



Optimal Energy Management of a Residential Prosumer: A Robust Data-Driven Dynamic Programming Approach

Zhongjie Guo , Wei Wei , Senior Member, IEEE, Laijun Chen , Member, IEEE, Zhaojian Wang , Member, IEEE, João P. S. Catalão , Senior Member, IEEE, and Shengwei Mei , Fellow, IEEE

Abstract—Prosumers are agents that both consume and produce energy. This article studies the optimal energy management of a residential prosumer which consists of a renewable power plant and an energy storage unit. Energy could stream among power grid, renewable plant, storage unit, and demand, providing a highly flexible energy supply and the opportunity of arbitrage. To capture the uncertainty of renewable generation and electricity price, as well as the rolling horizon feature of the multiperiod energy management, the problem is formulated as a robust data-driven dynamic programming (RDDP). Kernel regression is utilized to build the empirical conditional distribution in a data-driven manner, and all candidates that reside in a Wasserstein metric-based ambiguity set are taken into account to tackle the inexactness of the empirical distribution. The RDDP can be transformed into a series of convex optimization problems with cost-to-go functions in their constraints. The piecewise linear expression of the cost-to-go function is retrieved from dual linear programs. Through such an analytical expression of cost-to-go functions, the RDDP can be solved via backward induction, unlike the popular stochastic dual dynamic programming technique that incorporates forward and backward passes. Case studies validate the performance and advantage of the proposed RDDP approach.

Index Terms—Energy storage, prosumer, robust data-driven dynamic programming, uncertainty, value function approximation.

NOMENCLATURE

A. Abbreviations

ADP Approximate dynamic programming.
DG Distributed generator.

Manuscript received August 24, 2020; revised October 27, 2020; accepted December 3, 2020. The work of Zhongjie Guo, Wei Wei, and Shengwei Mei was supported by NSFC under Grant 51807101 and Grant U1766203 and in part by Shanxi Province Key Research and Development Project under Grant 201903D421029. The work of João P. S. Catalão was supported by the FEDER through COMPETE 2020 and in part by the FCT under Grant POCI-01-0145-FEDER-029803(02/SAICT/2017). (Corresponding author: Zhongjie Guo.)

Zhongjie Guo, Wei Wei, Zhaojian Wang, and Shengwei Mei are with the State Key Laboratory of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (e-mail: gzj18@mails.tsinghua.edu.cn; wei-wei04@mails.tsinghua.edu.cn; wangzhaojian@163.com; meishengwei@mail.tsinghua.edu.cn).

Laijun Chen is with the New Energy Industry Research Center, Qinghai University, Xi'ning 810016, China (e-mail: chenlaijun@tsinghua.edu.cn).

João P. S. Catalão is with the Faculty of Engineering of the University of Porto and INESC TEC, Porto 4200-465, Portugal (e-mail: catalao@fe.up.pt).

Digital Object Identifier 10.1109/JSYST.2020.3043342

ESU Energy storage unit.
LP linear programming.
PWL Piecewise linear.
RDDP Robust data-driven dynamic programming.
SDDP Stochastic dual dynamic programming.
SoC State of charge.

B. Parameters and Indices

x_{\max}^c Maximum charging power of ESU.
 x_{\max}^{dc} Maximum discharging power of ESU.
 s_{\min} Minimum SoC of ESU.
 s_{\max} Maximum SoC of ESU.
 η_c/η_{dc} Charging/discharging efficiency of ESU.
 Δ Duration of a time slot.
 w_t Available renewable power, random parameter.
 d_t Demand in time slot t , random parameter.
 π_s Selling price of electricity, random parameter.
 π_p Buying price of electricity, random parameter.
 ξ_t Random vector of exogenous states.
 N Number of historical observations.
 H Number of sampled state variables.
 $\mathcal{M}_t^\varepsilon$ Ambiguity set of the conditional distribution in period t with radius ε .

C. Variables

x_t^{wd} Power from renewable DG to demand.
 x_t^{wg} Power from renewable DG to power grid.
 x_t^{ws} Power from renewable DG to ESU.
 x_t^{sd} Power from ESU to demand.
 x_t^{sg} Power from ESU to power grid.
 x_t^{gd} Power from power grid to demand.
 x_t^{gs} Power from power grid to storage.
 s_t SoC of ESU.
 u_t Vector of control variables.

I. INTRODUCTION

WITH the development of technology, renewable distributed generators (DGs) have been widely deployed at the demand side, bringing various benefits not only to the distribution system but also to end consumers [1]. Private-owned DGs, like rooftop solar panels, small wind turbines, and micro gas-fired units, endow traditional end consumers with the ability

to produce, precipitating the advent of prosumers [2]. Energy storage unit (ESU) owned by the prosumer greatly improves the flexibility and reliability of maintaining power balance. On the one hand, the ESU helps to compensate the volatile output of renewable DGs; on the other hand, DG and ESU offer the prosumer a unique opportunity to actively participate in system operation through trading energy with the power grid [3]. Recently, prosumer has become a hot topic and attracted a lot of attention from researchers [4].

Because of low operation costs, renewable DGs have won the favor of prosumers. However, given the volatility of renewable generation, the energy management of prosumer becomes more challenging. Two-stage stochastic optimization and two-stage robust optimization are the mainstream methods for decision making under uncertainty. SO requires the probability distribution of uncertain factors and usually relies on scenario sampling [5] or chance-constraints [6]; robust optimization depicts uncertain parameters via an uncertainty set [7] and optimize the target under the worst-case scenario. Both of them have been widely applied in unit commitment [8], [9], economic dispatch [10], [11], optimal active-reactive power flow [12], [13], energy storage operation [14] and so on. A tutorial on two-stage robust optimization can be found in Appendix C of [15].

In the above two-stage optimization approaches, a decision is made with perfect knowledge of the realized uncertainty over the entire time horizon. However, in the energy management problem, renewable generation is observed period-by-period, and the dispatch strategy in period t cannot depend on the information after period t , which is called nonanticipativity; if nonanticipativity is neglected, the solution could be infeasible, as is illustrated in [16] by an example to demonstrate such phenomenon in the two-stage robust unit commitment problem. To overcome this difficulty, an affine policy-based multiperiod robust optimization method is proposed in [16] to enforcing nonanticipativity, in which the generator real-time adjustments are linear functions in the forecast errors. The method is applied to unit commitment with energy storage in [17]. Applying the affine policy restricts system flexibility and thus may give sub-optimal strategies. More detailed discussions on the optimality performance of affine policy can be found in [18].

Another remedy is to perform multiperiod optimization or dynamic programming in a rolling horizon fashion. The main difficulty is the nested optimization structure that requires evaluating the optimal value for the remaining periods (also called the cost-to-go function) recursively. Two prevalent methods are approximate dynamic programming (ADP) and stochastic dual dynamic programming (SDDP) [19], depending on how the cost-to-go function is approximated.

ADP approximates the cost-to-go function with a smaller number of parameters; general strategies include aggregation and continuous value function approximation [20]. The authors in [21] proposed a heuristic interpolation-based method. However, interpolation-based regression does not guarantee the performance of approximation (optimistic or pessimistic). The method in [21] is applied to battery storage management in [22]. An inspiring method in [23] focuses on finding the slopes of a piece of function rather than the value, but this approach does

not account for uncertainty. Machine learning methods are also found in value function approximation, such as deep recurrent neural network learning [24] and reinforcement learning [25]. ADP has been used in [26] for the strategic operation of energy hubs with uncertainties, in [27] for microgrid economic dispatch, in [28] for load management in smart grids, and so on.

SDDP approximates the cost-to-go function via multidimensional benders cuts. Lower bounds and upper bounds are generated in the forward and backward swapping process until a convergence criterion is met [29]. The applications of SDDP are also reported in [30] and [31] to economic dispatch, in [32] to transmission expansion planning and in [33] to ESU management strategy for microgrids. As variants of SDDP, stochastic dual dynamic integer programming [34], and distributionally robust SDDP [35] are also explored. In summary, SDDP is an elegant method that can solve dynamic programs under uncertainties with a provable guarantee on the optimality gap. However, as the forward pass and backward pass should repeat many times, the computation time may not meet the requirement of online use. Although some dedicated strategies have been proposed to accelerate SDDP, such as scenario reduction [36] and cut selection [37], it is still not clear how fast SDDP would converge in advance.

Except for ADP and SDDP, robust dual dynamic programming is proposed in [38]. It describes uncertain parameters using an uncertainty set and is a multistage robust optimization. The solution also relies on forward and backward swapping. A multiresolution dynamic programming is introduced in [39] for managing ESU. The optimization horizon is divided into a series of sub-horizons and discretized with different temporal resolutions, enabling a reduced computational complexity compared to the single-resolution approach. The authors in [40] discuss a data-driven multistage stochastic optimization approach for seasonal energy storage operation; the cost-to-go function is evaluated based on available historical data of uncertain parameters instead of modeling their underlying distributions.

In view of the limitation of two-stage optimization in the dynamic environment faced by prosumers, we resort to the dynamic programming modeling paradigm in [21]. Compared with existing work, the novelty of this article is twofold.

1) A robust data-driven dynamic programming (RDDP) formulation for energy management of a residential prosumer. To model the uncertainty in the dynamic environment, kernel regression is used to estimate the empirical conditional distribution using historical data following the paradigm in [21]; a Wasserstein metric-based ambiguity set is built accounting for the inexactness of conditional distribution. Unlike SDDP, the proposed method relies on data and is robust against the inexactness of conditional distribution. Compared with the robust dual dynamic programming in [38], the proposed RDDP is less conservative as the distributional property of uncertainty is taken into account.

2) A systematic approach to solve the RDDP, which entails solving a series of convex optimization problems with cost-to-go functions in its constraints. According to the multiparametric programming theory, we construct convex PWL expressions for the cost-to-go functions from dual variables extracted from

LPs, which enables to solve the RDDP via backward induction. The computational efficiency outperforms SDDP and robust dual dynamic programming in [38] which conduct forward pass and backward pass repeatedly. Unlike machine learning-based value function approximation, the approximation quality of the proposed PWL expression does not rely on the quality of historical data and training process. Data is only used to retrieve conditional distribution.

This article proposes a paradigm for the optimal energy management of the prosumer-like energy systems, including but not limited to residential users, microgrids, energy hubs, and industrial parks. Technically, the proposed method can approximate the cost-to-functions at a very low computational expense relying on historical data; the computational time of RDDP is $K \times \mathcal{O}(N)$, which is much shorter than that of SDDP, $\mathcal{O}(N^2)$. More importantly, the RDDP formulation enjoys a degree of robustness: even in case of the scarcity of historical data, the strategies can still guarantee an acceptable operating economy; results show that with only ten observations of historical data, the suboptimality of RDDP is less than 4%.

This article extends the work in [21] and [22] in two aspects. First, we consider the energy streams from the residential system to the main grid, improving the operational flexibility and capturing the arbitrage opportunity of the prosumer. Second, we develop a thorough scheme for approximating the cost-to-go function according to the multiparametric programming theory; compared with the interpolation-based approximation method in [21] and [22], which overly estimates the function value, such a method leads to an under-estimator and exhibits better accuracy without increasing complexity, especially when the sampling set is small.

The rest of this article is organized as follows: the architecture of the prosumer is established in Section II, followed by the RDDP formulation of the optimal energy management problem. The solution method for the RDDP is presented in Section III. The results of numerical tests are reported in Section IV. Finally, conclusions are drawn in Section V.

II. MATHEMATICAL FORMULATION

In this section, the architecture of the residential prosumer is introduced first; then the energy management problem in the deterministic case is presented; finally, the RDDP is set forth by incorporating uncertainty.

A. Architecture of the Residential Prosumer

The architecture of residential prosumer studied in this article is depicted in Fig. 1. We assume that the interfaces between wind turbine and storage, wind turbine and the load devices, wind turbine and grid, storage and grid (bidirectional), storage and load devices, as well as grid and load devices, are available. Demand is supplied by the DG, ESU, and the grid; ESU can store energy from renewable generation and grid; specifically, the energy in storage or generated from wind turbine can be sold back to the grid, which is practicable in a decentralized power grid. The ESU plays an important role in this residential-level energy system due to its ability to shift demand over time. The

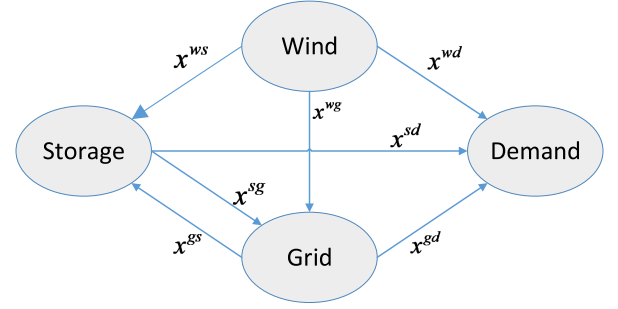


Fig. 1. Configuration of the residential prosumer.

model of ESU is given as

$$0 \leq x_t^{gs} + x_t^{ws} \leq x_{\max}^c, \forall t \quad (1a)$$

$$0 \leq x_t^{sg} + x_t^{sd} \leq x_{\max}^{dc}, \forall t \quad (1b)$$

$$s_{\min} \leq s_t \leq s_{\max}, \forall t \quad (1c)$$

$$s_{t+1} = s_t + (x_t^{gs} + x_t^{ws})\Delta\eta_c - \frac{(x_t^{sg} + x_t^{sd})\Delta}{\eta_{dc}}, \forall t \quad (1d)$$

where constraints (1a) and (1b) impose non-negativity and upper limits on charging power and discharging power of the ESU; constraint (1c) stipulates the feasible range of SoC; (1d) describes the SoC dynamics over time. Here we do not exert strict complementarity on charging and discharging power, which can be a barrier to unleash the flexibility of prosumer. The condition that the terminal SoC equals to the initial state is relaxed [41].

The total power supply from the renewable plant cannot exceed the available renewable generation w_t , yielding

$$x_t^{wd} + x^{ws} + x_t^{wg} \leq w_t, \forall t. \quad (2)$$

Demand balance boils down to

$$x_t^{sd} + x^{wd} + x_t^{gd} = d_t, \forall t. \quad (3)$$

Such an architecture enjoys high flexibility of reconfiguration. For example, if a real prosumer system does not have the interface between wind turbine and grid, the variable x^{wg} is set to 0. Therefore, the proposed architecture and its model are very general.

B. Prosumer Energy Management in Stochastic Form

Given the rolling horizon feature and uncertain nature, the energy management problem can be cast as a stochastic DP. In period t , storage SoC $s_t \in \mathbb{S}$ is the endogenous state, where \mathbb{S} is defined in (1c). Vector $\xi_t = \{w_t, \pi_t^s, \pi_t^p, d_t\} \in \mathbb{R}^4$ which encapsulates all external random parameters is the exogenous state, where π_t^s/π_t^p denotes selling and purchasing prices of electricity. ξ_t is observed at the beginning of period t . Based on (s_t, ξ_t) , the prosumer determines control action $u_t = \{x_t^{gs}, x_t^{gd}, x_t^{sg}, x_t^{sd}, x_t^{wg}, x_t^{wd}, x_t^{ws}\} \in \mathbb{U}_t(\xi_t)$ which includes energy flow variables in period t , where feasible set \mathbb{U}_t depending on w_t and d_t corresponds to constraints (1a), (1b), (2), and (3). After prosumer deploys action u_t , the system moves to a new state s_{t+1} at the beginning of period $t + 1$ according to

the state transition equation $s_{t+1} = g_t(s_t, u_t)$ described in (1d). Then, ξ_{t+1} is observed. The action u_t incurs an instantaneous payoff in period t which represents the cost for purchasing electricity from the power grid and the income for selling electricity to the power grid, i.e., $c_t(\xi_t, u_t) = \pi_t^p(x_t^{gs} + x_t^{gd}) - \pi_t^s(x_t^{sg} + x_t^{wg})$.

Based on the above notations and Bellman's principle of optimality [19], the stochastic dynamic programming model of prosumer energy management is cast in a backward recursive manner

$$\begin{aligned} V_t(s_t, \xi_t) &= \min c_t(\xi_t, u_t) + \mathbb{E}[V_{t+1}(s_{t+1}, \xi_{t+1}) | \xi_t] \\ \text{s.t. } u_t &\in \mathbb{U}_t(\xi_t), s_{t+1} \in \mathbb{S} \\ s_{t+1} &= g_t(s_t, u_t) \end{aligned} \quad (4)$$

for $t = T, \dots, 1$ with $V_{T+1} \equiv 0$, where $V_t(s_t, \xi_t)$ is the optimal payoff for the remaining $T - t$ periods and is called the cost-to-go function or value function. It quantifies the optimal value based on (s_t, ξ_t) revealed at the beginning of period t . However, as the payoff depends on the future realization of uncertain parameters, we resort to minimize the conditional expectation in problem (4) for each period.

C. RDDP Model

The stochastic DP model (4) encounters two difficulties in practical use. From the modeling perspective, the exact conditional probability distribution of ξ_{t+1} given ξ_t is not available; from the computational perspective, to perform backward induction, we need the analytical expression of cost-to-go function. In this section, we use the kernel regression method to construct the empirical conditional distribution based on historical data and formulate an RDDP model, inspired by the paradigm in [21]. Specifically, we consider all possible distributions in a Wasserstein-metric based ambiguity set which needs fewer data than the χ^2 -distance based one in [21]. The computational issue is left for the next section.

Given a set of N historical trajectories of exogenous states $\{\xi_t^i\}_{i=1}^N, \forall t$, the conditional expectation in (4) can be estimated via Nadaraya-Watson kernel regression method [21], [25]

$$\sum_{i=1}^N q_{ti}(\xi_t) V_{t+1}(s_{t+1}, \xi_{t+1}^i) \quad (5)$$

where the weight coefficients $q_{ti}(\xi_t)$ are obtained from

$$q_{ti}(\xi_t) = \frac{\mathcal{K}(\xi_t - \xi_t^i)}{\sum_{k=1}^N \mathcal{K}(\xi_t - \xi_t^k)} \quad (6)$$

where

$$\mathcal{K}(y) = \exp\left(-\frac{(\|y\|_2)^2}{2\sigma^2}\right) \quad (7)$$

is the Gaussian kernel function; $\|\cdot\|_2$ stands for the Euclidean norm of a vector; $\sigma > 0$ is a bandwidth parameter. The kernel function guarantees that the observation nearer to ξ_t holds a larger weight. Substituting (5) into (4) we have

$$\begin{aligned} [b]V_t(s_t, \xi_t) &= \min c_t(\xi_t, u_t) \\ &+ \sum_{i=1}^N q_{ti}(\xi_t) V_{t+1}(s_{t+1}, \xi_{t+1}^i) \\ \text{s.t. } u_t &\in \mathbb{U}_t(\xi_t), s_{t+1} \in \mathbb{S} \\ s_{t+1} &= g_t(s_t, u_t). \end{aligned} \quad (8)$$

Problem (8) is reminiscent of SDDP and can be solved via forward pass and backward pass algorithm in [29]. In practice, however, the historical data may be sparse, and the conditional probabilities $q_{ti}(\xi_t)$ could be inexact with the variance scalings with $\mathcal{O}(\frac{1}{N})$ [21].

To cope with this issue, when calculating the expectation in (5), not only the empirical distribution in (6), but also other distributions that are close to the empirical one, should be taken into account. To this end, we employ the Wasserstein metric to quantify the distance between two probability distributions. The Wasserstein metric for two discrete probability distributions P_1 and P_2 in probability space $\mathcal{P}(\Xi)$ supported on Ξ is defined through an optimal transport problem

$$\begin{aligned} d_W(P_1, P_2) &= \inf_{\pi \geq 0} \sum_i \sum_j \pi_{ij} \|\zeta_1^i - \zeta_2^j\|_p \\ \text{s.t. } \sum_j \pi_{ij} &= P_1^i, \forall i \\ \sum_i \pi_{ij} &= P_2^j, \forall j \\ \sum_i \sum_j \pi_{ij} &= 1 \end{aligned} \quad (9)$$

where decision variable $\pi \in \mathbb{R}_+^{I \times J}$ is a joint distribution of random variables ζ_1 and ζ_2 with marginal distributions P_1 and P_2 , respectively; $\|\cdot\|_p$ is the vector p -norm. We adopt $p = 1$ in this article, and d_W is called 1-Wasserstein metric or Kantorovich metric [42]. If we regard π_{ij} as the probability mass transported from slot i to slot j , and $\|\zeta_1^i - \zeta_2^j\|_p$ as the corresponding energy consumption, then problem (9) aims to find the energy optimal plan for reshaping distribution P_1 to distribution P_2 , and the minimal energy consumption is defined as the distance between P_1 and P_2 .

Equipped with the Wasserstein metric, we can build the ambiguity set which includes all probability distributions near the empirical distribution $Q = [q_{t1}(\xi_t), \dots, q_{tN}(\xi_t)], \forall t$

$$\mathcal{M}_t^\varepsilon = \{P \in \mathcal{P}(\{\xi_t^i\}_{i=1}^N) : d_W(P, Q) \leq \varepsilon\} \quad (10)$$

where distribution P is described by the vector of probabilities $[p_{t1}(\xi_t), \dots, p_{tN}(\xi_t)]$ associated with historical trajectories $\{\xi_t^i\}_{i=1}^N$. According to (10), the distance between empirical distribution Q and any $P \in \mathcal{M}_t^\varepsilon$ is no greater than ε in the sense of 1-Wasserstein metric. Parameter ε is called the radius of the ambiguity set, and depends on the prosumer's attitude on risk. With the increase of ε , more distributions are contained in $\mathcal{M}_t^\varepsilon$, and the model is more conservative, reflecting a risk-averse attitude. If the attitude towards risk is not clear, a proper value

of ε is recommended by [43]

$$\varepsilon = -\log(\alpha^*)/N \quad (11)$$

where N is the total number of historical observations. With such a selection, the probability for the true distribution belonging to $\mathcal{M}_t^\varepsilon$ is no less than $1 - \alpha^*$.

With the ambiguity set $\mathcal{M}_t^\varepsilon$, the robust version of stochastic DP (8), which is called RDDP, can be written as

$$\begin{aligned} [b]V_t(s_t, \xi_t) &= \min c_t(\xi_t, u_t) \\ &+ \max_{p_t(\xi_t) \in \mathcal{M}_t^\varepsilon} \sum_{i=1}^N p_{ti}(\xi_t) V_{t+1}(s_{t+1}, \xi_{t+1}^i) \\ \text{s.t. } u_t &\in \mathbb{U}_t(\xi_t), s_{t+1} \in \mathbb{S} \\ s_{t+1} &= g_t(s_t, u_t). \end{aligned} \quad (12)$$

In RDDP (12), the expectation is evaluated at the worst-case distribution in the ambiguity set $\mathcal{M}_t^\varepsilon$, therefore, the statistical performance of the optimal strategy is robust against the inexactness of conditional probability distribution. Compared to (8), an inner maximization over conditional distribution is performed accounting for uncertainty. Compared to a similar model in [21], the Wasserstein metric-based ambiguity set adapts to various cases with different data availability.

III. SOLUTION METHOD

In this section, we develop a systematic methodology to solve RDDP (12) via LP solver and backward induction. The key steps entail the elimination of inner maximization and the approximation of the cost-to-go function.

A. Eliminating the Inner Maximization

The maximization over $\mathcal{M}_t^\varepsilon$ is an infinite-dimensional optimization problem and must be reformulated. To this end, we summarize the observations from the concrete model provided as follows.

- 1) The cost function c_t is linear in u_t .
- 2) The state transition equation g_t is linear in s_t and u_t .
- 3) The feasible set \mathbb{U}_t is a polyhedron.

According to Proposition 2 in [44], problem (12) can be reduced to a finite-dimensional optimization problem

$$\begin{aligned} V_t(s_t, \xi_t) &= \min c_t(\xi_t, u_t) + \lambda\varepsilon - \mu + \sum_{i=1}^N q_{ti}(\xi_t) y_i \\ \text{s.t. } u_t &\in \mathbb{U}_t(\xi_t), s_{t+1} \in \mathbb{S} \\ \mu &\in \mathbb{R}, \lambda \in \mathbb{R}_+, z, y \in \mathbb{R}^N \\ s_{t+1} &= g_t(s_t, u_t) \\ V_{t+1}(s_{t+1}, \xi_{t+1}^i) + \mu &\leq z_i, \forall i = 1, \dots, N \\ z_i - \lambda \|\xi_t^i - \xi_t^j\|_1 &\leq y_j, \forall i, j = 1, \dots, N \end{aligned} \quad (13)$$

where μ, λ, z , and y are auxiliary variables; $q_{ti}(\xi_t), \forall i, t$ and $\|\xi_t^i - \xi_t^j\|_1, \forall i, j, t$ are constants calculated from historical observations. In the next section, we prove the cost-to-go function

$V_{t+1}(s_{t+1}, \xi_{t+1}^i)$ can be expressed via a convex PWL function, so that problem (13) gives rise to an LP and can be solved very efficiently. Solve (13) via backward induction (from the last period to the first period), we obtain the solution of RDDP (12). The strategy in the first period is deployed. The remaining task is to construct the PWL expression of the cost-to-go function in period t with respect to each observation.

B. PWL Expression of the Cost-to-go Function

We need an analytical expression of $V_{t+1}(s_{t+1}, \xi_{t+1}^i)$ to solve the problem (13). We know $V_{T+1} \equiv 0$, so problem (13) in period T is an LP. Suppose $V_{t+1}(s_{t+1}, \xi_{t+1}^i)$ is a convex PWL function in s_{t+1} , we prove $V_t(s_t, \xi_t^i)$ remains a convex PWL function in s_t . For notation brevity, let s_t denote the parameter and x represents all decision variables in (13). Under the assumption on $V_{t+1}(s_{t+1}, \xi_{t+1}^i)$, all constraints in (13) can be written as linear equalities and inequalities, resulting in an explicit LP. Proceeding one step back, we solve

$$\begin{aligned} [b]V_t(s_t, \xi_t) &= \min c^\top x \\ \text{s.t. } Ax &\leq b + Bs_t \end{aligned} \quad (14)$$

where A, B, b , and c are constant coefficients corresponding to the concrete model. We aim to demonstrate that the optimal value function $V_t(s_t, \xi_t)$ can be analytically expressed as a convex PWL function in s_t .

Recalled the basic property of LP, the optimal solution can always be found at one of the vertices of the polyhedral feasible region. At the optimal solution, the constraints in (14) are categorized into active ones and inactive ones

$$\begin{aligned} A'x^* &= b' + B's_t \\ A''x^* &< b'' + B''s_t. \end{aligned} \quad (15)$$

Hence, the optimal solution and optimal value are

$$x^* = A'^{-1}b' + A'^{-1}B's_t, s_t \in \Theta \quad (16a)$$

$$\begin{aligned} [b]V_t^*(s_t, \xi_t) &= c^\top A'^{-1}b' + c^\top A'^{-1}B's_t \\ &= m_t + n_t s_t, s_t \in \Theta \end{aligned} \quad (16b)$$

where Θ is called a critical region in which the set of active constraints remains unchanged. From (16) we can see that the optimal solution and the optimal value are linear functions of s_t . In fact, the feasible interval \mathbb{S} is covered by several disjoint subintervals, i.e., $\mathbb{S} = \bigcup_{i=1}^I \Theta_i$, where $\Theta_i \cap \Theta_j = \emptyset, \forall i \neq j$. Therefore, the cost-to-go function must be PWL in s_t

$$V_t(s_t, \xi_t) = \begin{cases} m_1 + n_1 s_t, & \forall s_t \in \Theta_1 \\ \vdots \\ m_I + n_I s_t, & \forall s_t \in \Theta_I \end{cases} \quad (17)$$

where the coefficients m_i and n_i depend on ξ_t . Computing the critical regions can be tricky in case of degeneracy. Here we use (17) to demonstrate the PWL structure of $V_t(s_t, \xi_t)$; and critical region is not actually used in the proposed method, once we prove the convexity of $V_t(s_t, \xi_t)$.

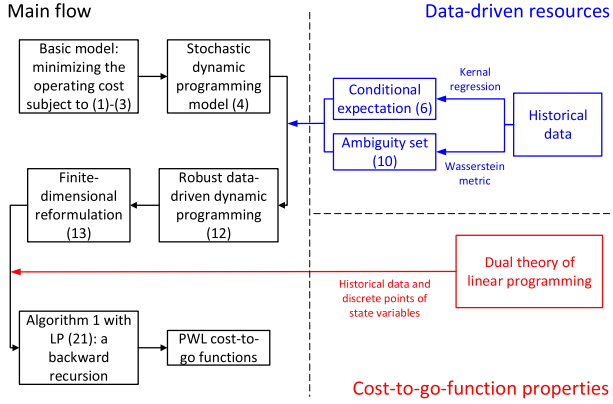


Fig. 2. Flowchart of the proposed RDDP approach.

Consider the dual problem of LP (14)

$$\begin{aligned} V_t(s_t, \xi_t) = \max_{\gamma} \gamma^\top (b + B s_t) \\ \text{s.t. } A^\top \gamma = c, \gamma \leq 0 \end{aligned} \quad (18)$$

where γ is the vector of dual variables. By strong duality, the optimal value of LP (18) coincides with $V_t(s_t, \xi_t)$, the optimal value of its primal problem. For any given γ , the objective function of (18) is a linear function in s_t . Therefore, $V_t(s_t, \xi_t)$ is the point-wise maximum of infinitely many linear functions, and thus convex in s_t .

The dual problem (18) provides a more convenient way to build PWL expression (17) without the information on critical region. Let $\Gamma = \{\gamma | A^\top \gamma = c, \gamma \leq 0\}$ be the polyhedral feasible set of dual variable and $\text{vert}(\Gamma)$ the set of vertices. Because the optimum is finite, the optimal solution must be found at some $\gamma \in \text{vert}(\Gamma)$, although Γ may contain extreme rays. In this regard

$$V_t(s_t, \xi_t) = \max_i \{\gamma_i^\top b + \gamma_i^\top B s_t\}, \forall \gamma_i \in \text{vert}(\Gamma). \quad (19)$$

Compare (17) and (19), coefficients m and n can be retrieved from dual variable via

$$m_i = \gamma_i^\top b, n_i = \gamma_i^\top B, \text{ for some } \gamma_i \in \text{vert}(\Gamma). \quad (20)$$

However, we do not know in prior which vertex produces the linear functions in (19). We propose the following heuristic method: choose uniformly distributed points $s_t^k \in \mathbb{S}$, $k = 1, \dots, K$; solve problem (14) for each s_t^k , the corresponding dual variable is γ_k , $k = 1, \dots, K$. After deleting duplicated elements in $\{\gamma_k\}_{k=1}^K$, we can retrieve the PWL expression according to (19).

Now we can claim: $V_t(s_t, \xi_t)$ is a convex PWL function in $s_t, \forall t$. The flowchart of constructing PWL expression for the cost-to-go functions is summarized in Algorithm 1. For clarity, an overall flowchart of the proposed RDDP approach is given in Fig. 2.

Additional discussions are given as follows.

1) Algorithm 1 performs backward induction from period T to period 2, unlike SDDP algorithm which performs forward pass and backward pass repeatedly. Hence, there is no convergence issue.

Algorithm 1: PWL expression of Cost-to-go Function.

1. Input:

K discrete points: $\{s_t^k\}_{k=1}^K \in \mathbb{S}, \forall t = 1, \dots, T$.

N historical observations: $\{\xi_t^l\}_{t=1}^T, \forall l = 1, \dots, N$.

$V_{T+1}(s_{T+1}, \xi_{T+1}^l) \equiv 0, \forall l = 1, \dots, N$.

2. For $t = T, \dots, 2$

For $l = 1, \dots, N$

For $k = 1, \dots, K$

Solve the following LP with $(s_t, \xi_t) = (s_t^k, \xi_t^l)$

$$\min c_t(\xi_t, u_t) + \lambda \varepsilon$$

$$- \mu + \sum_{i=1}^N q_{ti}(\xi_t) y_i$$

$$\text{s.t. } u_t \in \mathbb{U}_t(\xi_t), s_{t+1} \in \mathbb{S}$$

$$\mu \in \mathbb{R}, \lambda \in \mathbb{R}_+, z, y \in \mathbb{R}^N$$

$$s_{t+1} = g_t(s_t, u_t)$$

$$m_{t+1,i}^j + n_{t+1,i}^j s_{t+1} + \mu \leq z_i,$$

$$\forall j = 1, \dots, K, \forall i = 1, \dots, N$$

$$z_i - \lambda \|\xi_t^i - \xi_t^j\|_1 \leq y_j,$$

$$\forall i, j = 1, \dots, N$$

(21)

where $m_{t+1,i}^j$ and $n_{t+1,i}^j$ are obtained in the previous iteration. Calculate $m_{t,i}^j$ and $n_{t,i}^j$ by (20).

End

End

End

3. Output: Coefficients $m_{t,i}^j$ and $n_{t,i}^j, \forall i, t, j$.

2) The output of Algorithm 1 is the PWL expressions of cost-to-go functions, not the optimal strategy. The execution process can be viewed as training. When the real trajectory of uncertainty $\{\xi_t\}_{t=1}^T$ is observed period-by-period, the optimal action can be determined by solving LP (21) moving forward.

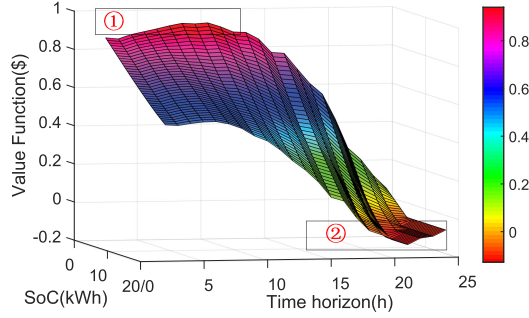
3) An interpolation-based method is proposed in [21] to approximate cost-to-go functions. It actually yields an over-estimator for univariate cost-to-go functions. The proposed method uses a subset of $\text{vert}(\Gamma)$ to construct PWL expressions and results in an under-estimator for cost-to-go functions (either univariate or multivariate). The heuristic method mentioned in [29] can help initiate discrete points s_t^k . With proper discrete points, the PWL expression is exact, at least is very close to the exact one.

IV. CASE STUDIES

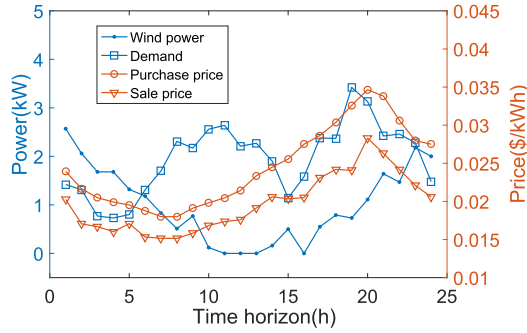
A residential prosumer with a storage and a wind turbine is used to validate the proposed method. Some parameters are given in Table I; complete system data are available in [45]. LPs are solved by CPLEX 12.8 on a laptop with Intel i5-8250U CPU and 8 GB memory.

TABLE I
SYSTEM PARAMETERS

Parameter	value(unit)	Parameter	value(unit)
x_{max}^c	2.5kW	x_{max}^{dc}	2kW
s_{min}	6kWh	s_{max}	20kWh
η_c	90%	η_{dc}	90%



(a)



(b)

Fig. 3. Analysis on approximate PWL function. (a) PWL cost-to-go functions across the day. (b) Corresponding observation.

A. Performance of the PWL Expression

To conduct Algorithm 1, we collect ten couples of historical observations of wind power and demand from a real residential energy system [46]; so $N = 10$. The electricity price is collected from the PJM database from July 1st to July 10th, 2019 [46]. One observation is plotted in Fig. 3, whose corresponding cost-to-go functions over the entire time horizon are drawn in Fig. 3, from which we can see the following.

1) In period t , the cost-to-go function $V_t(\cdot)$ is monotonically decreasing in its argument s_t . The reason is clear: A higher SoC can supply more energy in the future for demand or for sale, and thus reduce the purchasing cost while circumventing potential risks brought by market price uncertainty.

2) With time rolling on, cost-to-go functions in each period shows a decreasing trend in general, because less energy is needed in the rest of the day.

3) Distinguished from the overall trend, the first exception marked by ① exhibits an increase when battery SoC is low in the early morning. Such a phenomenon is more clearly displayed in Fig. 4. The reason can be deduced from Fig. 3 that in related periods, the available wind power drops considerably while

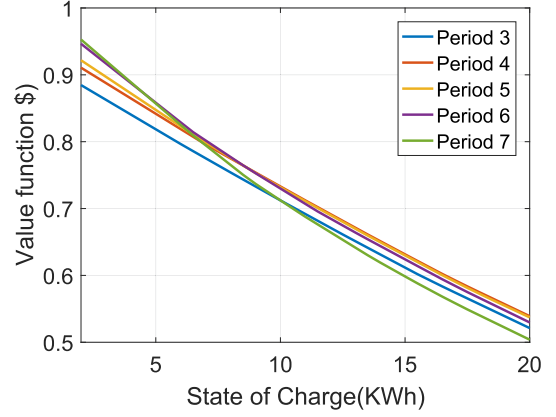


Fig. 4. PWL cost-to-go functions in periods 3–7.

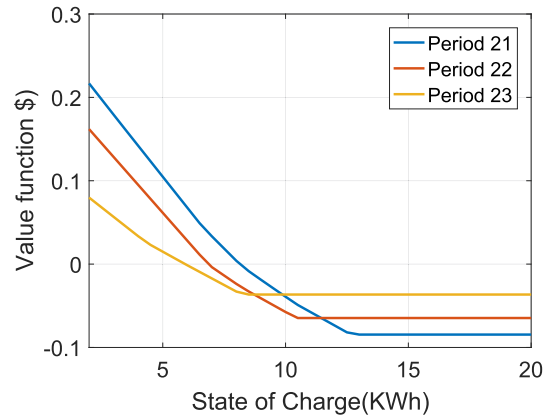


Fig. 5. PWL cost-to-go functions in periods 21–23.

the demand grows oppositely; therefore, the prosumer has to purchase energy from the grid if the energy stored ahead of the demand rise is not much.

4) The second exception marked by ② occurs at the end of a day. Referring to Fig. 5, the cost-to-functions in the last a few periods have values higher than those in period 21 when the level of SoC is relatively high. Because we assume $V_{T+1} \equiv 0$, the prosumer will sell out the energy before a new day. Given a maximum discharging power, the energy is sold across multiple periods. Consequently, the prosumer enjoys a higher revenue (represented by negative cost) in periods 22 and 23 than in the last period.

B. Impact of Data Availability and Parameter ε

By Algorithm 1 with ten groups of historical observations mentioned above, all the cost-to-go functions are already calculated, with which we can solve the reformulated dynamic program (13) as time recedes from period 1–24 with new observations.

To this end, we employ another five sequences of uncertain renewable power, demand and electricity price, denoted by $S1 \sim S5$; they act as the observations, which are given period by period. The initial level of SoC is set to 10 kWh. The prosumer

TABLE II
COMPARISONS ON INDEX 1 AND 2 WITH DIFFERENT ε

Test	Index	S1	S2	S3	S4	S5
$\varepsilon = 0$	Index 1(\$)	0.328	0.315	0.331	0.321	0.326
	Index 2(\$)	0.286	0.345	0.384	0.375	0.269
$\varepsilon = 0.3$	Index 1(\$)	0.585	0.572	0.587	0.578	0.582
	Index 2(\$)	0.294	0.330	0.368	0.381	0.270
$\varepsilon = 1$	Index 1(\$)	0.634	0.621	0.636	0.628	0.631
	Index 2(\$)	0.295	0.341	0.368	0.395	0.282

TABLE III
COMPARISONS ON ENERGY STREAMS WITH DIFFERENT ε

Test	Energy streams(kWh)						
	e^{wg}	e^{wd}	e^{ws}	e^{gs}	e^{gd}	e^{sd}	e^{sg}
$\varepsilon = 0$	3.180	14.97	6.046	5.053	16.75	15.66	0.528
$\varepsilon = 0.3$	3.354	14.68	6.171	4.914	16.64	16.07	0.109
$\varepsilon = 1$	3.500	14.48	6.224	4.874	16.81	16.10	0.093

can freely choose ε , or use the recommended value in (11), which depends on data availability. Three values $\varepsilon = 0$, $\varepsilon = 0.3$ (recommended value), and $\varepsilon = 1.0$ are tested.

To validate the performance of the RDDP method, we test three representative indices. Index 1: operation cost evaluated by problem (21) at period 1; Index 2: actual operation cost calculated from $\sum_{t=1}^T c_t$.

The comparison results of Index 1 and 2 are displayed in Table II. Index 1 increases with the growth of ε , which defines the size of the ambiguity set. The larger the value of ε , the more distributions are taken into account, resulting in a worst extreme distribution as well as a higher cost. For Index 2, a risk-neutral prosumer with $\varepsilon = 0$ may enjoy a lower cost in scenario 1 and scenario 5; while he has to pay a price for the strategy without robustness: In scenario 2 or 3, the cost is higher not only than those corresponding to two risk-averse cases, but also higher than its predictive expectation.

Furthermore, we explore the impact of ε on energy flows (in kWh, for example, $e^{wg} = \sum_{s=1}^5 \sum_{t=1}^T x_t^{wg}/5$). Results are listed in Table III, from which we can observe the following.

1) The allocation of wind energy changes with the value of ε ; specifically, as ε becomes larger, more volatile wind energy is sold to the grid, and more demand is supplied by either the grid power or the storage, which is controllable.

2) The energy exchange between storage and grid decreases as ε grows larger. Originally, the storage plays a prime role in arbitrage; when the uncertainty increases, more storage capacity is used to eliminate the negative impact of uncertainty and ensure a reliable energy supply.

The change of SoC from another perspective demonstrates how the robustness impacts the energy management strategy. The average SoC trajectories throughout the whole time horizon in three tests are drawn in Fig. 6; in the test with a larger ε , the SoC is kept at a higher level to avoid potential risks.

C. Comparison Between SDDP and RDDP

We compare the effectiveness and efficiency of the proposed RDDP and the popular SDDP. Please refer to [29] for a complete

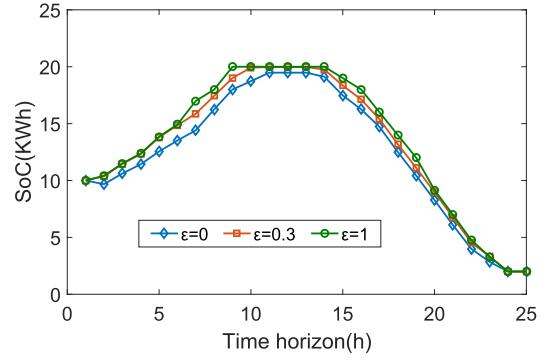


Fig. 6. Comparison on average SoC with different ε .

TABLE IV
COMPARISONS ON COMPUTATIONAL TIME WITH DIFFERENT SIZE OF HISTORICAL INFORMATION

Method	N=10	N=15	N=20	N=50
RDDP	48s	75s	99s	245s
SDDP	52s	108s	204s	1228s

TABLE V
COMPARISONS ON INDEX 2 BETWEEN RDDP AND SDDP

Method	S1	S2	S3	S4	S5
RDDP (\$)	0.294	0.330	0.368	0.381	0.270
SDDP (\$)	0.297	0.345	0.390	0.392	0.271
Relative error	1%	4%	5%	2.8%	0.3%

description of the forward-backward pass algorithm for SDDP. We use the same ten observations in SDDP. The algorithm terminates when the upper bound is close enough to the lower bound. Basically, LP in each step can be solved efficiently within 0.02 s. As we use $N = 10$ observations, the optimality gap between lower and upper bounds shrinks to 6×10^{-4} after ten forward passes and ten backward passes. Because all the scenarios must be visited in every backward pass, the computational time of SDDP is $O(N^2)$, which grows quickly if more historical observations are available. As for RDDP, the computational time is $K \times O(N)$ where the number of discrete points $K = 10$ is a constant. Comparatively, results in Table IV verify that RDDP is more efficient while the computational time of SDDP rapidly soars with the growth of the size of historical information.

Finally, we compare Index 2 in the five newly observed trajectories. The forward-backward pass algorithm of SDDP constructs accurate cost-to-go functions. From Table V, we can see that the relative errors between Index 2 offered by RDDP and SDDP in all five scenarios are less than 5%, demonstrating that the PWL expressions used in RDDP are accurate enough. Furthermore, the values offered by RDDP are slightly smaller, because in Algorithm 1, the dual variables for constructing PWL expression are extracted from discrete samples of parameters, and some critical vertices [as in (19)] that produces binding linear functions in the PWL expression may be missed. In this regard, the proposed method offers an under estimator for the cost-to-go function in theory. Nevertheless, as the sampled parameters are well distributed, the approximation is very close

to the accurate one. To improve accuracy, we can incorporate more sampled points of state variables to generate candidate dual variables.

V. CONCLUSION

This article proposes a robust data-driven dynamic programming method for prosumer energy management considering uncertainties of electricity price, renewable generation, and demand. The problem can be solved via backward induction using only linear programming solver, and the key step entails a piecewise linear approximation for the cost-to-go function, which is constructed from dual linear programs.

Case studies corroborate the effectiveness of the proposed RDDP method, revealing that the dispatch strategy of RDDP can keep a proper balance between optimality and computational expense: Compared to SDDP, the computational time of RDDP is $K \times O(N)$, which is much shorter than that of SDDP, $O(N^2)$; besides, results show the suboptimality of RDDP is less than 4%. Such results are based on only ten observations of historical data; therefore, the RDDP formulation enjoys a degree of robustness: even in case of the scarcity of historical data, the strategies can still guarantee an acceptable operating economy.

However, the disadvantage of the proposed method is that when the dimension of state variables is high, the computational complexity will exponentially grow.

In general, such a study establishes a paradigm for the optimal energy management of the prosumer-like energy systems, including but not limited to residential users, micro grids, energy hubs, and industrial parks. Our ongoing work is to apply RDDP to storage operation in large power systems with network flow constraints and develop more efficient solution algorithms.

REFERENCES

- [1] T. Adefarati and R. C. Bansal, "Integration of renewable distributed generators into the distribution system: A review," *IET Renew. Power Gener.*, vol. 10, no. 7, 2016, Art. no. 873884.
- [2] Y. Parag and B. K. Sovacool, "Electricity market design for the prosumer era," *Nat. Energy*, vol. 1, no. 16032, 2016.
- [3] I. Atzeni, L. G. Ordez, G. Scutari, D. P. Palomar, and J. R. Fonollosa, "Demand-side management via distributed energy generation and storage optimization," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 866–876, Jun. 2013.
- [4] L. Bhanidi and S. Sivasubramani, "Optimal sizing of smart home renewable energy resources and battery under prosumer-based energy management," *IEEE Syst. J.*, to be published. doi: [10.1109/JSYST.2020.2967351](https://doi.org/10.1109/JSYST.2020.2967351).
- [5] A. Ahmadi-Khatir, A. J. Conejo, and R. Cherkaoui, "Multi-area energy and reserve dispatch under wind uncertainty and equipment failures," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4373–4383, Nov. 2013.
- [6] H. Wu, M. Shahidehpour, Z. Li, and W. Tian, "Chance-constrained day-ahead scheduling in stochastic power system operation," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1583–1591, Jul. 2014.
- [7] Y. Guan and J. Wang, "Uncertainty sets for robust unit commitment," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1439–1440, May 2014.
- [8] R. Jiang, J. Wang, and Y. Guan, "Robust unit commitment with wind power and pumped storage hydro," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 800–810, May 2012.
- [9] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 52–63, Feb. 2013.
- [10] A. Lorca and X. A. Sun, "Adaptive robust optimization with dynamic uncertainty sets for multi-period economic dispatch under significant wind," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1702–1713, 2015.
- [11] W. Wei, F. Liu, S. Mei, and Y. Hou, "Robust energy and reserve dispatch under variable renewable generation," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 369–380, Jan. 2015.
- [12] R. A. Jabr, "Adjustable robust OPF with renewable energy sources," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4742–4751, Nov. 2013.
- [13] S. Pirouzi, J. Aghaei, M. A. Latify, G. R. Yousefi, and G. Mokryani, "A robust optimization approach for active and reactive power management in smart distribution networks using electric vehicles," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2699–2710, Sep. 2018.
- [14] X. Yan, C. Gu, X. Zhang, and F. Li, "Robust optimization-based energy storage operation for system congestion management," *IEEE Syst. J.*, vol. 14, no. 2, pp. 2694–2702, Jun. 2020.
- [15] W. Wei and J. Wang, *Modeling and Optimization of Interdependent Energy Infrastructures*. Berlin, Germany: Springer, Switzerland, 2020.
- [16] A. Lorca, X. A. Sun, E. Litvinov, and T. Zheng, "Multistage adaptive robust optimization for the unit commitment problem," *Oper. Res.*, vol. 64, no. 1, pp. 32–51, Jan. 2016.
- [17] A. Lorca and X. A. Sun, "Multistage robust unit commitment with dynamic uncertainty sets and energy storage," *Comput. Manag. Sci.*, vol. 11, no. 3, pp. 197–220, 2014.
- [18] D. Bertsimas, D. A. Iancu, and P. A. Parrilo, "Optimality of affine policies in multistage robust optimization," *Math. Oper. Res.*, vol. 34, no. 2, pp. 363–394, 2010.
- [19] T. Asamov, D. F. Salas, and W. B. Powell, "SDDP vs. ADP: The effect of dimensionality in multistage stochastic optimization for grid level energy storage," 2016, *arXiv:1605.01521*.
- [20] W. B. Powell, *Approximate Dynamic Programming: Solving the curses of dimensionality*. New York, NY, USA: John Wiley & Sons, 2007.
- [21] G. Hanasusanto and D. Kuhn, "Robust data-driven dynamic programming," *Adv. Neural Inf. Process. Syst.*, Jan. 2013, pp. 827–835.
- [22] N. Zhang, B. D. Leibowicz, and G. A. Hanasusanto, "Optimal residential battery storage operations using robust data-driven dynamic programming," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1771–1780, Mar. 2020.
- [23] J. Nascimento and W. B. Powell, "An optimal approximate dynamic programming algorithm for concave, scalar storage problems with vector-valued controls," *IEEE Trans. Autom. Control*, vol. 58, no. 12, pp. 2995–3010, Dec. 2013.
- [24] P. Zeng, H. Li, H. He, and S. Li, "Dynamic energy management of a micro-grid using approximate dynamic programming and deep recurrent neural network learning," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4435–4445, Jul. 2019.
- [25] X. Xu, C. Lian, L. Zuo, and H. He, "Kernel-based approximate dynamic programming for real-time online learning control: An experimental study," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 1, pp. 146–147, Jan. 2014.
- [26] S. Moazeni, A. H. Miragha, and B. Defourny, "A risk-averse stochastic dynamic programming approach to energy hub optimal dispatch," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 2169–2178, May 2019.
- [27] H. Shuai *et al.*, "Stochastic optimization of economic dispatch for micro-grid based on approximate dynamic programming," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2440–2452, May 2019.
- [28] W. Zhang, *et al.*, "A distributed dynamic programming-based solution for load management in smart grids," *IEEE Syst. J.*, vol. 12, no. 1, pp. 402–413, Mar. 2018.
- [29] M. V. F. Pereira and M. V. G. Pinto, "Multi-stage stochastic optimization applied to energy planning," *Math. Program.*, vol. 52, no. 1–3, pp. 359–375, May 1991.
- [30] R. Lu *et al.*, "Multi-stage stochastic programming to joint economic dispatch for energy and reserve with uncertain renewable energy," *IEEE Trans. Sustain. Energy*, vol. 11, no. 3, pp. 1140–1151, Jul. 2020.
- [31] A. Papavasiliou, Y. Mou, L. Cambier, and D. Scieur, "Application of stochastic dual dynamic programming to the real-time dispatch of storage under renewable supply uncertainty," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 547–558, Apr. 2018.
- [32] Z. Wu, P. Zeng, and X. Zhang, "Two-stage stochastic dual dynamic programming for transmission expansion planning with significant renewable generation and n-k criterion," *CSEE J. Power Energy Syst.*, vol. 2, no. 1, pp. 3–10, 2016.
- [33] A. Bhattacharya, J. P. Kharoufeh, and B. Zeng, "Managing energy storage in microgrids: A multistage stochastic programming approach," *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 483–496, Jan. 2018.
- [34] J. Zou, S. Ahmed, and X. A. Sun, "Multistage stochastic unit commitment using stochastic dual dynamic integer programming," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 1814–1823, May 2019.

- [35] A. B. Philpott, V. L. de Matos, and L. Kapelevich, "Distributionally robust SDDP," *Comput. Manag. Sci.*, vol. 15, no. 3–4, pp. 431–454, 2018.
- [36] J. D. Kozmík, "SDDP for multistage stochastic programs: Preprocessing via scenario reduction," *Math. Program.*, vol. 52, vol. 1–3, pp. 359–375, May 1991.
- [37] V. Guigues and M. Bandarra, "Single cut and multicut SDDP with cut selection for multistage stochastic linear programs: Convergence proof and numerical experiments," 2019, *arXiv:1902.06757*.
- [38] A. Georghiou, A. Tsoukalas, and W. Wiesemann, "Robust dual dynamic programming," *Oper. Res.*, vol. 67, no. 3, pp. 813–830, Oct. 2019.
- [39] K. Abdulla, J. D. Hoog, and K. Steer, "Multi-resolution dynamic programming for the receding horizon control of energy storage," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 333–343, Jan. 2019.
- [40] G. Darivianakis, A. Eichler, R. S. Smith, and J. Lygeros, "A data-driven stochastic optimization approach to the seasonal storage energy management," *IEEE Control Syst. Lett.*, vol. 1, no. 2, pp. 1722–1730, Oct. 2017.
- [41] A. Lorce and X. A. Sun, "Multistage robust unit commitment with dynamic uncertainty sets and energy storage," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 1678–1688, May 2017.
- [42] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations," *Math. Program.*, vol. 171, pp. 115–166, 2017.
- [43] N. Fournier and A. Guillin, "On the rate of convergence in wasserstein distance of the empirical measure," *Probability Theory Related Fields*, vol. 162, no. 3–4, pp. 707–738, Aug. 2015.
- [44] R. Gao and A. J. Kleywegt, "Distributionally robust stochastic optimization with Wasserstein distance," 2016, *arXiv:1604.02199*.
- [45] Z. Guo, "Historical observations of prosumer," 2019 [Online]. Available: <https://github.com/ZhongjieGuo/Papers>
- [46] PJM, "Monthly real-time locational marginal pricing," July 2019. [Online]. Available: <https://www.pjm.com/markets-and-operations/data-dictionary>



Zhongjie Guo received the B.S. degree in electrical engineering from Central South University, Changsha, China, in 2018. He is currently working toward the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China.

His research interests include energy storage, renewable power generation, and power system operation.



Wei Wei (Senior Member, IEEE) received the B.Sc. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2008 and 2013, respectively.

He was a Postdoctoral Research Associate with Tsinghua University from 2013 to 2015. He was a Visiting Scholar with Cornell University, Ithaca, NY, USA, in 2014, and a Visiting Scholar with Harvard University, Cambridge, MA, USA, in 2015. He is currently an Associate Professor with Tsinghua University. His research interests include applied optimization and energy system economics.



Laijun Chen (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2006 and 2011, respectively.

He is currently an Associate Professor with Tsinghua University. His research interests include power system analysis and control, and renewable energy integration.



Zhaojian Wang (Member, IEEE) received the B.S. degree in electrical engineering from Tianjin University, Tianjin, China, in 2013, and the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2018.

He was a Postdoctoral Research Associate with Tsinghua University from 2018 to 2020. From 2016 to 2017, he was a joint Ph.D. student at California Institute of Technology, CA, USA. He is currently an Assistant Professor with Shanghai Jiaotong University, Shanghai, China. His research interests include power system distributed control and optimization.



João P. S. Catalão (Senior Member, IEEE) received the M.Sc. degree in electrical engineering from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 2003, and the Ph.D. and Habilitation degrees in electrical engineering from the University of Beira Interior (UBI), Covilha, Portugal, in 2007 and 2013, respectively.

He is currently a Professor with the Faculty of Engineering, University of Porto (FEUP), Porto, Portugal, and a Research Coordinator at INESC TEC. He was also appointed as a Visiting Professor by North China Electric Power University, Beijing, China. His research interests include power system operations and planning, hydro and thermal scheduling, wind and price forecasting, distributed renewable generation, demand response, and smart grids.



Shengwei Mei (Fellow, IEEE) received the B.Sc. degree in mathematics from Xinjiang University, Urumqi, China, in 1984, the M.Sc. degree in operations research from Tsinghua University, Beijing, China, in 1989, and the Ph.D. degree in automatic control from the Chinese Academy of Sciences, Beijing, in 1996.

He is currently a Professor with Tsinghua University. His research interests include power system complexity and control, game theory, and its application in power systems.