Optimal Day-Ahead Power Procurement with Renewable Energy and Demand Response

Soongeol Kwon, Lewis Ntaimo, and Natarajan Gautam, *Senior Member, IEEE*

***Abstract*—This study proposes the demand-side power procurement problem to optimally reduce consumer's energy cost. The motivation stems from pressing issues on an increase of energy cost in an industrial section. From an energy consumer's perspective, there exists an opportunity to reduce energy cost by adjusting purchase and consumption of energy in response to time-varying electricity price while utilizing renewable energy, which is called demand response. In this case, energy storage can be used to mitigate fluctuation of intermittent renewable supply and volatile electricity price. Although it is anticipated to serve a significant amount of energy consumption from renewable energy and to avoid peak electricity price, variability and uncertainty in power demand, renewable supply, and electricity price, make it challenging to determine an optimal power procurement. The main objective of this study is to suggest a decision-making methodology that enables energy consumers to optimally determine power procurement against time-varying and stochastic electricity price and renewable supply. Specifically, this study formulates an optimal day-ahead power procurement as a two-stage stochastic mixed integer program and proposes a solution approach based on Benders decomposition. The proposed methodology can be successfully applied to energy-intensive industries, such as data centers.**

***Index Terms*—Day-Ahead Power Procurement, Demand Response, Renewable Energy, Energy Storage, Two-stage Stochastic Integer Programming, Benders Decomposition.**

NOMENCLATURE

- Sets and Indices
  - T: Index set of time periods $t{\in}T$
  - ω: Index set of scenarios $\omega{\in}\Omega$
- Deterministic Parameters
  - $\overline{D}_t$: Forecasted power demand at time $t{\in}T$
  - $\overline{R}_t$: Forecasted renewable supply at time $t{\in}T$
  - C_t^{DA}: Day-ahead electricity price at time $t{\in}T$
  - M^{char}, M^{dis}: Charging and discharging rate of storage
  - S^{max}: Maximum level of energy storage
  - η^{char}, η^{dis}: Charging and discharging inefficiency of storage
  - P_t^{loss}: Penalty cost for power loss at time $t{\in}T$
  - L^{max}: Allowed number of time periods for shifting demand
  - TW: Time window to meet shifted power demand
  - ϵ: Maximum fraction of amount of shifted load
- Stochastic Parameters (for each scenario $\omega \in \Omega$)
  - $D_t(\omega)$: Actual power demand at time $t{\in}T$
  - $R_t(\omega)$: Actual renewable supply at time $t{\in}T$
  - $C_t^{RT}(\omega)$: Real-time electricity price at time $t{\in}T$
- First-stage Decision Variables (Day-Ahead Operations)
  - x_t: Day-ahead purchase commitment at time $t{\in}T$
  - u_t: Binary variable indicates whether demand load at time $t{\in}T$ can be shifted by demand response
  - zc_t^{DA}, zd_t^{DA}: Amount to be charged/discharged at time $t{\in}T$
  - s_t^{DA}: Level of storage at the beginning of time $t{\in}T$
- Second-stage Decision Variables (Real-Time Operations)
  - y_t: Real-time electricity purchase at period $t{\in}T$
  - y_t^{loss}: Power loss at period $t{\in}T$
  - $v_{t\ell}$: Amount of load shifted from time $t{\in}T$ will be satisfied at time $\ell{\in}T$ $(t{<}\ell)$
  - w_t: Amount of shifted load at the beginning of time $t{\in}T$
  - zc_t^{RT}, zd_t^{RT}: Amount to be charged/discharged at time $t{\in}T$
  - s_t^{RT}: Level of storage at the beginning of time $t{\in}T$

Authors are with the Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, 77843-3131 USA (e-mail: soongeol@tamu.edu; ntaimo@tamu.edu; gautam@tamu.edu).

I. INTRODUCTION

In recent years, many industries have witnessed a tremendous increase in energy consumption that has resulted in enormous expenses as well as carbon pollution. In 2013, U.S. data centers, one of the today's fastest-growing industries, consumed an estimated 91 billion kilowatt-hours of electricity, which is equivalent to the annual output of 34 large (500-megawatt) coal-fired power plants. Moreover, data centers energy consumption is projected to increase to roughly 140 billion kilowatt-hours annually by 2020, costing \$13 billion annually in electricity bills and emitting nearly 100 million metric tons of carbon pollution per year [1]. For this reason, many energy-intensive industries are striving to reduce energy cost and to have a positive impact on the environment. In this situation, renewable energy is considered as a promising solution for them to be energy-efficient. In other words, industries have an opportunity to utilize renewable energy to partially or fully serve their demand load to curtail expenses for procuring energy. In fact, U.S. renewable electricity has grown up to 13.5% of total electricity, and 7.4% of energy consumption in the industrial sector is currently met by renewable energy [2]. In addition, the amount of industrial energy consumption saved by renewable energy has been continuously increasing, and this trend is expected to continue in the future. In addition, from the energy-consumers perspective, there exists an opportunity for industries to adjust purchase and consumption of energy in response to time-varying price in the energy market. Traditionally, power consumers use electricity with a flat rate offered by utility companies or energy market for their

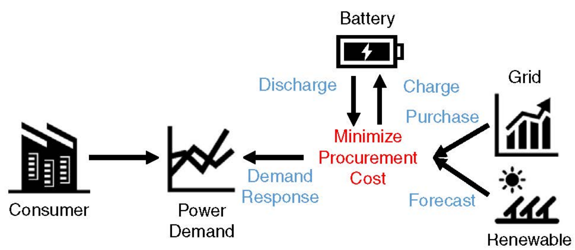


Fig. 1. Demand-side Power Procurement

usage. However, in recent years, it is becoming common for many utilities to offer day-ahead and real-time prices for smart pricing [3], and some independent system operators, such as ERCOT [4] and California ISO [5], have recently allowed consumers to purchase electricity directly form the market while providing price information. Therefore, industries get a chance to procure energy by participating in the market while being fully aware of the time-varying price, and they may have an opportunity to determine the amount of their energy consumption depending on the electricity price. This opportunity is called *demand response*. Moreover, considering an opportunity to use renewable energy, demand response can also be successfully implemented to utilize renewable energy by consuming more renewable energy when it is available. In addition, by applying demand response to energy procurement, energy storage can be used to mitigate fluctuation of intermittent renewable supply and volatile electricity price. Data centers are one of promising application areas for *demand response*, since they have manageable and flexible workloads [6] and are currently using renewable energy to supply power demand by installing on-site renewable generation facility or make contracts with solar or wind farms [7]. Applying demand response in demand-side power system management is studied under the concept of "Virtual Power Plant" [8], [9], and [10].

To realize the aforementioned opportunities, practitioners are strongly encouraged to develop new technologies for planning, design, control, and operations of power systems against variability and uncertainty in renewable energy and electricity price. In other words, since the conventional systems and techniques have not been designed while considering integration of renewable energy and demand response into power system operations, intermittent renewable generation and volatile electricity price challenge power system engineers' decision making. In this context, current research in the power system has been focused on integrating optimization techniques to yield reliable and robust energy generation and procurement. It is anticipated that application of optimization techniques will have a significant impact on planning, design, control, and operations of power systems. For these reasons, this study focuses on developing a decision-making methodology for demand-side power procurement with renewable energy, storage, and demand response using a stochastic optimization technique. Specifically, this study considers a two-stage power procurement composed of day-ahead and real-time procurements. Note that there is a body of literature on demand-side power procurement based on Markov decision process, [11], [12], [13], [14], [15], and Lyapunov optimization [16], [17], [18], [19]. While all of the aforementioned literature focuses on modeling the sequential stochastic control problem and designing optimal policy tailored to real-time power procurement, this study proposes a two-stage stochastic optimization problem tailored to day-ahead power procurement and suggests a solution approach based on Benders decomposition. To the best of our knowledge, the two-stage stochastic optimization approach for day-ahead power procurement problem with renewable energy, storage, and demand response has not been addressed in the literature. Thus, this study would be a good starting point to study demand-side power procurement problem based on the framework of two-stage stochastic program. The rest of this paper is organized as follows: Section II gives a detailed description and assumption of the proposed two-stage power procurement problem and formulates the problem as a mathematical model. Section III introduces an algorithm based on Benders decomposition and suggests strategies designed to improve the algorithm. Section IV analyzes the results obtained by numerical experiments, and Section V ends the paper with concluding remarks and future research directions.

II. Problem Description

A. Scenario and Assumption

Based on scenario considered in this study, consumer's power demand can be met by the following sources: (i) purchase from energy market, (ii) renewable energy, and (iii) discharge from energy storage as depicted in Figure 1. In practice, energy market includes day-ahead and real-time markets that work together as follows:

- *Day-ahead energy market* lets participants commit to buy electricity one day before the operating day to help avoid price volatility.
- *Real-time energy market* allows participants to buy electricity during the course of the operating day to balance mismatch between day-ahead purchase commitment actual demand load.

Considering the operations of energy market, we consider a two-stage framework that consists of day-ahead and real-time power procurement, and propose day-ahead procurement problem. Based on a two-stage stochastic program, the proposed day-ahead power procurement problem is designed so that the first-stage problem determines day-ahead purchase commitment (here-and-now decisions) based on the forecasted demand load and renewable supply, while the second-stage determines the real-time purchase (recourse decisions) to adjust the mismatch between purchase commitments and the actual power demand and renewable supply. We assume that day-ahead electricity price, forecasted power demand and renewable supply are known in the first-stage, but real-time electricity price, actual power demand and renewable supply are time-varying and stochastic. Note that forecasting power demand and renewable supply are out of the scope of this study. In addition, we consider energy storage operations with finite capacity, maximum charging and discharging rates, and inefficiency in charging and discharging. In fact, the frequent cycle of charging or discharging causes the degradation of the energy storage in terms of lifetime and efficiency. However, this study does not consider the degradation since it is assumed to be negligible within one-day operations. Moreover, we implement demand response into the proposed day-ahead power procurement so that consumer assigns time periods in day-ahead and allows demands to be shifted in real-time at assigned time periods, but should be met by the deadline in real time operation. According to the proposed two-stage power procurement framework, based on day-ahead purchase commitment, power loss (i.e. procured power that could not be used to neither serve power demand nor charge storage) might be occurred depending on actual demand load and renewable generations. In our study, we define a penalty cost charged for power loss to ensure that both day-ahead purchase commitment and renewable energy are fully used in real-time operations.

B. Mathematical Model

We formulate the proposed day-ahead power procurement problem as a two-stage stochastic mixed-integer programming (SMIP) problem. The first-stage problem determines the purchase commitment and assign periods for shifting demand based on the day-ahead electricity price, forecasted demand and renewable supply considering storage operation to minimize day-ahead purchase cost and the expected recourse cost caused by the real-time procurement for each possible scenario. In the second-stage, the subproblem is defined to adjust mismatch caused by forecasting errors against actual power demand and renewable supply by purchasing electricity from a real-time market and shifting consumers demand based on operations of energy storage (charging/discharging). Our proposed day-ahead power procurement problem can be formulated as a two-stage SMIP as follows:

$$\text{Min} \sum_{t\in T} C_t^{DA} x_t + \mathbb{E}[f(x,u,\tilde{\omega})] \tag{1}$$

$$\text{s.t. } x_t + zd_t^{DA} - zc_t^{DA} = \overline{D}_t - \overline{R}_t \quad \forall t \in T \tag{2}$$

$$\sum_{t\in T} u_t \le L^{max} \tag{3}$$

$$zc_t^{DA} \le \min\{M^{char}, S^{max} - s_t^{DA}\} \quad \forall t \in T \tag{4}$$

$$zd_t^{DA} \le \min\{M^{dis}, s_t^{DA}\} \quad \forall t \in T \tag{5}$$

$$s_{t+1}^{DA} - s_t^{DA} - \eta^{char} zc_t^{DA} + \frac{1}{\eta^{dis}} zd_t^{DA} = 0 \quad \forall t \in T \tag{6}$$

$$x_t, s_t^{DA}, zc_t^{DA}, zd_t^{DA} \ge 0 \quad \forall t \in T \tag{7}$$

$$u_t \in \{0,1\} \quad \forall t \in T \tag{8}$$

where for each scenario $\omega \in \Omega$

$$f(x,u,\omega) = \sum_{t\in T} \left(C_t^{RT}(\omega) y_t + P_t^{loss} y_t^{loss} \right) \tag{9}$$

$$\text{s.t. } y_t - y_t^{loss} + zd_t^{RT} - zc_t^{RT} + \sum_{\ell=t+1}^{t+TW} v_{t\ell} - \sum_{\ell=t-TW}^{t-1} v_{\ell t} = D_t(\omega) - R_t(\omega) - x_t \quad \forall t \in T \tag{10}$$

$$\sum_{\ell=t+1}^{t+TW} v_{t\ell} \le D_t(\omega) u_t \quad \forall t \in T \tag{11}$$

$$w_{t+1} - w_t - \sum_{\ell=t+1}^{t+TW} v_{t\ell} + \sum_{\ell=t-TW}^{t-1} v_{\ell t} = 0 \quad \forall t \in T \tag{12}$$

$$w_t \le \epsilon \sum_{\ell=1}^{t-1} D_\ell(\omega) \quad \forall t \in T \tag{13}$$

$$zc_t^{RT} \le \min\{M^{char}, S^{max} - s_t^{RT}\} \quad \forall t \in T \tag{14}$$

$$zd_t^{RT} \le \min\{M^{dis}, s_t^{RT}\} \quad \forall t \in T \tag{15}$$

$$s_{t+1}^{RT} - s_t^{RT} - \eta^{char} zc_t^{RT} + \frac{1}{\eta^{dis}} zd_t^{RT} = 0 \quad \forall t \in T \tag{16}$$

$$y_t, y_t^{loss}, v_{\ell t}, w_t, s_t^{RT}, zc_t^{RT}, zd_t^{RT} \ge 0 \quad \forall \ell, t \in T. \tag{17}$$

In the above formulation, the objective function (1) is composed of day-ahead power procurement costs and the expected recourse cost for real-time power procurement in the second-stage corresponding to the one-day operation cycle. Constraint (2) is the power balance equation for day-ahead power procurement plan ensuring that day-ahead purchase commitment is determined so that forecasted power demand is fully satisfied by considering forecasted renewable supply and energy storage operations. Constraint (3) assigns time periods for shifting demand in real time with a maximum allowed number of time periods. Constraints (4)-(6) are for day-ahead storage operations. Constraint (7) are the non-negativity restrictions and constraint (8) gives the binary restrictions on the first-stage decision variables. In the second-stage, the objective function of the subproblem for each scenario is formulated to minimize real-time operations cost, which is composed by real-time purchase cost and penalty cost for power loss as (9). Constraint (10) is the power balance equation for real-time power procurement operation including shifting and serving power demand (for demand response) corresponding to the actual power demand and wind power supply given day-ahead purchase commitment. Note that, power loss may happen when the amount of total power procurement is exceeding the actual power demand and the maximum charging amount. Constraints (11)-(13) are for demand response. Constraint (11) defines a condition that power demand can be shifted only at pre-assigned time periods, and constraint (12) is the balance equation for demand shifting under demand response. We define the quality of usage constraint as (13) so that the fraction of the amount of shifted demand (but not yet served) to the total amount of power demand cannot be exceeded a pre-agreed level. Constraints (14)-(16) are for real-time energy storage operations, and constraint (17) are the non-negativity restrictions on the second-stage decision variables.

Note that "$\min\{\}$" function used in constraints (4), (5), (14), and (15) can simply be linearized by two separate constraints. For example, constraint (4) is equivalent to $zc_t^{DT} \le M^{char}$ and $zc_t^{DT} + s_t^{DT} \le S^{max}$ $\forall t \in T$. We would like to emphasize that in the two-stage SMIP formulation, only the fist-stage problem includes integer variables and the subproblem is formulated without any integer variables, and thus, the proposed two-stage SMIP problem has continuous recourse. In addition, the two-stage SMIP has relatively complete recourse [20] such that every solution obtained by solving the master problem always results in a feasible subproblem.

III. Solution Approach

As described in Section II-B, our proposed day-ahead procurement problem is formulated as a two-stage SMIP problem with continuous recourse where only the master problem includes binary decision variables. Note that the two-stage SMIP problem can be modelled as a deterministic equivalent problem (DEP) that is formulated as a large mixed integer programming problem with a finite number of scenarios. In general, solving a DEP of the two-stage SMIP problem is inefficient with a large number of scenarios, and in this case, decomposition techniques can be used to solve the problem efficiently. Specifically, for the continuous recourse, the L-shaped algorithm [21] and the multicut L-shaped algorithm [22] can be used to solve the two-stage stochastic programming problem based on Benders decomposition [23]. The main idea of the L-shaped algorithm and the multicut algorithm is to solve the decomposed master and subproblems separately by approximating a recourse function by adding Benders cuts within the course of solving the master problem. However, both algorithms based on Benders decomposition may lead the slow convergence to get an optimal solution depending on problem structure as well as scenario data. For these reasons, there has been a body of literature that the proposed techniques to generate stronger Benders cuts that accelerate the convergence of the algorithm [24], [25], [26], and [27].

In this study, we propose cut generation strategy (Section III-A) that introduces valid inequalities to generate stronger Benders cuts and define valid optimality cuts that can be added to the master problem in addition to Benders cuts during the course of the multicut L-shaped algorithm. In addition to cut generation strategy, we suggest cut aggregation strategy (Section III-B) based on the relative trade-off between the single cut and multicut methods [28], while investigating the optimal aggregation level of Benders cuts. Let us redefine decision variables used in formulation (1)-(17) as a set of vectors, $\mathbf{x}, \mathbf{u}$ and $\mathbf{y}$, such that $\mathbf{x}$ denotes vectors of continuous variables (i.e. $x_t, s_t^{DA}, zc_t^{DA}, zd_t^{DA}$ for all $t \in T$) and $\mathbf{u}$ denotes binary variables (i.e. u_t for all $t \in T$) in the first stage, and $\mathbf{y}$ denotes vectors of continuous variables in the second stage (i.e. $y_t, y_t^{loss}, v_{\ell t}, w_t, s_t^{RT}, zc_t^{RT}, zd_t^{RT}$ for all $\ell, t \in T$). Then, with suitable matrices, $\mathbf{A}, \mathbf{D}, \mathbf{W}, \mathbf{T}, \mathbf{H}(\omega)$, and vectors, $\mathbf{c}, \mathbf{b}, \mathbf{e}, \mathbf{q}(\omega), \mathbf{r}(\omega)$, our proposed two-stage day-ahead power procurement problem (1)-(17) can be defined as follows,

$$\text{Min } \mathbf{c}^\top \mathbf{x} + \mathbb{E}[f(\mathbf{x}, \mathbf{u}, \tilde{\omega})] \quad (18)$$

$$\text{s.t. } \mathbf{A}\mathbf{x} \le \mathbf{b} \quad (19)$$

$$\mathbf{D}\mathbf{u} \le \mathbf{e} \quad (20)$$

$$\mathbf{x} \ge 0, \mathbf{u} \in \{0,1\}^n \quad (21)$$

where for each scenario $\omega \in \Omega$

$$f(\mathbf{x}, \mathbf{u}, \omega) = \text{Min } \mathbf{q}(\omega)^\top \mathbf{y} \quad (22)$$

$$\text{s.t. } \mathbf{W}\mathbf{y} \le \mathbf{r}(\omega) - \mathbf{T}\mathbf{x} - \mathbf{H}(\omega)\mathbf{u} \quad (23)$$

$$\mathbf{y} \ge 0, \quad (24)$$

where $\tilde{\omega}$ is a multivariate random variable defined on a probability space with outcome scenarios $\omega \in \Omega$. Let s denote index of scenarios such that $s = 1, \ldots, S$ ($S = |\Omega| < \infty$) and p_s denote the probability of occurrence for each scenario, then based on the multicut L-shaped algorithm, we solve the following master problem iteratively,

$$\text{Min } \mathbf{c}^\top \mathbf{x} + \sum_{s=1}^{S} p_s \eta_s \quad (25)$$

$$\text{s.t. } \mathbf{A}\mathbf{x} \le \mathbf{b} \quad (26)$$

$$\mathbf{D}\mathbf{u} \le \mathbf{e} \quad (27)$$

$$\beta_{\mathbf{t(s)}}^\top \mathbf{x} + \gamma_{\mathbf{t(s)}}^\top \mathbf{u} + \eta_s \ge \alpha_{t(s)} \quad t(s) = 1, ..., u(s), \; s = 1, ..., S \quad (28)$$

$$x \ge 0, u \in \{0,1\}^n, \eta_s \text{ free}, \quad s = 1, \ldots, S, \quad (29)$$

where $t(s)$ is an index of Benders optimality cuts generated by solving the sub problem with scenarios $s \in S$ and $u(s)$ is the number of Benders optimality cuts added to the master problem during the course of algorithm. Note that Benders optimality cuts (28) are generated by solving the following dual subproblem for each possible scenario,

$$f_s(x) = \text{Max } \pi_{\mathbf{s}}^\top (\mathbf{r_s} - \mathbf{T}\mathbf{x} - \mathbf{H}\mathbf{u}) \quad (30)$$

$$\text{s.t. } \pi_{\mathbf{s}}^\top \mathbf{W} \le \mathbf{q} \quad (31)$$

$$\pi_{\mathbf{s}} \le 0, \quad (32)$$

with $\alpha_s = p_s(\pi_{\mathbf{s}}^*)\mathbf{r_s}$, $\beta_{\mathbf{s}}^\top = \mathbf{p_s}(\pi_{\mathbf{s}}^*)^\top \mathbf{T}$, and $\gamma_{\mathbf{s}}^\top = \mathbf{p_s}(\pi_{\mathbf{s}}^*)^\top \mathbf{H_s}$ with $\pi_{\mathbf{s}}^*(\mathbf{x})$ an optimal solution of the dual subproblem. We would like to emphasize that our proposed two-stage SMIP problem has relatively complete recourse, and thus, only optimality cuts (28) are generated and added to the master problem based on the multicut L-shaped algorithm.

In this study, we implement the Benders decomposition based on single search tree referred to as "Branch-and-Benders-cut" (B&BC) algorithm [29] by using the lazy constraints pool provided by CPLEX Concert Technology (IBM ILOG CPLEX [30]). The main advantage of B&BC is that Benders cuts can be added to the master problem during the course of branch-and-cut algorithm (i.e. single search tree) rather than re-solving the master problem as a new problem at each iteration when Benders cuts are generated and added by solving the subproblems. This can expedite solving the master program. However, there also might be disadvantages of using the lazy constraints pool due to the following reasons. During the course of branch-and-cut algorithm, Benders cuts are generated and added each time when the integer (and fractional) solutions are encountered, and the algorithm check the lazy constraint pool for the fractional solution. This might take longer computational time than the classical implementation of Benders decomposition. Therefore, we conducted preliminary experiments, and results showed that the B&BC algorithm using the lazy constraints pool outperforms the classical implementation for solving the proposed problem. Hence, we implement the multicut L-shaped algorithm by using the lazy constraints pool. The details of our proposed cut generation and aggregation strategies are described in the following Sections III-A and III-B, respectively.

A. *Cut Generation Strategy*

1) Valid Inequalities: The key idea for improving performance of the multicut L-shaped algorithm is to generate stronger Benders cuts so that the solution space of the master problem can be significantly restricted. For the purpose of generating stronger Benders cuts, we propose the following valid inequalities (33) and (34). By adding valid inequalities (33) and (34), and projecting them into the solution space of the subproblem, the additional effects of the master problem's solution can be reflected in the subproblem's solution, and thus, stronger Benders cuts can be generated and added.

$$\sum_{\ell=t+1}^{t+TW} v_{t\ell} \leq \epsilon\Big(\sum_{\ell=1}^{t-1} D_\ell(\omega)\Big) - w_t + \sum_{\ell=t-TW}^{t-1} v_{\ell t} + \epsilon D_t(\omega) u_t \quad \forall t \in T \tag{33}$$

$$w_{t+1} \leq \epsilon \sum_{\ell=1}^{t-1} D_\ell(\omega) + \epsilon D_t(\omega) u_t \quad \forall t \in T. \tag{34}$$

For demand response, the two set of valid inequalities ensure that the amount of shifted demand at time period $t \in T$ does not exceed the actual allowable limit that is restricted by the quality of usage constraint (13). We have the following propositions and proofs to show the validity of the proposed inequalities (33).

Proposition 1: The following inequality,

$$\sum_{\ell=t+1}^{t+TW} v_{t\ell} = \epsilon\left(\sum_{\ell=1}^{t-1} D_\ell(\omega)\right) - w_t + \sum_{\ell=t-TW}^{t-1} v_{\ell t} + \epsilon D_t(\omega) \quad \forall t \in T, \tag{35}$$

is valid for problem (1)-(17).

Proof: By plugging (13) into (12), we can show that

$$\begin{aligned}
&\sum_{\ell=t+1}^{t+TW} v_{t\ell} = w_{t+1} - w_t + \sum_{\ell=t-TW}^{t-1} v_{\ell t} \\
&\leq \epsilon\left(\sum_{\ell=1}^{t} D_\ell(\omega)\right) - w_t + \sum_{\ell=t-TW}^{t-1} v_{\ell t} \\
&= \epsilon\left(\sum_{\ell=1}^{t-1} D_\ell(\omega)\right) - w_t + \sum_{\ell=t-TW}^{t-1} v_{\ell t} + \epsilon D_t(\omega) \quad \forall t \in T,
\end{aligned}$$

which proves the result. ■

Proposition 2: The inequality,

$$\sum_{\ell=t+1}^{t+TW} v_{t\ell} \leq \epsilon\Big(\sum_{\ell=1}^{t-1} D_\ell(\omega)\Big) - w_t + \sum_{\ell=t-TW}^{t-1} v_{\ell t} + \epsilon D_t(\omega) u_t \quad \forall t \in T, \tag{36}$$

is valid for problem (1)-(17).

Proof: Considering the value of decision variable u_t for all $t \in T$, we have the following two cases:

- Case 1: If $u_t = 0$, then

$$\sum_{\ell=t+1}^{t+TW} v_{t\ell} = 0, \tag{37}$$

  by constraint (11). Now, for $u_t = 0$, we have the inequality (36) as,

$$\sum_{\ell=t+1}^{t+TW} v_{t\ell} \leq \epsilon\Big(\sum_{\ell=1}^{t-1} D_\ell(\omega)\Big) - w_t + \sum_{\ell=t-TW}^{t-1} v_{\ell t}. \tag{38}$$

  Note that the RHS of (38) would be positive by constraint (13), and $v_{\ell t} \geq 0$ for all $\ell, t \in T$. Hence, inequality (36) is valid for $u_t = 0$ for all $t \in T$.
- Case 2: If $u_t = 1$, then inequality (36) is equivalent to inequality (35) which is valid for the proposed problem (1)-(17). Hence, inequality (36) is valid for $u_t = 1$ for all $t \in T$.

■

In addition, the following proposition shows the validity of inequality (34).

Proposition 3: The inequality,

$$w_{t+1} \leq \epsilon \sum_{\ell=1}^{t-1} D_\ell(\omega) + \epsilon D_t(\omega) u_t \quad \forall t \in T, \tag{39}$$

is valid for problem (1)-(17).

Proof: By plugging (12) into valid inequality (36), we obtain inequality (39). ■

Note that our proposed valid inequalities (36) and (39) are equivalent because of equation (12), however, their contribution to improve the performance of Benders decomposition might be different. In Section IV, we will compare performance improvement by applying each of valid inequalities (36) and (39).

2) Valid Optimality Cuts: In addition to Benders optimality cuts, we introduce a set of valid optimality cuts designed to be added to approximate the recourse function in the first-stage problem of the two-stage SMIP problem. Note that Laporte and Louveaux [31] developed the optimality cut for approximating the expected recourse function with the binary first-stage problem (i.e. the first-stage problem includes only binary decision variables). In this study, we extend their optimality cut so that it can be used to approximate the expected continuous recourse $F(\mathbf{x}, \mathbf{u}) = \mathbb{E}[f(\mathbf{x}, \mathbf{u}, \tilde{\omega})]$ for the mixed-binary first-stage problem where x is continuous and u is binary decision variables. To introduce the proposed valid optimality cuts, we assume that a lower bound L on $\mathbb{E}[f(\mathbf{x}, \mathbf{u}, \tilde{\omega})]$ is known, that is,

$$L \leq \min_{\mathbf{x},\mathbf{u}} \{\mathbb{E}[f(\mathbf{x},\mathbf{u},\tilde{\omega})] \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{D}\mathbf{u} \leq \mathbf{e}, \mathbf{x} \geq \mathbf{0}, \mathbf{u} \in \{0,1\}^n\}.$$

Let $\mathbf{x^k}$ and $\mathbf{u^k}$ denote the master problem's solution at kth iteration during the course of the multicut L-shaped algorithm, then we have recourse function for $\mathbf{x^k}$ and $\mathbf{u^k}$ as,

$$F(\mathbf{x^k}, \mathbf{u^k}) = \mathbb{E}[f(\mathbf{x^k}, \mathbf{u^k}, \tilde{\omega})],$$

and define the set S^k for kth binary decision variables as,

$$S^k = \{t \mid u_t^k = 1\}.$$

We summarize our proposed optimality cut in the following theorem.

Theorem 1: The following cut is a valid cut for $F(x, u)$:

$$\eta \geq (F(\mathbf{x^k}, \mathbf{u^k}) - L)\left(\sum_{t\in S^k} u_t - \sum_{t\notin S^k} u_t - |S^k| + 1\right) + L - \mathbf{c}^\top(\mathbf{x} - \mathbf{x^k}). \tag{40}$$

Proof:

1) If $u = u^k$, $(\sum_{t\in S^k} u_t - \sum_{t\notin S^k} u_t - |S^k| + 1) = 1$.

- If $\mathbf{x} = \mathbf{x^k}$, then the cut $\eta \geq F(\mathbf{x^k}, \mathbf{u^k})$ is tight (i.e. active).

- If $\mathbf{x}{\neq}\mathbf{x^k}$, then the cut $\eta{\geq}F(\mathbf{x^k},\mathbf{u^k}){-}c^\top(\mathbf{x}{-}\mathbf{x^k})$ is valid for $\mathbf{x}{\in}\{\mathbf{Ax}{\leq}\mathbf{b},\mathbf{x}{\in}\mathbb{R}^\mathbf{n}_+,\mathbf{x}{\neq}\mathbf{x^k}\}$. In addition, for incumbent solution $\mathbf{x^k}$ and $\mathbf{u^k}$ obtained during the course of branch-and-cut algorithm, the following inequality is valid,

$$\mathbf{c}^\top\mathbf{x}{+}F(\mathbf{x},\mathbf{u^k}){\geq}c^\top\mathbf{x^k}{+}F(\mathbf{x^k},\mathbf{u^k}), \tag{41}$$

for all $\mathbf{x}{\in}\{\mathbf{Ax}{\leq}\mathbf{b},\mathbf{x}{\geq}\mathbf{0}\}$. Note that $\mathbf{c}^\top\mathbf{x}{+}F(\mathbf{x},\mathbf{u})$ represents objective function value of overall problem (i.e. including the first and second-stage objective function value) decision variable u is fixed as $u{=}u^k$ in both left-hand-side and right-hand-side of the above inequality (41). Now we have the following inequality:

$$\eta{\geq}F(\mathbf{x},\mathbf{u^k}){\geq}F(\mathbf{x^k},\mathbf{u^k}){-}c^\top(\mathbf{x}{-}\mathbf{x^k}). \tag{42}$$

This shows that the cut $\eta{\geq}F(\mathbf{x^k},\mathbf{u^k}){-}c^\top(\mathbf{x}{-}\mathbf{x^k})$ is valid.

2) If $\mathbf{u}{\neq}\mathbf{u^k}$, then $(\sum_{t\in S^k}u_t{-}\sum_{t\notin S^k}u_t{-}|S^k|{+}1){\leq}0$. And let $M{=}(F(\mathbf{x^k},\mathbf{u^k}){-}L)\big(\sum_{t\in S^k}u_t{-}\sum_{t\notin S^k}u_t{-}|S^k|{+}1\big)$, then $M{\leq}0$ since $F(x^k,u^k){\geq}L$.

- If $\mathbf{x}{=}\mathbf{x^k}$, then the cut is $\eta{\geq}M{+}L$ and it must be valid.
- If $\mathbf{x}{\neq}\mathbf{x^k}$, then the cut $\eta{\geq}L{+}M{-}\mathbf{c}^\top(\mathbf{x}{-}\mathbf{x^k})$ is valid since,

$$\begin{aligned}\eta{\geq}F(\mathbf{x},\mathbf{u^k})&{\geq}M{+}L{-}F(\mathbf{x^k},\mathbf{u^k}){+}F(\mathbf{x},\mathbf{u^k})\\&{\geq}M{+}L{-}\mathbf{c}^\top(\mathbf{x}{-}\mathbf{x^k}),\end{aligned}$$

based on inequality (41).

■

We would like to emphasize that optimality cut (40) is weak, therefore it should be used together with Benders cuts to improve the performance. In addition, optimality cut (40) can be implemented into the cut aggregation scheme based on the multicut L-shaped algorithm. Let $j \in J$ be the index of cut aggregate and define the expected recourse function for the subset of scenarios corresponding to each cut aggregate as,

$$F_j(\mathbf{x},\mathbf{u}) = \mathbb{E}[f(\mathbf{x},\mathbf{u},\tilde{\omega}_j)]. \tag{43}$$

Assuming that a lower bound L_j is known, that is,

$$L_j{\leq}\mathrm{Min}_{x,u}\{\mathbb{E}[f(\mathbf{x},\mathbf{u},\tilde{\omega}_j)]|\mathbf{Ax}{\leq}\mathbf{b},\mathbf{D}{\leq}\mathbf{e},\mathbf{x}{\geq}\mathbf{0},\mathbf{u}{\in}\{0,1\}^n\}. \tag{44}$$

Then, the following cut is a valid optimality cut for $F_j(x,u)$:

$$\eta_j \geq (F(\mathbf{x^k},\mathbf{u^k})_j - L_j)\left(\sum_{t\in S^k}u_t - \sum_{t\notin S^k}u_t - |S^k| + 1\right) + L_j - \mathbf{c}^\top(\mathbf{x}-\mathbf{x^k}). \tag{45}$$

Optimality cut (45) can be added to the master problem together with Benders optimality cuts for the approximated recourse function of each cut aggregate, η_j. To implement optimality cuts (45), lower bound L_j can be determined by solving the following relaxed problem:

$$\begin{aligned}&L_j = \mathrm{Min}\sum_{\omega\in\Omega_j}p(\omega)\mathbf{q}(\omega)^\top\mathbf{y}(\omega)\\&\text{s.t. } \mathbf{Wy}(\omega) \leq \mathbf{r}(\omega) - \mathbf{T}(\omega)\mathbf{x} - \mathbf{H}(\omega)\mathbf{u} \quad \forall\omega\in\Omega\\&\mathbf{Ax}\leq\mathbf{b}\\&\mathbf{x}\geq 0, \mathbf{u}\in[0,1]^n, y(\omega)\geq 0 \quad \forall\omega\in\Omega,\end{aligned} \tag{46}$$

where Ω_j represents subset of scenarios corresponding to cut aggregate $j \in J$. Note that problem (46) is relatively easy to solve with relaxed binary decision variables.

B. Cut Aggregation Strategy

The motivation of cut aggregation stems from the relative advantages of the L-shaped (single cut) algorithm and the multicut L-shaped algorithm. In general, the multicut L-shaped algorithm has less major iterations via passing more information by allowing for cuts up to the number of scenarios than the L-shaped algorithm, however, solving the master problem requires more computation time. On the other hand, when we aggregate cuts and less number of optimality cuts are added to the master problem, the algorithm may have more major iterations due to loss of information caused by aggregation. However, the master problem can be solved easier than when the multicut L-shaped algorithm is used. Based on the trade-off in terms of computational time, authors of [28] suggested an adaptive optimality multicut method that dynamically adjusts the level of aggregation of the optimality cuts in the master problem during the course of the algorithm. The numerical results of [28] show that the optimal computational time is achieved on some middle level of aggregation, but this level is not known a priori and depends on problem structure. In a similar fashion, we try to investigate an appropriate aggregation levels based on the trade-off of algorithm performance in terms of computational time.

In this study, we propose a cut aggregation strategy that assigns Benders optimality cuts to be aggregated for the given aggregation level during the course of the algorithm. The fundamental idea of our suggested strategy is to aggregate Benders cuts while minimizing loss of information caused by cut aggregation. This can be accomplished by aggregating Benders optimality cuts obtained from the subproblem defined by "similar" scenario data. Each scenario consists of three-dimensional vectors, power demand, renewable supply, and electricity prices, respectively. These vectors show time-varying patterns across 24 hours periods corresponding to one-day time horizon. We would like to emphasize that relations among power demand, renewable supply, and electricity prices have a significant impact on the solution of the subproblem due to the problem structure. For example, if there exists a negative correlation between power demand and electricity price, then the optimal solution of subproblem is determined so that storage is charged and discharged more frequently as well as more power demand is shifted to minimize expense. In this context, we characterize the structure of each scenario data using pairwise correlations between power demand, renewable supply, and electricity prices and measure similarity of scenario data based on those correlations. For example, correlation between series of $D_t(\omega)$ and $C_t^{RT}(\omega)$ across time periods $t \in T$ for each scenario $\omega \in \Omega$, $\rho_{DC}(\omega)$, can be computed as follows:

$$\rho_{DC}(\omega){=}\frac{\sum_{t=1}^{24}(D_t(\omega){-}\overline{D}(\omega))(C_t(\omega){-}\overline{C}(\omega))}{\sqrt{\sum_{t=1}^{24}(D_t(\omega){-}\overline{D}(\omega))^2\sum_{t=1}^{24}(C_t(\omega){-}\overline{C}(\omega))^2}}, \tag{47}$$

where $\overline{D}(\omega)$ is the average power demand and $\overline{C}(\omega)$ is the average electricity prices for each scenario $\omega \in \Omega$. Likewise, we can determine pairwise correlation between power demand and renewable supply, $\rho_{DR}(\omega)$, and renewable supply and

TABLE I
PARAMETER SETTING

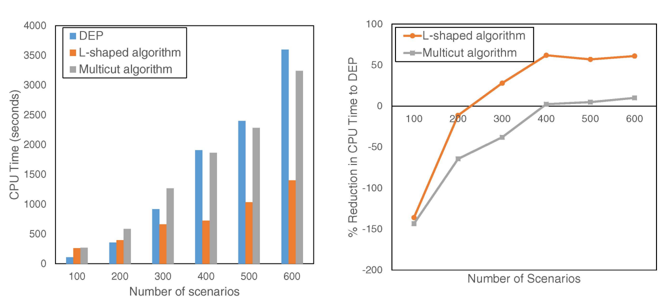
Energy Storage	S^{max}	$S^{max} = \frac{E[D_t(\omega)]}{2}$
	M^{char}, M^{dis}	$M^{char} = M^{dis} = \frac{E[D_t(\omega)]}{4}$
	η^{char}, η^{dis}	$\eta^{char} = \eta^{dis} = 0.9$
Demand Response	L^{max}	$L^{max} = 4$
	TW	$TW = 4$
	ϵ	$\epsilon = 0.05$
Forecasted Demand	$\overline{D}_t$	$\overline{D}_t = E[D_t(\omega)]\ \forall t \in T$
Forcasted Renewable	$\overline{R}_t$	$\overline{R}_t = E[R_t(\omega)]\ \forall t \in T$
Penalty Cost	P_t^{loss}	$P_t^{loss} = C_t^{DA}$

electricity price, $\rho_{RC}(\omega)$, for each scenario $\omega \in \Omega$.

To implement our idea for cut aggregation, we cluster scenarios using k-means clustering algorithm based on pairwise correlation values of each scenario. As mentioned above, three pairwise correlations are computed for each scenario, and thus we can cluster scenarios using k-means up to 3-dimensions based on selection of those pairwise correlations. For example, for 1-dimensional clustering, we can pick one of $\rho_{DC}(\omega)$, $\rho_{DR}(\omega)$, and $\rho_{RC}(\omega)$), and for 2-dimensional clustering, we can choose combination of two correlations, $\rho_{DC}(\omega)$ and $\rho_{DR}(\omega)$, $\rho_{DR}(\omega)$ and $\rho_{RC}(\omega)$), $\rho_{DC}(\omega)$ and $\rho_{RC}(\omega)$). Note that original k-means clustering algorithm does not guarantee to generate equal-sized cluster, therefore we implemented k-means algorithm by using an open source data mining software [32] so that it yields equal-sized k clusters (i.e. each cluster consists of n/k where n is the number of scenarios) for balanced aggregation. Once the scenarios are clustered, Benders optimality cuts generated by solving the subproblem for scenarios in the same cluster will be aggregated and added to the master problem.

IV. NUMERICAL EXPERIMENTS

For numerical experiments, we investigated the performance of the proposed algorithm using scenarios generated by the probabilistic model introduced by Kwon et al. [15]. Through analyzing real historical data, it is evident that power demand, wind generation, and electricity price are time-varying and stochastic, however, it is also reasonable to assume that there exist daily cyclic patterns in power demand and electricity price. In other words, there are deterministic and stochastic variabilities in power demand, renewable generation, and electricity price. Kwon et al. [15] proposed the probabilistic model using on Markov chain to adequately capture the both deterministic and stochastic variabilities. To generate a set of representative scenarios that adequately captures the both deterministic and stochastic variabilities, we train the probabilistic model by using real historical data obtained from Pennsylvania New Jersey Maryland (PJM) interconnection [33], and randomly generate scenarios using Monte Carlo simulation for replications. Note that the proposed probabilistic model using discrete time Markov chains on discrete state spaces, and we mapped random variables to 20 discretized states ($M = 20$) so that power demand, wind supply, and electricity price have 20 different values for each time period. Once we generate a pool of 100,000 scenarios, we obtain 10 replications for each sample size, 100, 200, 300, 400, 500, and 600 by selecting instances randomly from scenario pool. In addition, in terms of parameters in the proposed problem, we set parameters' value as described in Table I. All the experiments were conducted on an Intel Core i7-3740 2.70GHz processor with 16GB memory. We summarize various numerical results for performance evaluation in the following Sections IV-B and IV-C.



(a) Computational Time (b) % Reduction against DEP

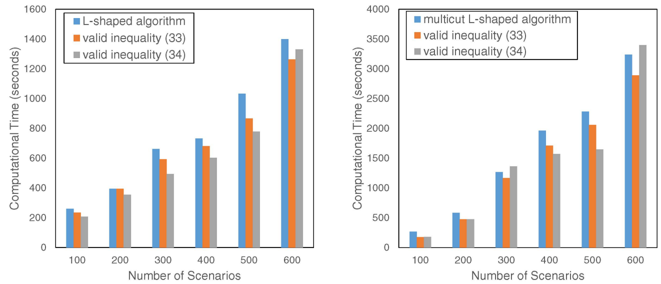
Fig. 2. Performance Comparison: DEP vs L-shaped vs Multicut L-shaped

A. Value of Stochastic Solution

Before analyzing performance of our proposed approach, we would like to discuss the value of stochastic solution (VSS) [34]. In general, stochastic programs are computationally difficult to solve, and thus, practitioners may want to formulate the real-world problem as simpler versions, e.g. deterministic optimization problem by using nominal values as you mentioned. The solution obtained from the simpler versions of problems may provide nearly optimal solutions, however, sometimes yield totally inaccurate solution due to the lack of considering uncertainties. In this case, we can measure the value of the stochastic program by using VSS which is the possible cost reduction obtained by solving the stochastic optimization problem. When no further in formation about the future is available, VSS becomes more practically relevant [20]. We conducted preliminary experiments to analyze the quantity of VSS and we checked that about 10-15% of procurement cost can be reduced by solving stochastic optimization problem instead of solving deterministic optimization problem using the expected value.

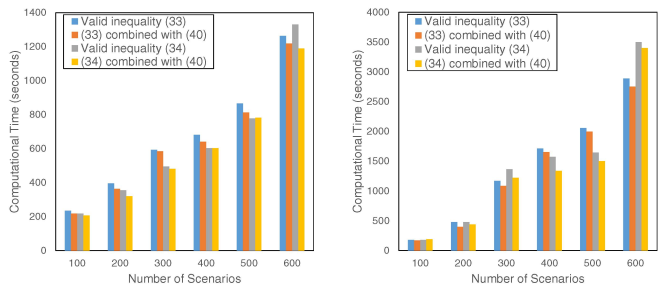
B. Performance Analysis of Cut Generation Strategy

We analyze the performance of the cut generation strategy introduced in Section III-A. Based on the L-shaped and multicut L-shaped algorithms, we solve the problems for various sizes of scenarios by applying (i) the proposed valid inequalities (33) and (34), (ii) the proposed valid optimality cut (45), and combination of both (i) and (ii). We compare the performance of the L-shaped and the multicut L-shaped algorithm against the DEP. As depicted in Figure 2, as the size of scenarios increases, the L-shaped and the multicut L-shaped algorithms outperform the DEP. Moreover, we can find that the L-shaped algorithm shows better performance than the multicut L-shaped algorithm, and this indicates that the performance of the multicut L-shaped algorithm can be improved with cut aggregation strategy as we conjectured. We will investigate the algorithm performance for the different



(a) L-shaped algorithm (b) Multicut algorithm

Fig. 3. Performance analysis of the proposed valid inequalities (33) and (34)



(a) L-shaped algorithm (b) Multicut algorithm

Fig. 4. Performance analysis of the proposed valid inequalities (33) and (34) combined with optimality cut (40)

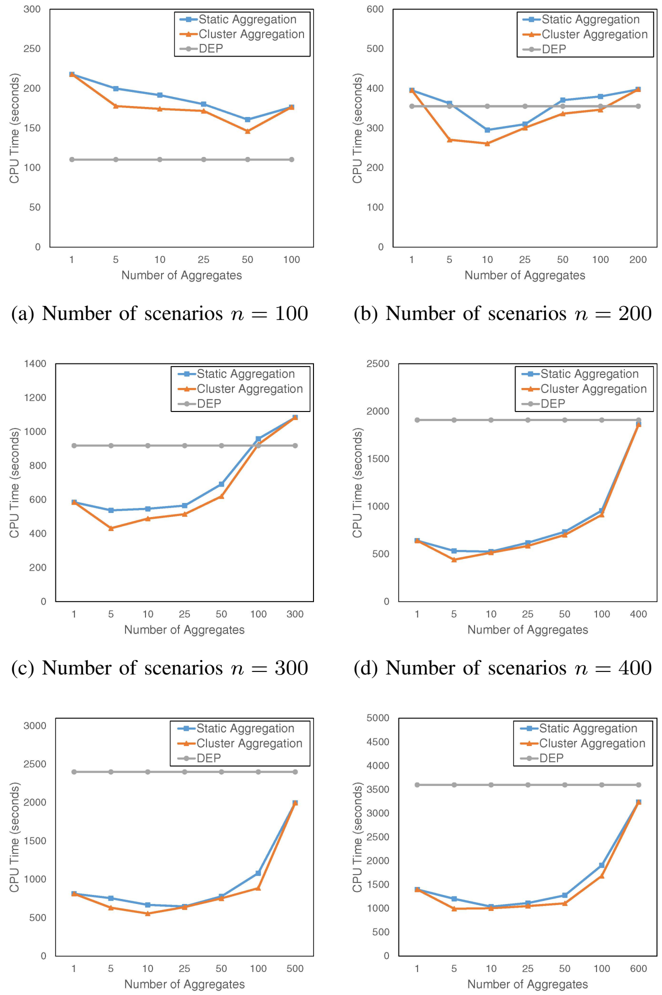
levels of aggregation in Section IV-C. Next, we investigate the performance improvement by the proposed valid inequalities for both the L-shaped and the multicut L-shaped algorithm in terms of computational time. As depicted in Figures 3 and 4, both the L-shaped and the multicut L-shaped algorithm are improved by applying our proposed valid inequalities. We found that valid inequality 34 performs better than valid inequality (33) in many instances, however, it does not dominate. In addition, we analyzed the performance improvement by applying the proposed valid optimality cut (45) combined with the valid inequalities (33) and (34). As depicted in Figure 4, we can see the most improved performance when applying both the proposed valid inequalities and valid optimality cut simultaneously during the course of the algorithm. Based on these findings, we use the proposed valid inequalities and valid optimality cuts when we investigate the effect of cut aggregation on the performance in Section IV-C.

C. Performance Analysis of Cut Aggregation Strategy

We conducted experiments aimed at studying the performance of the proposed cut aggregation strategy using k-means clustering algorithm introduced in Section III-B. Specifically, we evaluated the performance of the proposed cut aggregation strategy comparing with the static multicut aggregation used by Trukhanov et al. [28]. Under the static multicut aggregation, total n Benders optimality cuts that are generated from n scenarios were aggregated into k cuts so that each of k cuts is composed of n/k Benders cuts. For example, for 100 possible scenarios ($n = 100$), static cut aggregation with $k = 1$ corresponds to the L-shaped algorithm, $k = 100$ corresponds to the multicut algorithm, and $1 < k < 100$ corresponds to the partial aggregation that resigns between full aggregation (i.e. L-shaped algorithm) and full disaggregation (i.e. multicut algorithm). In addition, for implementation of the proposed cut aggregation, we use k as an input parameter (i.e. number of clusters) of k-means clustering algorithm, and Benders optimality cuts would be aggregated based on clustered scenarios. As described in Section III-B, we have an option to choose dimensions of k-means clusters (up to 3-dimensions) for the combinations of three pairwise correlations, $\rho_{DC}(\omega)$, $\rho_{DR}(\omega)$, and $\rho_{RC}(\omega)$). Note that we use one-dimensional k-means clustering for the pairwise correlation between power demand and renewable supply, $\rho_{DR}(\omega)$, that shows the most improved performance for the scenarios used in this study. Figures 5 and 6 show the computational times to obtain an optimal solution using the multicut L-shaped algorithm with various aggregation level k where $1 \leq k \leq n$ for each size of scenarios $n = 100, 200, 300, 400, 500$ and 600. Through analyzing the results of numerical experiments, we find that (i) both the static and the proposed cut aggregation improve the performance of algorithm at certain level of aggregate k, where $1 \leq k \leq n$, (ii) the proposed cut aggregation strategy shows better performance improvement than the static aggregation, and (iii) the best k exists between both extreme cases. We would like to emphasize that the multicut L-shaped algorithm shows better performance at higher aggregation level for scenario data used in this study.

V. Concluding Remark and Future Work

This study is motivated by an opportunity to reduce the energy cost and carbon pollution by utilizing renewable energy and adopting demand response from the demand-side perspective. While utilizing renewable energy to meet power demand, consumers may be willing to adjust their demand load, which is called as demand response, to avoid peak electricity price as well as optimally utilize renewable energy to reduce procurement cost. In addition, energy storage can be used to mitigate fluctuations of intermittent renewable supply and volatile electricity price. Considering renewable energy, demand response, and energy storage, the main objective of this study is to propose decision-making models that enable energy consumers to procure energy in a cost-efficient manner in response to variability and uncertainty of renewable supply as well as electricity price. In summary, the main contributions of this paper are: (i) propose day-ahead power procurement problem and formulate it as a two-stage SMIP problem; (ii) introduce cut generation and cut aggregation strategies that can be integrated with the course of the multicut L-shaped algorithm to improve algorithm performance; and (iii) implement the proposed algorithm by using lazy constraints pool provided by CPLEX Concert Technology and investigate performance by conducting numerical experiments with various settings. The proposed day-ahead power procurement problem and solution approach can be applied to many industries (e.g. data centers and manufacturing) and also extended to grid-level power system operations (e.g. micro grid) to curtail expenses of procuring energy to meet demand load. We believe that

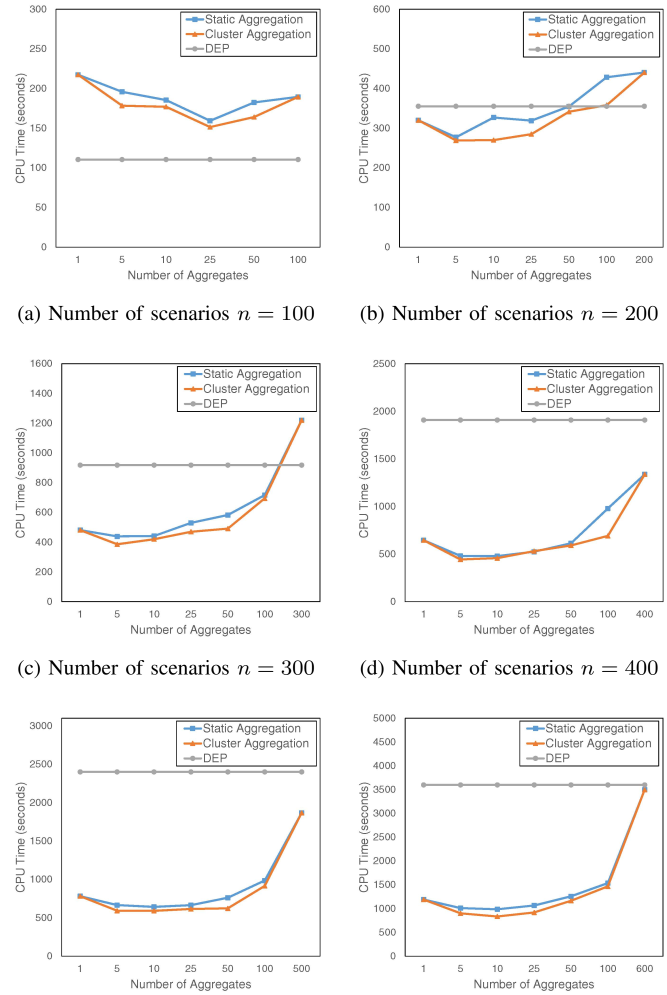


(a) Number of scenarios $n = 100$ (b) Number of scenarios $n = 200$

(c) Number of scenarios $n = 300$ (d) Number of scenarios $n = 400$

(e) Number of scenarios $n = 500$ (f) Number of scenarios $n = 600$

Fig. 5. % Reduction in CPU time: Static versus Cluster aggregations combined with valid inequality (33) and valid optimality cut (40)



(a) Number of scenarios $n = 100$ (b) Number of scenarios $n = 200$

(c) Number of scenarios $n = 300$ (d) Number of scenarios $n = 400$

(e) Number of scenarios $n = 500$ (f) Number of scenarios $n = 600$

Fig. 6. % Reduction in CPU time: Static versus Cluster aggregations combined with valid inequality (34) and valid optimality cut (40)

this study would be a good starting point to study demand-side power procurement problem based on the framework of two-stage stochastic program and will have a significant impact on study for the utilization of renewable energy and implementation of demand response.

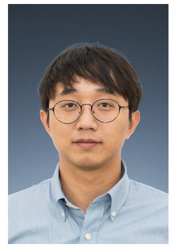
ACKNOWLEDGEMENTS

We thank the anonymous reviewers and editors for their insightful comments that lead to significant improvements in the content and presentation of this work.

REFERENCES

[1] J. Whitney and P. Delforge, "Data center efficiency assessment," *Natural Resources Defense Council, New York City, New York*, 2014.
[2] US Department of Energy, "2014 renewable energy data book," 2014.
[3] A. Energy, "Power smart pricing," 2012, [Online], Available:https://www.powersmartpricing.org/prices/.
[4] "Electricity reliability council of texas," [Online], Available:http://www.ercot.com/.
[5] "California independent system operator," [Online], Available:http://www.caiso.com/Pages/default.aspx.
[6] A. Wierman, Z. Liu, I. Liu, and H. Mohsenian-Rad, "Opportunities and challenges for data center demand response," in *Green Computing Conference (IGCC), 2014 International*. IEEE, 2014, pp. 1–10.
[7] Google.com, "Renewable energy - google green," [Online], Available:https://www.google.com/green/energy/\#power.
[8] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *Industrial Informatics, IEEE Transactions on*, vol. 7, no. 3, pp. 381–388, 2011.
[9] J. Aghaei and M.-I. Alizadeh, "Demand response in smart electricity grids equipped with renewable energy sources: A review," *Renewable and Sustainable Energy Reviews*, vol. 18, pp. 64–72, 2013.
[10] N. Ruiz, I. Cobelo, and J. Oyarzabal, "A direct load control model for virtual power plant management," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 959–966, 2009.
[11] I. Koutsopoulos, V. Hatzi, and L. Tassiulas, "Optimal energy storage control policies for the smart power grid," in *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*, 2011, pp. 475–480.
[12] P. M. Van de Ven, N. Hegde, L. Massoulié, and T. Salonidis, "Optimal control of end-user energy storage," *Smart Grid, IEEE Transactions on*, vol. 4, no. 2, pp. 789–797, 2013.
[13] Y. Xu and L. Tong, "On the operation and value of storage in consumer demand response," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, 2014, pp. 205–210.
[14] N. Gautam, Y. Xu, and J. T. Bradley, "Meeting inelastic demand in systems with storage and renewable sources," in *Smart Grid Communications (SmartGridComm), 2014 IEEE International Conference on*, 2014, pp. 97–102.
[15] S. Kwon, N. Gautam, and Y. Xu, "Meeting inelastic demand in systems with storage and renewable sources," *IEEE Transactions on Smart Grid*, 2015, accepted.
[16] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam,

"Optimal power cost management using stored energy in data centers," in *Proceedings of the ACM SIGMETRICS joint international conference on Measurement and modeling of computer systems*, 2011, pp. 221–232.
[17] L. Huang, J. Walrand, and K. Ramchandran, "Optimal demand response with energy storage management," in *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*, 2012, pp. 61–66.
[18] Y. Guo and Y. Fang, "Electricity cost saving strategy in data centers by using energy storage," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 24, no. 6, pp. 1149–1160, 2013.
[19] Y. Guo, Y. Gong, Y. Fang, P. P. Khargonekar, and X. Geng, "Energy and network aware workload management for sustainable data centers with thermal storage," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 25, no. 8, pp. 2030–2042, 2014.
[20] J. R. Birge and F. Louveaux, *Introduction to stochastic programming*. Springer Science & Business Media, 2011.
[21] R. M. Van Slyke and R. Wets, "L-shaped linear programs with applications to optimal control and stochastic programming," *SIAM Journal on Applied Mathematics*, vol. 17, no. 4, pp. 638–663, 1969.
[22] J. R. Birge and F. V. Louveaux, "A multicut algorithm for two-stage stochastic linear programs," *European Journal of Operational Research*, vol. 34, no. 3, pp. 384–392, 1988.
[23] J. F. Benders, "Partitioning procedures for solving mixed-variables programming problems," *Numerische mathematik*, vol. 4, no. 1, pp. 238–252, 1962.
[24] D. McDaniel and M. Devine, "A modified benders' partitioning algorithm for mixed integer programming," *Management Science*, vol. 24, no. 3, pp. 312–319, 1977.
[25] T. L. Magnanti and R. T. Wong, "Accelerating benders decomposition: Algorithmic enhancement and model selection criteria," *Operations research*, vol. 29, no. 3, pp. 464–484, 1981.
[26] G. Zakeri, A. B. Philpott, and D. M. Ryan, "Inexact cuts in benders decomposition," *SIAM Journal on Optimization*, vol. 10, no. 3, pp. 643–657, 2000.
[27] G. K. Saharidis and M. G. Ierapetritou, "Speed-up benders decomposition using maximum density cut (mdc) generation," *Annals of Operations Research*, vol. 210, no. 1, pp. 101–123, 2013.
[28] S. Trukhanov, L. Ntaimo, and A. Schaefer, "Adaptive multicut aggregation for two-stage stochastic linear programs with recourse," *European Journal of Operational Research*, vol. 206, no. 2, pp. 395–406, 2010.
[29] R. Rahmaniani, T. G. Crainic, M. Gendreau, and W. Rei, "The benders decomposition algorithm: A literature review," Tech. Rep., 2016.
[30] IBM, "Ibm ilog cplex optimization studio v12.6.3 documentation," *International Business Machines Corporation*, 2015.
[31] G. Laporte and F. V. Louveaux, "The integer l-shaped method for stochastic integer programs with complete recourse," *Operations research letters*, vol. 13, no. 3, pp. 133–142, 1993.
[32] E. Schubert, A. Koos, T. Emrich, A. Züfle, K. A. Schmid, and A. Zimek, "A framework for clustering uncertain data," *PVLDB*, vol. 8, no. 12, pp. 1976–1979, 2015. [Online]. Available: http://www.vldb.org/pvldb/vol8/p1976-schubert.pdf
[33] P. new jersey maryland interconnection, [Online], Available:http://www.pjm.com.
[34] J. R. Birge, "The value of the stochastic solution in stochastic linear programs with fixed recourse," *Mathematical programming*, vol. 24, no. 1, pp. 314–325, 1982.



Soongeol Kwon received B.S. and M.S. degrees in Information and Industrial Engineering from Yonsei University, Seoul, Korea, in 2005 and 2007, respectively, and he is currently pursuing a Ph.D. degree in Industrial and Systems Engineering at Texas A&M University, College Station, Texas. His research interests include mixed-integer programming, stochastic programming, stochastic process and control, and simulation. Research domains include but are not limited to (i) energy-efficiency in system operations; (ii) energy procurement with renewable energy; (iii) smart grid; and (iv) sustainable manufacturing.

Lewis Ntaimo is an Associate Professor in the Department of Industrial and Systems Engineering at Texas A&M University. He received his Ph.D. in Systems and Industrial Engineering from the University of Arizona (2004), M.S. in Mining and Geological Engineering (2000), and B.S. in Mining Engineering (1998), both from the University of Arizona. He received his Engineer-In-Training (EIT) certification (07760) in June 1999, State of Arizona. His research interests are in stochastic programming, systems modeling and engineering processes, and discrete event modeling and simulation. He is on the editorial board of the Journal of Global Optimization and he is vice president of the INFORMS Minority Issues Forum.

Natarajan Gautam is a Professor in the Department of Industrial and Systems Engineering at Texas A&M University with a courtesy appointment in the Department of Electrical and Computer Engineering. Prior to joining Texas A&M University in 2005, he was on the Industrial Engineering faculty at Penn State University for 8 years. He received his M.S. and Ph.D. degrees in Operations Research from the University of North Carolina at Chapel Hill and his B.Tech. from the Indian Institute of Technology, Madras. His research interests are in the areas of modeling, analysis, and performance evaluation of stochastic systems with special emphasis on optimization and control in computer, telecommunication, and information systems. He is an Associate Editor for the INFORMS Journal on Computing, IIE Transactions, and OMEGA.