pi_n(t)P_n(t)$, where $\pi_n(t)$ is the electricity price (cents per unit of active power) in time slot t for customer n in the demand response program.

The discomfort cost in (2) is a quadratic function of $\mathbf{W}_n(t)$. To obtain its SDP form, we introduce the auxiliary variable $\theta_n(t)$, $n \in \mathcal{N}^-$, $t \in \mathcal{T}$, such that $d_n(\mathbf{W}_n(t)) \leq \theta_n(t)$. The DNO's objective function in time slot t is obtained as

$$f^{\rm obj}(t) = \textstyle\sum_{n \in \mathcal{N}^-} \left(\theta_n(t) - \pi_n(t)\,{\rm Tr}\{\mathbf{Y}_n \mathbf{W}_n(t)\}\right). \tag{9}$$

We include the following constraint corresponding to inequality $d_n(\mathbf{W}_n(t)) \leq \theta_n(t)$, $t \in \mathcal{T}$ into the constraint set of the DNO's problem:

$$\begin{bmatrix} \theta_n(t)/\omega_n(t) & P_n^{\rm des}(t) + {\rm Tr}\{\mathbf{Y}_n \mathbf{W}_n(t)\} \\ P_n^{\rm des}(t) + {\rm Tr}\{\mathbf{Y}_n \mathbf{W}_n(t)\} & 1 \end{bmatrix} \succeq 0,\; n \in \mathcal{N}^-. \tag{10}$$

The centralized load control problem in time slot $t \in \mathcal{T}$ is as follows:

$$\begin{aligned} \boldsymbol{\mathcal{P}}_1(t): \quad & \underset{\substack{\mathbf{W}_n(t),\, n \in \mathcal{N} \\ \theta_n(t),\, n \in \mathcal{N}^-}}{\text{minimize}} \quad f^{\rm obj}(t) \\ & \text{subject to constraints (3)–(8b) and (10),} \\ & \qquad \mathbf{W}_n(t) \succeq 0, \qquad n \in \mathcal{N}, \\ & \qquad {\rm rank}(\mathbf{W}_n(t)) = 1, \qquad n \in \mathcal{N}. \end{aligned}$$

Problem $\boldsymbol{\mathcal{P}}_1(t)$ is a nonconvex optimization problem due to the rank-one constraint. We relax the rank constraint to obtain the following SDP relaxation form of $\boldsymbol{\mathcal{P}}_1(t)$:

$$\begin{aligned} \boldsymbol{\mathcal{P}}_2(t): \quad & \underset{\substack{\mathbf{W}_n(t),\, n \in \mathcal{N} \\ \theta_n(t),\, n \in \mathcal{N}^-}}{\text{minimize}} \quad f^{\rm obj}(t) \\ & \text{subject to constraints (3)–(8b) and (10),} \\ & \qquad \mathbf{W}_n(t) \succeq 0, \qquad n \in \mathcal{N}. \end{aligned}$$

We can show practical distribution networks (including IEEE test feeders) satisfy the sufficient conditions given in [15, Sec. IV-C] for the network topology, constraints, and their corresponding dual variables. Hence, the SDP relaxation gap between problems $\boldsymbol{\mathcal{P}}_1(t)$ and $\boldsymbol{\mathcal{P}}_2(t)$ is zero, i.e., the solution matrices $\mathbf{W}_n^{\rm opt}(t)$, $n \in \mathcal{N}$ to problem $\boldsymbol{\mathcal{P}}_2(t)$ are all rank one.

In practice, the DNO has uncertainty about the customers' load demand and preferences. In the following, we apply online convex optimization technique to address the uncertainty.

B. Stochastic Centralized Load Control

In the online convex optimization model, we consider the DNO's uncertainty about parameters $P_n^{\min}(t)$, $P_n^{\max}(t)$, $\omega_n(t)$, and $P_n^{\rm des}(t)$ for customer $n \in \mathcal{N}^-$ in time slot t. During time slot t, the DNO receives information about the *actual* values of the uncertain parameters from its customers, which may be different from the predicted values. To cope with the uncertainty, the DNO can decide to adjust the customers' load demands $P_n(t)$, $n \in \mathcal{N}^-$ by $\Delta P_n(t)$, i.e., either purchase power from the spot market with price $\pi^{\rm b}(t)$ (if $\Delta P_n(t) > 0$) or sell power to the spot market with price $\pi^{\rm s}(t)$ (if $\Delta P_n(t) < 0$). We assume that $\pi^{\rm b}(t) \geq \pi^{\rm s}(t)$, and thus the monetary transaction with the spot market is obtained as a piece-wise linear function $\max\{-\pi^{\rm b}(t)\Delta P_n(t),\, \pi^{\rm s}(t)\Delta P_n(t)\}$. By adjusting the load demands, the discomfort cost of customer n becomes $\omega_n(t)\left(P_n^{\rm des}(t) - \big(P_n(t)+\Delta P_n(t)\big)\right)^2$.

We substitute $P_n(t) + \Delta P_n(t) = -\text{Tr}\{\mathbf{Y}_n\mathbf{W}_n(t)\}$, $n \in \mathcal{N}^-$ into the monetary transaction with the spot market and the customers' discomfort cost function. To transform the piece-wise linear function into an SDP, we introduce an auxiliary variable $\delta_n(t)$, and consider the following constraints:

$$-\pi^{\rm b}(t)\big(-\text{Tr}\{\mathbf{Y}_n\mathbf{W}_n(t)\} - P_n(t)\big) \leq \delta_n(t),\ n \in \mathcal{N}^- \quad (11a)$$
$$\pi^{\rm s}(t)\big(-\text{Tr}\{\mathbf{Y}_n\mathbf{W}_n(t)\} - P_n(t)\big) \leq \delta_n(t),\ n \in \mathcal{N}^-. \quad (11b)$$

The discomfort cost is obtained as a quadratic function of matrix $\mathbf{W}_n(t)$. Recall that we introduced auxiliary variable $\theta_n(t)$ for the discomfort cost of customer $n \in \mathcal{N}^-$ and included constraint (10). Hence, the total cost of DNO in the online convex optimization model is obtained as:

$$\widetilde{f}^{\rm obj}(t) = \textstyle\sum_{n\in\mathcal{N}^-} \pi_n(t)P_n(t) + g^{\rm obj}(t),\ \ t \in \mathcal{T}, \qquad (12)$$

where $g^{\rm obj}(t)$ is the optimal value of the following optimization problem under the given values of $P_n(t)$, $n \in \mathcal{N}^-$:

$$\boldsymbol{\mathcal{P}}_3(t): \quad \underset{\substack{\mathbf{W}_n(t),\, n\in\mathcal{N} \\ \theta_n(t),\, \delta_n(t),\, n\in\mathcal{N}^-}}{\text{minimize}} \quad \sum_{n\in\mathcal{N}} c_n(t)$$
$$\text{subject to constraints (3)}-\text{(8b) and (10)}-\text{(11b)},$$
$$\mathbf{W}_n(t) \succeq 0, \qquad n \in \mathcal{N},$$

where $c_n(t) = \theta_n(t) + \delta_n(t)$ for customer $n \in \mathcal{N}^-$ and $c_N(t) = 0$ for substation bus N. Problem $\boldsymbol{\mathcal{P}}_3(t)$ is a convex optimization problem. We can also show that it satisfies the sufficient conditions given in [15, Sec. IV-C]. Thus, the solution matrices $\mathbf{W}_n^{\rm opt}(t)$, $n \in \mathcal{N}$ to problem $\boldsymbol{\mathcal{P}}_3(t)$ are all rank one. Hence, the optimal values of $\Delta P_n(t)$, $n \in \mathcal{N}^-$ can be determined *under the given* values of $P_n(t)$, $n \in \mathcal{N}^-$.

Next the DNO computes the *updated* scheduled load demands $P_n(t+1)$, $n \in \mathcal{N}^-$ for time slot $t+1$ using the following projected gradient-based update rule [12]:

$$P_n(t+1) = \Big[P_n(t) - \alpha(t)\big(\nabla_{P_n(t)}\widetilde{f}^{\rm obj}(t) + \mu_n(t)\big)\Big]^+, \quad (13)$$

where $\alpha(t) > 0$ is the step size in time slot t, $\nabla_{P_n(t)}\widetilde{f}^{\rm obj}(t)$ is the gradient of $\widetilde{f}^{\rm obj}(t)$. Also in (13), $\mu_n(t) = \overline{\mu}_n(t) - \underline{\mu}_n(t)$, where $\overline{\mu}_n(t)$ and $\underline{\mu}_n(t)$ are the dual variables associated with the upper and lower inequalities in (3), respectively. $[\cdot]^+$ is the projection onto the positive orthant. Let $f^{\rm obj,opt}(t)$ denote the optimal value of problem $\boldsymbol{\mathcal{P}}_2(t)$.

Remark 1 (Competitive Analysis): Under the mild conditions that $\widetilde{f}^{\rm obj}(t)$, $n \in \mathcal{N}^-$ are all Lipschitz and convex functions, we can show that solving problem $\boldsymbol{\mathcal{P}}_3(t)$ and the update in (13) achieve a regret $R(T) = \sum_{t\in\mathcal{T}}(\widetilde{f}^{\rm obj}(t) - f^{\rm obj,opt}(t))$ that grows sublinearly in T. That is, $R(T)/T$ tends to zero as T approaches infinity [11, Ch. 5].

Algorithm 1 ECCs and DNO Interaction in time slot t.

1: ECC $n \in \mathcal{N}^-$ randomly set $P_n(1)$ in time slot $t := 1$.

Load adjustment phase:

2: Set $i := 1$ and $\epsilon := 10^{-3}$.
3: ECC $n \in \mathcal{N}$ initializes $\mathbf{W}_n^1(t)$. DNO initializes $\boldsymbol{\lambda}^1(t)$.
4: **Repeat**
5: DNO computes control signals $\mathbf{S}_n^i(t)$ according to (15a)−(15d) and sends to the ECC at bus $n \in \mathcal{N}$.
6: ECC n solves problem $\boldsymbol{\mathcal{P}}_{3,n}^i(t)$, and sends the updated vector $\mathbf{v}_n^{i+1}(t)$ of the bus voltages to the DNO.
7: DNO updates $\boldsymbol{\lambda}^i(t)$ according to (18).
8: $i := i + 1$.
9: **Until** $||\boldsymbol{\lambda}^i(t) - \boldsymbol{\lambda}^{i-1}(t)|| \leq \epsilon$.

Load scheduling phase:

10: ECC $n \in \mathcal{N}^-$ determines $P_n(t+1)$ according to (13).

IV. Decentralized Algorithm Design

In this section, we develop a decentralized algorithm to solve problem $\boldsymbol{\mathcal{P}}_3(t)$ and perform the update in (13) in a parallel fashion. Algorithm 1 describes the interactions between the DNO and ECC $n \in \mathcal{N}^-$ in time slot t. In time slot $t = 1$, ECC $n \in \mathcal{N}^-$ randomly initializes $P_n(1)$ for the scheduled load demand in Line 1. Lines 2 to 9 correspond to the load adjustment phase, in which problem $\boldsymbol{\mathcal{P}}_3(t)$ is solved by the ECCs at buses $n \in \mathcal{N}$ in an iterative and parallel fashion.

Problem $\boldsymbol{\mathcal{P}}_3(t)$ has a separable objective function and coupling constraints (8a) and (8b). The VS-ADMM algorithm [13, Algorithm 1] has been commonly used to develop decentralized algorithms. However, the number of variables and constraints substantially increases by introducing splitting variable. To address this issue, we use PJ-ADMM [13, Algorithm 4] with prox-linear method [14] that does not require the splitting variables. We define the vector of Lagrange multipliers $\boldsymbol{\lambda}(t) = (\lambda_m^n(t),\, \overline{\lambda}_m^n(t),\ m \in \mathcal{N}_n,\, n \in \mathcal{N})$, where $\lambda_m^n(t)$ and $\overline{\lambda}_m^n(t)$ denote the Lagrange multipliers associated with equality constraints (8a) and (8b) for buses $n \in \mathcal{N}$ and $m \in \mathcal{N}_n$ in time slot t, respectively. We express (8a) and (8b) in the form of $\mathbf{g}(t) = \mathbf{0}$, where $\mathbf{g}(t) = (\mathbf{W}_m^{n,n}(t) - \mathbf{W}_n^{n,n}(t)$, $\mathbf{W}_m^{n+N,n+N}(t) - \mathbf{W}_n^{n+N,n+N}(t),\, m \in \mathcal{N}_n,\, n \in \mathcal{N})$.

Let i denote the iteration index in the load adjustment phase. Lines 2 and 3 of Algorithm 1 describe the initialization in iteration $i = 1$. The loop involving Lines 4 to 9 describes the interactions between the DNO and ECC $n \in \mathcal{N}$. In Line 5, the DNO broadcasts control signal matrix $\mathbf{S}_n^i(t)$ to ECC n, where $\mathbf{S}_n^i(t)$ is a $2N \times 2N$ *diagonal* matrix. We use the approach in [13, Algorithm 4] to obtain the nonzero entries of the diagonal matrix $\mathbf{S}_n^i(t)$ in iteration i. We develop the augmented Lagrangian of $\boldsymbol{\mathcal{P}}_3(t)$ with parameter $\rho > 0$. We decompose the augmented Lagrangian into N subproblems corresponding to buses in set $\mathcal{N}$. To the objective function of the subproblem associated with bus $n \in \mathcal{N}$, we add a proximal term $\frac{\tau_n}{2}\left\|\mathbf{w}_n(t) - \mathbf{w}_n^i(t)\right\|_2^2$ with weight τ_n, where $\mathbf{w}_n(t) = (\mathbf{W}_n^{m,m}(t), \mathbf{W}_n^{m+N,m+N}(t),\, m \in \mathcal{N}_n \cup \{n\})$ is the vector of nonzero diagonal elements of matrix $\mathbf{W}_n(t)$. The objective function of the subproblem for bus $n \in \mathcal{N}$ in iteration i is obtained as follows:

$$\widetilde{f}_n^{\text{obj},i}(t) = c_n(t) + \text{Tr}\{\mathbf{S}_n^i(t)\,\mathbf{W}_n(t)\} + \frac{\tau_n}{2}\left\|\mathbf{w}_n(t) - \mathbf{w}_n^i(t)\right\|_2^2, \tag{14}$$

where the nonzero elements of matrix $\mathbf{S}_n^i(t)$ are computed as:

$$\mathbf{S}_n^{n,n,i}(t) = \rho \sum_{m\in\mathcal{N}_n}\Big(\mathbf{W}_n^{n,n,i}(t) - \mathbf{W}_m^{n,n,i}(t) + \frac{\lambda_m^{n,i}(t)}{\rho}\Big), \tag{15a}$$

$$\mathbf{S}_n^{n+N,n+N,i}(t) = \rho \sum_{m\in\mathcal{N}_n}\Big(\mathbf{W}_n^{n+N,n+N,i}(t) - \mathbf{W}_m^{n+N,n+N,i}(t) + \frac{\overline{\lambda}_m^{n,i}(t)}{\rho}\Big), \tag{15b}$$

and for $m \in \mathcal{N}_n$, we have

$$\mathbf{S}_n^{m,m,i}(t) = \rho\Big(\mathbf{W}_n^{m,m,i}(t) - \mathbf{W}_m^{m,m,i}(t) - \frac{\lambda_n^{m,i}(t)}{\rho}\Big), \tag{15c}$$

$$\mathbf{S}_n^{m+N,m+N,i}(t) = \rho\Big(\mathbf{W}_n^{m+N,m+N,i}(t) - \mathbf{W}_m^{m+N,m+N,i}(t) - \frac{\overline{\lambda}_n^{m,i}(t)}{\rho}\Big). \tag{15d}$$

To formulate the subproblem of bus n as an SDP, we define an auxiliary variable vector $\boldsymbol{\beta}_n(t) = (\beta_n^m(t),\, \beta_n^{m+N}(t),\, m \in \mathcal{N}_n \cup \{n\})$ associated with the quadratic proximal term in (14). The SDP form of (14) is obtained as follows.

$$\widetilde{f}_n^{\text{obj},i}(t) = c_n(t) + \text{Tr}\{\mathbf{S}_n^i(t)\,\mathbf{W}_n(t)\} + \frac{\tau_n}{2}\sum\nolimits_{m\in\mathcal{N}_n\cup\{n\}}\left(\beta_n^m(t) + \beta_n^{m+N}(t)\right). \tag{16}$$

For $k = m$ and $m+N$, $m \in \mathcal{N}_n \cup \{n\}$, we include the following constraint into the constraint set of subproblem $n \in \mathcal{N}$:

$$\begin{bmatrix} \beta_n^k(t) & \mathbf{w}_n^k(t) - \mathbf{w}_n^{k,i}(t) \\ \mathbf{w}_n^k(t) - \mathbf{w}_n^{k,i}(t) & 1 \end{bmatrix} \succeq 0. \tag{17}$$

Let $\boldsymbol{\psi}_n(t) = (\mathbf{W}_n(t),\, \boldsymbol{\beta}_n(t),\, \theta_n(t),\, \delta_n(t))$ denote the decision variable vector of customer $n \in \mathcal{N}^-$ in time slot $t \in \mathcal{T}$. Let $\boldsymbol{\psi}_N(t) = (\mathbf{W}_N(t),\, \boldsymbol{\beta}_N(t))$ denote the decision variable vector of substation bus N in time slot $t \in \mathcal{T}$. Let $\Psi_n(t)$ for customer $n \in \mathcal{N}^-$ denote the feasible space defined by constraints (3), (4), (6), (7), (10)$-$(11b), and (17). Also let $\Psi_N(t)$ for substation bus N denote the feasible space defined by constraints (5a)$-$(7) and (17). In Line 6, ECC $n \in \mathcal{N}$ solves the following optimization problem to determine the updated variable vector $\boldsymbol{\psi}_n^{i+1}(t)$:

$$\begin{aligned} \boldsymbol{\mathcal{P}}_{3,n}^i(t): \quad & \underset{\boldsymbol{\psi}_n(t)}{\text{minimize}} \quad \widetilde{f}_n^{\text{obj},i}(t) \\ & \text{subject to} \quad \boldsymbol{\psi}_n(t) \in \Psi_n(t), \\ & \qquad\qquad\quad\; \mathbf{W}_n(t) \succeq 0. \end{aligned}$$

Problem $\boldsymbol{\mathcal{P}}_{3,n}^i(t)$ is an SDP and can be solved efficiently. We can show that the solution matrix $\mathbf{W}_n^{i+1}(t)$ to subproblem $\boldsymbol{\mathcal{P}}_{3,n}^i(t)$ is rank-one [15, Sec. IV-C]. It guarantees that $\mathbf{W}_n^{n,n,i+1}(t)$ and $\mathbf{W}_n^{n+N,n+N,i+1}(t)$ are equal to the square of the real and imaginary parts of the bus voltage in time slot t. Moreover, $\mathbf{W}_n^{m,m,i+1}(t)$ and $\mathbf{W}_n^{m+N,m+N,i+1}(t)$ are equal to the square of the real and imaginary parts of the voltage at neighbouring bus $m \in \mathcal{N}_n$. Hence, in Line 6, ECC n can send the updated vector $\mathbf{v}_n^{i+1}(t)$ of the bus voltages to the DNO.

In Line 7, the DNO receives the updated information from the ECCs and determines the updated vector of Lagrange multipliers $\boldsymbol{\lambda}^{i+1}(t)$ as follows:

$$\boldsymbol{\lambda}^{i+1}(t) = \boldsymbol{\lambda}^i(t) - \eta\rho\,\mathbf{g}^{i+1}(t), \tag{18}$$

where $\eta \in (0,1)$ is a damping parameter. The iteration index is updated in Line 8. The stopping criterion for the loop is given in Line 9. Line 10 corresponds to the load scheduling for the next time slot. ECCs $n \in \mathcal{N}^-$ determine $P_n(t+1)$ using (13) in a parallel fashion.

Remark 2 (Convergence Analysis): Suppose that parameters ρ, η, and τ_n satisfy the inequality $\tau_n > \frac{\rho N}{2-\eta}$ for all $n \in \mathcal{N}$. Then, the loop involving Lines 5 to 10 converges to the global optimal solution of problem $\boldsymbol{\mathcal{P}}_3(t)$ [13, Lemma 2.2].

V. Performance Evaluation

In this section, we evaluate the performance of the proposed online demand response algorithm on the IEEE 37-bus distribution feeder. The network data can be found in [17]. A time period $\mathcal{T}$ represents 10 days. Each time slot is 15 minutes. To model the active load pattern for a customer, we use a load pattern from Ontario, Canada power grid database from Apr. 1, 2018 to Apr. 10, 2018 [18]. We scale the load profile for each bus such that the average load demand in each day is within 50% to 150% of the average load in [17]. The lower bound $P_n^{\min}(t)$ for the load demand of customer n is randomly chosen between 65% and 90% of its desirable load in time slot t. The upper bound $P_n^{\max}(t)$ for customer n is randomly chosen between 105% and 110% of its desirable load in time slot t. We assume that the power factor of customer n varies within interval $[0.7,\, 0.95]$ to determine the limits $Q_n^{\min}(t)$ and $Q_n^{\max}(t)$ for the reactive load in time slot t. Coefficients $\omega_n(t)$, $t \in \mathcal{T}$ for the discomfort cost function of customer n are randomly selected from interval $[0.1 \text{ cents/kW}^2,\, 0.5 \text{ cents/kW}^2]$. We consider RTP scheme combined with IBR. For customer n and time slot t, we consider the block rate $\pi_1(t)$ for the portion of load demand $P_n(t)$ smaller than or equal to the threshold l_n^{th} and the block rate $\pi_2(t)$ for the portion of load demand $P_n(t)$ above l_n^{th} [2]. The rates $\pi_1(t)$ and $\pi_2(t)$ during one day are shown in Fig. 1. We set l_n^{th} to the average load of customer n. For the spot market price in time slot t, we set $\pi^{\text{b}}(t) = 2\pi_2(t)$ and $\pi^{\text{s}}(t) = 0.5\pi_1(t)$. We perform simulations using MATLAB/CVX with MOSEK solver in a PC with processor Intel(R) Core(TM) i5-3337U CPU@1.8 GHz.

We first show how Algorithm 1 enables a customer to manage its load demand. For the sake of comparison, we consider the benchmark scenario where the DNO has complete information about the customers' load demand and preferences. The DNO solves problem $\boldsymbol{\mathcal{P}}_2(t)$ to determine the *optimal* scheduled load demands in time slot $t \in \mathcal{T}$. In the scenario with incomplete information, the customers execute Algorithm 1. We set $\rho = 0.1$, $\eta = 0.9$, and $\tau_n > 0.1N$, $n \in \mathcal{N}$ to satisfy the condition in Remark 2. The step size $\alpha(t)$, $t \in \mathcal{T}$ in (13) is set to $\gamma/\sqrt{t}$, where $\gamma = 0.18$ for our case study.

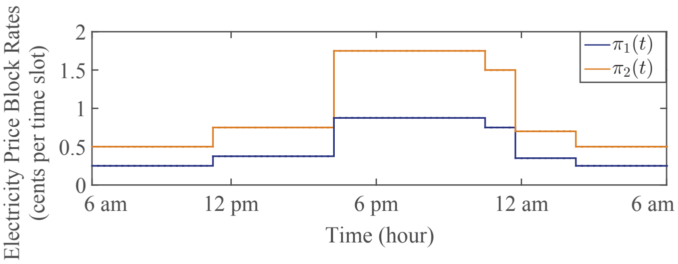

Figure 1. Electricity price block rates for demand response during one day.

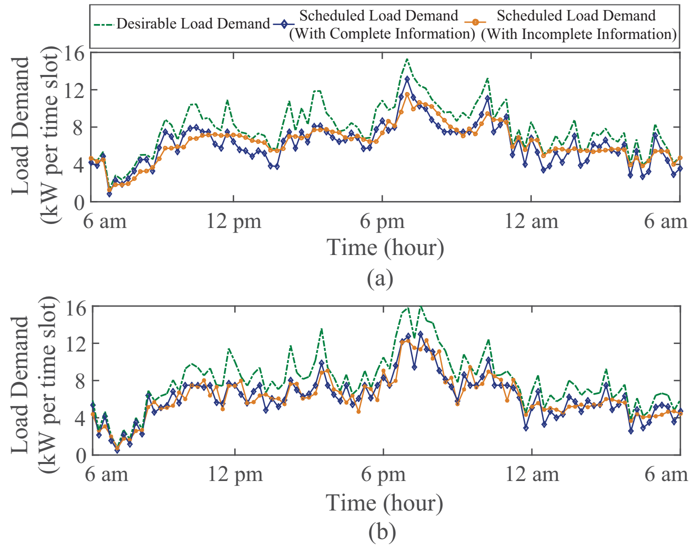

Figure 2. Scheduled load demand profile at bus with index 714 in (a) day 1, and (b) day 10 with full information and incomplete information.

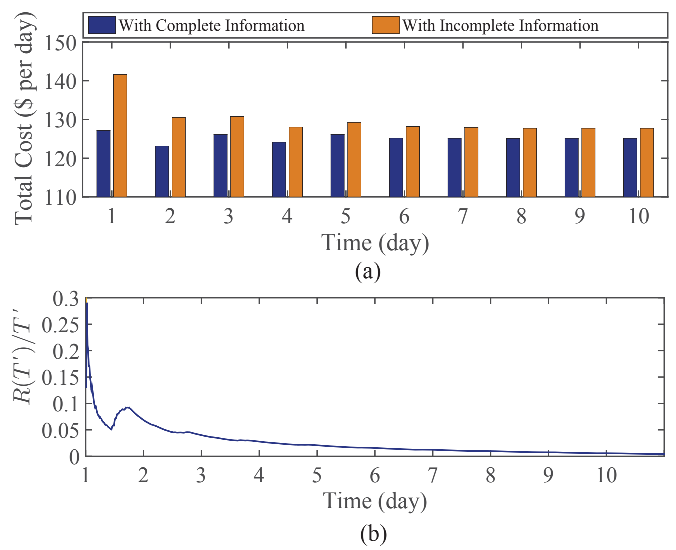

Figure 3. (a) Total daily cost with complete and incomplete information; and (b) the average regret versus time between day 1 and day 10.

Consider customer at bus with index 714 as an example. Figs. 2(a) and (b) show the desirable load demand, the scheduled load demand in the benchmark scenario, and the scheduled load demand with incomplete information in day 1 and day 10, respectively. Our results show that with load scheduling, the customers' load demand during peak hours (between 6 pm and 10 pm) is reduced by about 14% on average. Using Algorithm 1, the scheduled load demand approximates the optimal load profile in the benchmark scenario. As a similarity measure of the load profiles, we consider the average absolute difference between the scheduled load demand with uncertainty and the load demand in the benchmark scenario. Fig. 2(b) shows when compared to day 1, the scheduled load demand in day 10 is a tighter approximation of the optimal scheduled load demand. In particular, the average absolute differences between the scheduled load demand with uncertainty and the load demand in the benchmark scenario for day 1 and day 10 are 2.63 kW and 0.87 kW, respectively.

Next we study the changes in the total daily cost of all customers in the feeder, i.e., the objective values (9) and (12) during day 1 to day 10. Fig. 3(a) shows that the gap between the total cost with uncertainty using Algorithm 1 and the cost in the benchmark scenario is reduced from 10.9% in day 1 to 2.3% in day 10. It implies that Algorithm 1 based on online convex optimization converges to a near optimal solution to problem $\boldsymbol{\mathcal{P}}_2(t)$ and the gap is small. We also evaluate the convergence of the average regret $R(T')/T'$ for $T' = 1, \ldots, 960$ time slots. Fig. 3(b) shows that $R(T')/T'$ tends to zero as T' increases. That is, the average regret grows sublinearly in T'. Ten days (i.e., 960 time slots) are sufficient for Algorithm 1 to converge to a near optimal solution of problem $\boldsymbol{\mathcal{P}}_2(t)$. We emphasize that the average regret $R(T')/T'$ for 36 customers converges to 0.05 in the second day, which implies that one customer incurs a negligible regret. These results demonstrate the potential of Algorithm 1 based on online convex optimization for practical applications.

Finally, we evaluate the average number of iterations in the loop involving Lines 5 to 10 of Algorithm 1 with PJ-ADMM approach, which can be interpreted as an indicator of the number of message exchange between the customers and DNO in the load adjustment phase. For the sake of comparison, we consider the algorithm based on VS-ADMM [13, Algorithm 1], which has been commonly used in the literature (e.g., in [19], [20]). We solve problem $\boldsymbol{\mathcal{P}}_3(t)$ in a centralized fashion for day 1 and Fig. 4(a) shows the scheduled load demand profile for customer at bus with index 714 before and after load adjustment. Consider the load demand at 7 pm, when the customer adjusts its load demand from 8.9 kW to 11.5 kW. Fig. 4(b) depicts the load demand in each iteration of the loop involving Lines 5 to 10 of Algorithm 1. Algorithm 1 with both the PJ-ADMM and VS-ADMM converges to the solution of problem $\boldsymbol{\mathcal{P}}_3(t)$ in a reasonable number of iterations (in 14 and 22 iterations, respectively). Nevertheless, the PJ-ADMM is preferable to VS-ADMM as introducing the splitting variables increases the number of variables and constraints in each subproblem, resulting in a higher average running time for Algorithm 1 per time slot. Table I shows the average running time per time slot for the centralized approach and Algorithm 1 with the PJ-ADMM and VS-ADMM approaches versus the number of buses in the network. When compared with the centralized approach, results show that the proposed

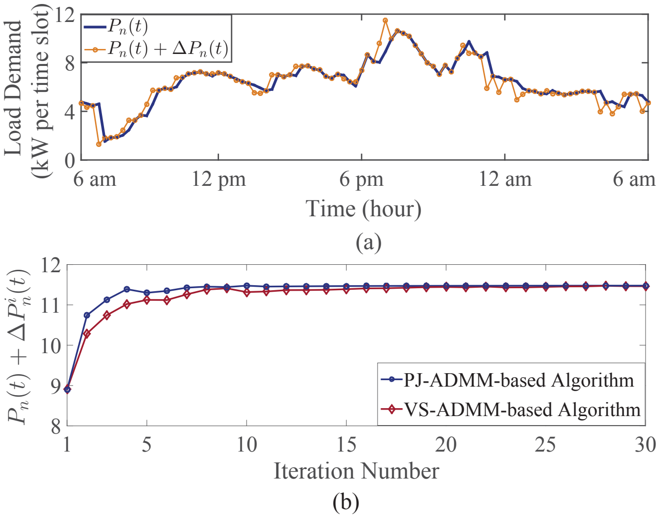

Figure 4. (a) Load demand profile before and after load adjustment during day 1; (b) convergence of load demand in the load adjustment phase at 7 pm using Algorithm 1 based on PJ-ADMM and VS-ADMM.

Table I
THE AVERAGE RUNNING TIME PER TIME SLOT FOR THE CENTRALIZED APPROACH, AND ALGORITHM 1 WITH VS-ADMM AND PJ-ADMM.

	Average Running Time per Time Slot (second)		
Test System	Centralized Approach	Algorithm 1 with VS-ADMM	Algorithm 1 with PJ-ADMM
37-bus	0.42	0.19	0.12
73-bus	1.1	0.68	0.4
361-bus	13.2	1.8	1.1
1801-bus	298.5	14.1	7.6

decentralized algorithm with either PJ-ADMM or VS-ADMM has significantly lower average running time per time slot. The reason is that the buses solve their subproblems in a parallel fashion in Algorithm 1. When compared with the VS-ADMM, Algorithm 1 based on PJ-ADMM has a lower running time *per time slot* due to smaller number of required iterations to converge as well as a lower running time per iteration.

VI. CONCLUSION

In this paper, we addressed the problem of load scheduling in a demand response program. We formulated the DNO's centralized load control problem as an SDP using convex relaxation techniques. We addressed the DNO's uncertainty about the load demand and preferences by using online convex optimization techniques to solve the original centralized load control problem. We applied PJ-ADMM approach to developed an algorithm to solve the stochastic centralized load control problem in a decentralized fashion. Simulation results showed that when compared with the scenario without the demand response, the load demand during peak hours is reduced by 14% on average. Furthermore, the scheduled load profile using the proposed algorithm approximates the optimal load profile in the benchmark scenario without uncertainty, and the approximation is tight. We showed that the average regret grows sublinearly in the number of time slots, and the gap between the total cost with the proposed algorithm and the cost in the benchmark scenario is 2.3%, which is negligible. When compared with the centralized approach and the VS-ADMM technique, our algorithm based on PJ-ADMM has a lower average running time per time slot. For future work, we plan to extend the model by considering the operational model of the customers' electric appliances and the integration of renewable generation in the network.

REFERENCES

[1] J.S. Vardakas, N. Zorba, and C.V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152–178, 1st quarter 2015.

[2] P. Samadi, A.H. Mohsenian-Rad, V.W.S. Wong, and R. Schober, "Real-time pricing for demand response based on stochastic approximation," *IEEE Trans. on Smart Grid*, vol. 5, no. 2, pp. 789–798, Mar. 2014.

[3] Z. Chen, L. Wu, and Y. Fu, "Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization," *IEEE Trans. on Smart Grid*, vol. 3, no. 4, pp. 1822–1831, Dec. 2012.

[4] B.G. Kim, Y. Zhang, M. van der Schaar, and J.W. Lee, "Dynamic pricing and energy consumption scheduling with reinforcement learning," *IEEE Trans. on Smart Grid*, vol. 7, no. 5, pp. 2187–2198, Sept. 2016.

[5] C. Eksin, H. Delic, and A. Ribeiro, "Demand response management in smart grids with heterogeneous consumer preferences," *IEEE Trans. on Smart Grid*, vol. 6, no. 6, pp. 3082–3094, Nov. 2015.

[6] S. Bahrami, V.W.S. Wong, and J. Huang, "An online learning algorithm for demand response in smart grid," *IEEE Trans. on Smart Grid*, vol. 9, no. 5, pp. 4712–4725, Sept. 2018.

[7] S.J. Kim and G.B. Giannakis, "An online convex optimization approach to real-time energy pricing for demand response," *IEEE Trans. on Smart Grid*, vol. 8, no. 6, pp. 2784–2793, Nov. 2017.

[8] W. Shi, N. Li, X. Xie, C. Chu, and R. Gadh, "Optimal residential demand response in distribution networks," *IEEE Journal on Selected Areas in Commun.*, vol. 32, no. 7, pp. 1441–1450, Jun. 2014.

[9] B.V. Solanki, A. Raghurajan, K. Bhattacharya, and C.A. Canizares, "Including smart loads for optimal demand response in integrated energy management systems for isolated microgrids," *IEEE Trans. on Smart Grid*, vol. 8, no. 4, pp. 1739–1748, Jul. 2017.

[10] N. Li, L. Gan, L. Chen, and S.H. Low, "An optimization-based demand response in radial distribution networks," in *Proc. of IEEE Globecom*, Anaheim, CA, Dec. 2012.

[11] E. Hazan, "Introduction to online convex optimization," *Foundations & Trends in Optimization*, vol. 2, no. 3-4, pp. 157–325, Aug. 2016.

[12] S. Shalev-Shwartz, "Online learning and online convex optimization," *Foundations & Trends in Maching Learning*, vol. 4, no. 2, pp. 107–194, Feb. 2012.

[13] W. Deng, M.-J. Lai, Z. Peng, and W. Yin, "Parallel multi-block ADMM with $o(1/k)$ convergence," *Journal of Scientific Computing*, vol. 71, no. 2, pp. 712–736, May 2017.

[14] G. Chen and M. Teboulle, "A proximal-based decomposition method for convex minimization problems," *Mathematical Programming*, vol. 64, no. 1, pp. 81–101, Mar. 1994.

[15] J. Lavaei and S.H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. on Power Systems*, vol. 27, pp. 92–107, Feb. 2012.

[16] L. Song, Y. Xiao, and M. van der Schaar, "Demand side management in smart grids using a repeated game framework," *IEEE Journal on Selected Areas in Commun.*, vol. 32, no. 7, pp. 1412–1424, Jul. 2014.

[17] IEEE PES AMPS DSAS test feeder working group. [Online]. Available: http://sites.ieee.org/pes-testfeeders/resources/

[18] Independent Electricity System Operator (IESO). [Online]. Available: http://www.ieso.ca

[19] T. Erseghe, "Distributed optimal power flow using ADMM," *IEEE Trans. on Power Systems*, vol. 29, no. 5, pp. 2370–2380, Sept. 2014.

[20] S. Magnusson, P.C. Weeraddana, and C. Fischione, "A distributed approach for the optimal power-flow problem based on ADMM and sequential convex approximations," *IEEE Trans. on Control of Network Systems*, vol. 2, no. 3, pp. 238–253, Sept. 2015.