ROBUST OPTIMIZATION WITH CONTINUOUS DECISION-DEPENDENT UNCERTAINTY WITH APPLICATIONS IN DEMAND RESPONSE PORTFOLIO MANAGEMENT

HONGFAN (KEVIN) CHEN *, XU ANDY SUN †, AND HAOXIANG YANG ‡

Abstract. We consider a robust optimization problem with continuous decision-dependent uncertainty (RO-CDDU), which has two new features: an uncertainty set linearly dependent on *continuous* decision variables and a convex piecewise-linear objective function. We prove that RO-CDDU is strongly $\mathcal{NP}$-hard in general and reformulate it into an equivalent mixed-integer nonlinear program (MINLP) with a decomposable structure to address the computational challenges. Such an MINLP model can be further transformed into a mixed-integer linear program (MILP) using extreme points of the dual polyhedron of the uncertainty set. We propose an alternating direction algorithm and a column generation algorithm for RO-CDDU. We model a robust demand response (DR) management problem in electricity markets as RO-CDDU, where electricity demand reduction from users is uncertain and depends on the DR planning decision. Extensive computational results demonstrate the promising performance of the proposed algorithms in both speed and solution quality. The results also shed light on how different magnitudes of decision-dependent uncertainty affect the demand response decision.

Key words. Robust Optimization, Decision-dependent Uncertainty, Demand Response

AMS subject classifications. 90C17, 90C11

1 Introduction

Robust optimization (RO) has emerged as a major modeling framework for decision-making under uncertainty [9]. In a RO model, the decision-maker optimizes the worst-case performance of an objective function within an uncertainty set. Often the RO problem is a semi-infinite program, which can be reformulated as the finite-dimensional *robust counterpart*. We can classify uncertainty models into decision-independent and decision-dependent ones. The decision-independent uncertainty, called *exogenous* uncertainty, has been discussed extensively in the literature [10; 11; 12]. As stated in [9], for many types of convex uncertainty sets independent of decisions, the RO model admits a computationally tractable robust counterpart.

Recently more theoretical developments have focused on the RO formulation with *decision-dependent uncertainty sets* [35], which admits a wide range of applications in pricing, scheduling, and electricity demand response [27; 48]. In this paper, we consider a class of mixed-integer robust optimization models with a continuous decision-dependent uncertainty set (RO-CDDU), which contains two features: (i) the uncertainty set depends on the continuous decision variables, and (ii) the objective function is piecewise-linear convex. We formulate the RO-CDDU model as follows:

$$\min_{\boldsymbol{x},\boldsymbol{y}} \max_{\boldsymbol{\xi}\in\Xi(\boldsymbol{x})} \max_{k=1,\dots,K} f_k(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\xi}) \tag{1.1a}$$

*Business School, The Chinese University of Hong Kong, Ma Liu Shui, Hong Kong, China (kevinchen@cuhk.edu.hk),

†Sloan School of Management, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02142 (sunx@mit.edu) corresponding author.

‡School of Data Science, The Chinese University of Hong Kong-Shenzhen, Shenzhen, China (yanghaoxiang@cuhk.edu.cn).

$$\text{s.t.} \quad (\boldsymbol{x}, \boldsymbol{y}) \in \Omega, \tag{1.1b}$$

$$\boldsymbol{x} \in \mathbb{R}^{n_x}, \ \boldsymbol{y} \in \mathbb{Z}^{n_y}. \tag{1.1c}$$

In problem (1.1), the feasibility set Ω is a polyhedron defined by m inequalities such that $\Omega = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{n_x+n_y} : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} \leq \boldsymbol{r}\}$. The uncertainty set $\Xi(\boldsymbol{x})$ is a polyhedron defined by l inequalities: $\Xi(\boldsymbol{x}) = \{\boldsymbol{\xi} \in \mathbb{R}^{n_\xi} : \boldsymbol{W}\boldsymbol{\xi} \leq \boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}\}$. The total number of pieces in the objective function is $K \in \mathbb{N}_+$, and the k-th piece $f_k(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$ is a linear function $f_k(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}) = \boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + \boldsymbol{c}_k^\top \boldsymbol{\xi} + d_k$. The piecewise linear convex objective function has been widely used in robust optimization applications, such as robust queuing networks [5; 8; 49], operating room scheduling [6], and inventory management [7; 31; 45]. In this paper, the piecewise linear objective function is motivated by different marginal costs for over- and under-commitment in an electricity market demand response application with details in Section 4. Model (1.1) returns a decision profile $(\boldsymbol{x}, \boldsymbol{y})$ that minimizes the worst-scenario cost given the uncertainty set. Here the RO-CDDU model (1.1) is different from the RO model with exogenous uncertainty, as the uncertainty set $\Xi(\boldsymbol{x})$ depends on the continuous decision $\boldsymbol{x}$.

The literature has extensively discussed robust optimization problems with decision-dependent uncertainty (RO-DDU). Reference [35] establishes that a robust linear optimization problem with the uncertainty set dependent on decision variables is $\mathcal{NP}$-hard by constructing a polynomial reduction from the 3-SAT problem. Reference [44] considers a software partitioning problem to minimize the run-time of a computer program, in which the scheduling of code execution depends on binary assignment decisions. Reference [38] extends the budget uncertainty set of [12] by allowing the protection level to be dependent on binary decision variables. Reference [48] proposes a decision-dependent uncertainty set as a Minkowski sum of static uncertainty sets. Reference [39] proposes a $(1 + \varepsilon)$-approximation algorithm for the robust optimization problem with a knapsack uncertainty set. Reference [28] generalizes the dependency from binary decision variables to general discrete ones. The uncertainty set dependent on discrete decisions with finite dimensions admits a computationally tractable robust counterpart that can be represented as a finite union of convex sets. Our work establishes that RO-CDDU is strongly $\mathcal{NP}$-hard and characterizes the structure of the adversary's problem that depends on continuous decisions in our algorithm design.

Another stream of research focuses on endogenous uncertainty in distributionally robust optimization settings, in which the ambiguity set characterized by the probabilistic distributions depends on the previous stages' decisions. For example, Reference [30] explores multiple types of ambiguity sets based on moments, covariance matrix, Wasserstein metric, Phi-divergence, and Kolmogorov–Smirnov test, for which they derive tractable dual reformulations. Reference [36] develops tractable formulations for ambiguity sets based on similar statistical distances. Reference [51] has a decision-dependent moment-based ambiguity set, and the formulation is extended to a multi-stage setting. However, those distributionally robust optimization models still require an estimation of the ambiguity set to compute the expectation based on the worst-case probability distribution, which may not satisfy the robustness requirement in some low-probability high-impact applications [54].

The formulation of the RO-CDDU model is motivated by the demand response management in electricity markets [1]. As the internet-of-things (IoT) and smart grid technologies develop, an increasing number of electric appliances, including air conditioners and space heaters in residential and commercial buildings, are eligible for real-time control. This allows flexible electric loads in different locations to be aggregated into a sizable portfolio of demand response (DR) resources. A company that creates and manages such a portfolio is called a DR aggregator, which balances supply and demand in electricity markets by adjusting DR resources' loads. DR aggregators constantly face issues of uncertainty in DR resources [37]: a DR resource commits to reducing its electricity consumption by a certain amount for a given time period, but the actual reduction can deviate from such a commitment and the deviation often depends on the committed reduction amount. If mishandled, this uncertainty can cause significant load shedding and financial loss. Therefore, we propose an RO-CDDU model, utilizing a convex piecewise-linear function to realistically model electric power generation cost functions [50], and develop computationally tractable algorithms for a DR aggregator to manage their large portfolios of DR resources.

We summarize the main contributions of this paper below.

1. We formulate the RO-CDDU model (1.1) and establish that RO-CDDU in a general form is strongly $\mathcal{NP}$-hard.
2. We establish that problem (1.1) has an equivalent decomposable formulation with an uncertainty set specific to each piece of the linear function.
3. We derive an MINLP formulation for RO-CDDU, and pose two assumptions on the dual polyhedron such that RO-CDDU is well-defined. Under those assumptions, we reformulate RO-CDDU into an MILP using the extreme points of the dual polyhedron. We characterize cases for RO-CDDU to be solvable in polynomial time even when the dual polyhedron has an exponential number of extreme points, and in addition, we develop two computationally efficient algorithms to numerically solve RO-CDDU.
4. We propose a novel RO-CDDU model for a demand response management problem in electricity markets. We present extensive computational experiments on our proposed algorithms to analyze the robust solution's properties.

The paper is organized as follows. In Section 2, we prove that the RO-CDDU problem is strongly $\mathcal{NP}$-hard. In Section 3, we discuss model reformulation and algorithm design. More specifically, in Section 3.1, we provide an exact MILP formulation for the RO-CDDU problem, and characterize the model reformulation for widely-studied uncertainty sets. We propose an alternating direction algorithm (ADA) and a column generation algorithm (CGA) in Section 3.2, and the McCormick relaxation for a lower bound of RO-CDDU in Section 3.3. In Section 4, we discuss the application of our model in a demand response scheduling problem in electricity markets and report the performance of the computational experiments. Section 5 concludes the paper with a summary and future directions of RO-CDDU.

2 Computational Complexity

We are interested in whether the RO-CDDU problem could be solved polynomially in $\mathcal{O}(n_x^{\alpha_1} n_y^{\alpha_2} m^{\alpha_3} n_\xi^{\alpha_4} l^{\alpha_5})$ steps for some $\alpha_i \geq 0$ with $i = 1, \ldots 5$. Besides the computational challenges caused by integer variables, it remains to show if the continuous decision-dependent uncertainty set makes the problem hard to solve. Using a polynomial reduction from the 3-partition problem, we prove that RO-CDDU is strongly $\mathcal{NP}$-hard, even with no integer decision variables.

THEOREM 2.1. *For any $n_y \in \mathbb{N}$ (including $n_y = 0$), the RO-CDDU problem in (1.1) is strongly $\mathcal{NP}$-hard.*

Proof of Theorem 2.1. To prove that model (1.1) is strongly $\mathcal{NP}$-hard for any

$n_y \in \mathbb{N}$, we consider a problem instance of (1.1) with $n_x = KN_x$ for some $N_x \in \mathbb{N}_+$, and for each $k = 1, \ldots, K$ we set the objective function $f_k(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$ as:

$$f_k(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi}) = -\sum_{i=1}^{N_x} \omega_i x_{ik} - \sum_{j=1}^{n_y} \nu_j y_j + \sum_{\ell=1}^{N_x} \xi_{\ell k}.$$

We define feasible region Ω in (1.1b) as $\Omega = \{(\boldsymbol{x}, \boldsymbol{y}) \in [0,1]^{KN_x+n_y} : \sum_{i=1}^{N_x} \omega_i x_{ik} + \sum_{j=1}^{n_y} \nu_j y_j = W, \sum_{i=1}^{N_k} x_{ik} = 3, \sum_{k=1}^{K} x_{ik} = 1\}$ with $\omega_i > 0$ for all $i = 1, \ldots N_x$ and $\nu_j > W > 0$ for all $j = 1, \ldots, n_y$, and the uncertainty set $\Xi(\boldsymbol{x})$ as $\Xi(\boldsymbol{x}) = \times_{i=1}^{N_x} \times_{k=1}^{K} \Xi_{ik}(x_{ik})$, where $\Xi_{ik}(x_{ik}) = \{\xi_{ik} \in \mathbb{R} : \xi_{ik} \leq x_{ik},\ \xi_{ik} \leq 1 - x_{ik}\}$ for $i = 1, \ldots, N_x$ and $k = 1, \ldots, K$. We can rewrite model (1.1) as:

$$\min_{V,\boldsymbol{x},\boldsymbol{y}} \quad V \tag{2.1a}$$

$$\text{s.t.} \quad V \geq \sum_{i=1}^{N_x} \max_{\boldsymbol{\xi}_{ik} \in \Xi_{ik}(x_{ik})} \{\xi_{ik}\} - \sum_{i=1}^{N_x} \omega_i x_{ik} - \sum_{j=1}^{n_y} \nu_j y_j, \quad \forall k = 1, \ldots, K, \tag{2.1b}$$

$$\sum_{i=1}^{N_x} \omega_i x_{ik} + \sum_{j=1}^{n_y} \nu_j y_j = W, \qquad \forall k = 1, \ldots, K, \tag{2.1c}$$

$$\sum_{i=1}^{N_k} x_{ik} = 3, \qquad \forall k = 1, \ldots, K, \tag{2.1d}$$

$$\sum_{k=1}^{K} x_{ik} = 1, \qquad \forall i = 1, \ldots, N_x, \tag{2.1e}$$

$$0 \leq x_{ik} \leq 1, \qquad \forall i = 1, \ldots, N_x, k = 1, \ldots, K, \tag{2.1f}$$

$$0 \leq y_j \leq 1,\ y_j \in \mathbb{Z}, \qquad \forall j = 1, \ldots, n_y. \tag{2.1g}$$

The objective function (2.1a) and constraint (2.1b) together reformulate the objective function in (1.1a). We observe that $\max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi} = \sum_{i=1}^{n_x} \max_{\boldsymbol{\xi}_{ik} \in \Xi_{ik}(x_{ik})} \xi_{ik}$ since the uncertainty set $\Xi_{ik}(x_{ik})$ is separable for $k = 1, \ldots, K$. By definition, we can establish $\max_{\boldsymbol{\xi}_{ik} \in \Xi_{ik}(x_{ik})} \{\xi_{ik}\} = \min\{x_{ik}, 1 - x_{ik}\}$. Constraints (2.1c)-(2.1f) characterize the feasible region Ω. Constraint (2.1g) specifies bounds and integrality for variables $\boldsymbol{y}$ in constraint (1.1c). Given that $\nu_j > W$ for all $j = 1, \ldots, n_y$, any feasible solution should satisfy $y_j = 0$ and we can omit y in our formulation.

We denote the decision problem associated with model (2.1) as $\mathcal{Q}$, in which we decide if there exists a feasible solution $(V, \boldsymbol{x}, \boldsymbol{y})$ such that $V = -W$. Computing $\sum_{i=1}^{N_x} \max_{\boldsymbol{\xi}_{ik} \in \Xi_{ik}(x_{ik})} \{\xi_{ik}\}$ and $\sum_{i=1}^{N_x} \omega_i x_{ik}$ takes polynomial time in the size of input, so the decision problem is in $\mathcal{NP}$. We next establish a polynomial reduction from a 3-partition problem, $\mathcal{Q}_{3par}$, to $\mathcal{Q}$ by verifying that the answer to $\mathcal{Q}$ is "yes" if and only if the answer to $\mathcal{Q}_{3par}$ is "yes". The 3-partition problem $\mathcal{Q}_{3par}$ asks if there exists a partition of set $\mathcal{S}$ into triplets for $\mathcal{S} = \mathcal{S}_1 \cup \cdots \cup \mathcal{S}_K$ with $|\mathcal{S}_k| = 3$ for all $k = 1, \ldots, K$ and $\mathcal{S}_k \cap \mathcal{S}_{k'} = \emptyset$ for all $k \neq k'$ such that $\sum_{\omega \in \mathcal{S}_k} \omega = W$ for each $k = 1, \ldots, K$.

($\Longrightarrow$) Suppose the answer to $\mathcal{Q}$ is "yes", i.e., there exists a feasible solution $(V, \boldsymbol{x}, \boldsymbol{y})$ such that $V = -W$. We can derive the following inequalities:

$$V \overset{(a)}{\geq} \sum_{i=1}^{N_x} \min\{1 - x_{ik}, x_{ik}\} - \sum_{i=1}^{N_x} \omega_i x_{ik} \overset{(b)}{=} \sum_{i=1}^{n_x} \min\{1 - x_{ik}, x_{ik}\} - W \overset{(c)}{\geq} -W, \tag{2.2}$$

where step (a) follows directly from constraint (2.1b), and step (b) follows directly from constraint (2.1c). In step (c), constraint (2.1f) suggests that $\sum_{i=1}^{N_x} \min\{1-x_{ik}, x_{ik}\} \geq 0$. Since $V = -W$, every inequality in (2.2) holds as an equality. From step (b) we have $\sum_{i=1}^{N_x} \omega_i x_{ik} = W$. Since step (c) holds as an equality, for each $i = 1, \ldots, N_x$, we either have $1-x_{ik} = 0$ or $x_{ik} = 0$, i.e., $x_{ik} \in \{0,1\}$. We set $\mathcal{S}_k = \{\omega_i : x_{ik} = 1\}, \forall k = 1, \ldots, K$. Each $\mathcal{S}_k$ forms a triplet by constraint (2.1d) and every element ω_i can find a unique triplet assignment by constraint (2.1e). The sum of elements of each triplet equals W by constraint (2.1c) and we obtain a solution to $\mathcal{Q}_{3par}$.

($\Longleftarrow$) Suppose that the answer to $\mathcal{Q}_{3par}$ is "yes", which implies that there exists a partition of set $\mathcal{S}$ into triplets $\mathcal{S}_1, \ldots, \mathcal{S}_K$ such that $\sum_{\omega \in \mathcal{S}_k} \omega = W$ for all $k = 1, \ldots, K$. We can then construct a tuple $(V, \boldsymbol{x}, \boldsymbol{y})$ as:

$$x_{ik} = \begin{cases} 1 & \text{if } \omega_i \in \mathcal{S}_k \\ 0 & \text{otherwise} \end{cases} \forall i = 1, \ldots, N_x, k = 1, \ldots, K, \quad y_j = 0,\ \forall j = 1, \ldots, n_y, \quad V = -W,$$

which is feasible for model (2.1), and thus we can answer "yes" to $\mathcal{Q}$.

In summary, we establish a polynomial reduction from $\mathcal{Q}_{3par}$ to $\mathcal{Q}$. Since $\mathcal{Q}_{3par}$ is strongly $\mathcal{NP}$-complete, the decision problem $\mathcal{Q}$ is also strongly $\mathcal{NP}$-complete and the optimization problem RO-CDDU is strongly $\mathcal{NP}$-hard for all $n_y \in \mathbb{N}$. ■

Theorem 2.1 suggests that the uncertainty set's dependency on continuous decisions makes RO-CDDU model (1.1) strongly $\mathcal{NP}$-hard. This strongly $\mathcal{NP}$-hardness also leads to the result that RO-CDDU does not admit a fully polynomial-time approximation scheme (FPTAS) unless $\mathcal{P} = \mathcal{NP}$ [22]. Note that our complexity result still holds when there is no integer variable, i.e., $n_y = 0$, or when $\boldsymbol{x}$ is integer.

To improve the computational tractability of RO-CDDU, we first establish a reformulation of model (1.1) to the following model with a decomposable structure:

$$\begin{aligned}
&(2.3\text{a}) && \min_{V, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}} && V && \\
&(2.3\text{b}) && \text{s.t.} && V \geq \boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + z_k + d_k, && \forall k = 1, \ldots, K, \\
&(2.3\text{c}) && && z_k \geq \max_{\boldsymbol{\xi}_k \in \Xi(\boldsymbol{x})} \{\boldsymbol{c}_k^\top \boldsymbol{\xi}_k\}, && \forall k = 1, \ldots, K, \\
&(2.3\text{d}) && && (\boldsymbol{x}, \boldsymbol{y}) \in \Omega,\ \boldsymbol{x} \in \mathbb{R}^{n_x},\ \boldsymbol{y} \in \mathbb{Z}^{n_y}. &&
\end{aligned}$$

We summarize the connections between model (1.1) and (2.3) in Proposition 2.2 below.

PROPOSITION 2.2. *The RO-CDDU problem in* (1.1) *has the same optimal value as model* (2.3). *Any optimal solution to model* (2.3) *is also optimal to model* (1.1).

Proof of Proposition 2.2. We first add an auxiliary variable V to represent the objective function of the RO-CDDU model in (1.1), and then lift the decision space into $(V, \boldsymbol{x}, \boldsymbol{y})$ to obtain the following equivalent formulation:

$$\begin{aligned}
&(2.4\text{a}) && \min_{V, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}} && V \\
&(2.4\text{b}) && \text{s.t.} && V = \max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{x})} \max_{k=1,\ldots,K} \{\boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + \boldsymbol{c}_k^\top \boldsymbol{\xi} + d_k\}, \\
&(2.4\text{c}) && && (\boldsymbol{x}, \boldsymbol{y}) \in \Omega,\ \boldsymbol{x} \in \mathbb{R}^{n_x},\ \boldsymbol{y} \in \mathbb{Z}^{n_y}.
\end{aligned}$$

Next, we show that problem (2.4) is equivalent to problem (2.5) below:

$$(2.5\text{a}) \quad \min_{V, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}} \quad V$$

$$\text{s.t.} \quad V \geq \max_{k=1,\ldots,K} \left\{ \boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + z_k + d_k \right\}, \tag{2.5b}$$

$$z_k \geq \max_{\boldsymbol{\xi}_k \in \Xi(\boldsymbol{x})} \left\{ \boldsymbol{c}_k^\top \boldsymbol{\xi}_k \right\}, \qquad \forall k = 1, \ldots, K, \tag{2.5c}$$

$$(\boldsymbol{x}, \boldsymbol{y}) \in \Omega, \ \boldsymbol{x} \in \mathbb{R}^{n_x}, \ \boldsymbol{y} \in \mathbb{Z}^{n_y}. \tag{2.5d}$$

To establish the claim, we prove the equivalence between model (2.4) and model (2.5). First, for problem (2.4), we consider the corresponding optimal solution $(\boldsymbol{x}', \boldsymbol{y}', \boldsymbol{z}', V_1^*)$ where $\boldsymbol{\xi}'$ satisfies that $\boldsymbol{\xi}' \in \arg\max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{x}')} \max_{k=1,\ldots,K} \left\{ \boldsymbol{a}_k^\top \boldsymbol{x}' + \boldsymbol{b}_k^\top \boldsymbol{y}' + \boldsymbol{c}_k^\top \boldsymbol{\xi} + d_k \right\}$. We then let the optimal solution to problem (2.5) be $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{z}^*, V_2^*)$ where $\boldsymbol{\xi}_k^* \in \arg\max_{\boldsymbol{\xi}_k \in \Xi(\boldsymbol{x}^*)} \left\{ \boldsymbol{c}_k^\top \boldsymbol{\xi}_k \right\}$ for all $k = 1, \ldots, K$. At optimality, the index k^* represents the piece where $\max_{k=1,\ldots,K} \left\{ \boldsymbol{a}_k^\top \boldsymbol{x}^* + \boldsymbol{b}_k^\top \boldsymbol{y}^* + z_k^* + d_k \right\}$ is achieved.

($\Longrightarrow$): We first prove that the optimal value V_2^* of model (2.5) is greater than or equal to the optimal value V_1^* of model (2.4). To establish the claim, we show that

$$V_2^* \overset{(a)}{\geq} \max_{k=1,\ldots,K} \boldsymbol{a}_k^\top \boldsymbol{x}^* + \boldsymbol{b}_k^\top \boldsymbol{y}^* + \boldsymbol{c}_k^\top \boldsymbol{z}_k^* + d_k \tag{2.6a}$$

$$\geq \max_{k=1,\ldots,K} \boldsymbol{a}_k^\top \boldsymbol{x}^* + \boldsymbol{b}_k^\top \boldsymbol{y}^* + \max_{\boldsymbol{\xi}_k \in \Xi(\boldsymbol{x}^*)} \left\{ \boldsymbol{c}_k^\top \boldsymbol{\xi}_k \right\} + d_k \tag{2.6b}$$

$$\overset{(b)}{\geq} \max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{x}^*)} \max_{k=1,\ldots,K} \boldsymbol{a}_k^\top \boldsymbol{x}^* + \boldsymbol{b}_k^\top \boldsymbol{y}^* + \boldsymbol{c}_k^\top \boldsymbol{\xi} + d_k \tag{2.6c}$$

$$\overset{(c)}{\geq} \min_{\substack{\boldsymbol{x}, \boldsymbol{y} \in \Omega \\ \boldsymbol{x} \in \mathbb{R}^{n_x}, \boldsymbol{y} \in \mathbb{Z}^{n_y}}} \max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{x})} \max_{k=1,\ldots,K} \boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + \boldsymbol{c}_k^\top \boldsymbol{\xi} + d_k \ = \ V_1^*, \tag{2.6d}$$

where step (a) follows directly from constraints (2.5b) and (2.5c) in model (2.5). In step (b), the inequality holds because we can view the function in (2.6c) as the function in (2.6b) with additional constraints $\boldsymbol{\xi}_k = \boldsymbol{\xi}$ for all $k = 1, \ldots, K$, which enables us to move the maximization operator over $\boldsymbol{\xi}$ outside $\max_{k=1,\ldots,K}$. In step (c), the inequality follows given that $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is only a feasible solution to the minimization problem in (2.6d) with an optimal objective value V_1^*, which also leads to the last equality. Summarizing the observations above, we obtain that $V_2^* \geq V_1^*$.

($\Longleftarrow$): To establish $V_1^* \geq V_2^*$, we deduce that

$$V_1^* \overset{(d)}{=} \max_{k=1,\ldots,K} \left\{ \boldsymbol{a}_k^\top \boldsymbol{x}' + \boldsymbol{b}_k^\top \boldsymbol{y}' + \boldsymbol{c}_k^\top \boldsymbol{\xi}' + d_k \right\} \tag{2.7a}$$

$$\overset{(e)}{\geq} \max_{k=1,\ldots,K} \left\{ \boldsymbol{a}_k^\top \boldsymbol{x}' + \boldsymbol{b}_k^\top \boldsymbol{y}' + \max_{\boldsymbol{\xi}_k \in \Xi(\boldsymbol{x}')} \left\{ \boldsymbol{c}_k^\top \boldsymbol{\xi}_k \right\} + d_k \right\} \tag{2.7b}$$

$$\overset{(f)}{\geq} \max_{k=1,\ldots,K} \left\{ \boldsymbol{a}_k^\top \boldsymbol{x}^* + \boldsymbol{b}_k^\top \boldsymbol{y}^* + \max_{\boldsymbol{\xi}_k \in \Xi(\boldsymbol{x}^*)} \left\{ \boldsymbol{c}_k^\top \boldsymbol{\xi}_k \right\} + d_k \right\} \overset{(g)}{=} V_2^*, \tag{2.7c}$$

where in step (d), we plug in the optimal solution $(\boldsymbol{x}', \boldsymbol{y}', V_1^*)$ from model (2.4). We let k' denote the index where the expression in (2.7a) achieves the maximum and k'' denotes the index where the expression in (2.7b) achieves the maximum. We can prove step (e) by discussing the following two scenarios: (i) if $k' = k''$, then we observe that the inequality holds with equality given that $\boldsymbol{\xi}' \in \arg\max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{x}')} \left\{ \boldsymbol{a}_{k'}^\top \boldsymbol{x}' + \boldsymbol{b}_{k'}^\top \boldsymbol{y}' + \boldsymbol{c}_{k'}^\top \boldsymbol{\xi} + d_{k'} \right\}$; (ii) if $k' \neq k''$, we can show the inequality by contradiction: assuming the inequality in (2.7b) does not hold, we obtain that

$$\boldsymbol{a}_{k''}^\top \boldsymbol{x}' + \boldsymbol{b}_{k''}^\top \boldsymbol{y}' + \boldsymbol{c}_{k''}^\top \boldsymbol{\xi}_{k''} + d_{k''} > \max_{k=1,\ldots,K} \left\{ \boldsymbol{a}_k^\top \boldsymbol{x}' + \boldsymbol{b}_k^\top \boldsymbol{y}' + \boldsymbol{c}_k^\top \boldsymbol{\xi}' + d_k \right\}.$$

However, this contradicts that $\boldsymbol{\xi}'$ maximizes the expression in (2.7a), since $\boldsymbol{\xi}_{k''}$ is a feasible solution and achieves a larger value for (2.7a). Therefore, the inequality in step (e) holds. Step (f) follows from the optimality of solution $(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{z}^*, V_2^*)$ to model (2.5). Step (g) matches the definition of V_2^*. Therefore, $V_1^* \geq V_2^*$.

Summarizing the two arguments above, we obtain that $V_1^* = V_2^*$. Furthermore, $(\boldsymbol{x}^*, \boldsymbol{y}^*)$ is an optimal solution to RO-CDDU model (1.1) based on the observation in step (c) that the solution is feasible, which also achieves the optimal value given that $V_2^* = V_1^*$. This completes the proof of the claims in this result. ■

Proposition 2.2 implies that we can solve model (2.3) instead of model (1.1) without loss of optimality. Such a reformulation is important as we can obtain a decomposable structure from model (2.3), while it is hard to do so for model (1.1). We will explain this structure with more details in Section 3 and consider model (2.3) in the discussion of RO-CDDU for the rest of the paper.

3 Model Reformulations and Algorithms

In the RO-CDDU model (1.1), the adversarial variable $\boldsymbol{\xi}$ influences the value of the piecewise-linear objective function $\max_{k=1,\ldots,K} f_k(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$. From Proposition 2.2, we observe that model (2.3), equivalent to model (1.1), allows us to establish a decomposable structure as in constraint (2.3c), such that the adversarial variable $\boldsymbol{\xi}_k$ is specific to each linear function $f_k(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\xi})$. However, even with a decomposable structure, Proposition 2.2 presents the fundamental challenge of solving the RO-CDDU problem: model (2.3) is a semi-infinite mixed-integer program, and the standard robust counterpart reformulation in Theorem 1.3.4 of [9] cannot be directly applied due to the uncertainty set's dependency on continuous decision variables. Another computational challenge is that the set of constraints in (2.3c) is nonconvex in decision $\boldsymbol{x}$. Moreover, since the decision vector $\boldsymbol{x}$ is continuous, we can neither directly apply the reformulation techniques in [35].

To address the issues above, we reformulate model (2.3) as the following MINLP, using the strong duality result in Theorem 1.3.4 of [9]:

$$
\begin{aligned}
&(3.1a) && \min_{V,\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{\pi}} \quad V \\
&(3.1b) && \text{s.t.} \quad V \geq z_k + \boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + d_k, && \forall k = 1,\ldots,K, \\
&(3.1c) && \qquad z_k \geq \boldsymbol{\pi}_k^\top (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}), && \forall k = 1,\ldots,K, \\
&(3.1d) && \qquad \boldsymbol{W}^\top \boldsymbol{\pi}_k = \boldsymbol{c}_k, \ \boldsymbol{\pi}_k \geq \boldsymbol{0} && \forall k = 1,\ldots,K, \\
&(3.1e) && \qquad (\boldsymbol{x}, \boldsymbol{y}) \in \Omega, \ \boldsymbol{x} \in \mathbb{R}^{n_x}, \ \boldsymbol{y} \in \mathbb{Z}^{n_y}.
\end{aligned}
$$

The bilinear terms $\boldsymbol{\pi}_k^\top \boldsymbol{T}\boldsymbol{x}$ make model (3.1) computationally challenging. In Section 3.1, by defining the dual polyhedron $\mathcal{H}_k := \{\boldsymbol{\pi} \in \mathbb{R}^l : \boldsymbol{W}^\top \boldsymbol{\pi} = \boldsymbol{c}_k, \ \boldsymbol{\pi} \geq \boldsymbol{0}\}$ from constraint (3.1d), we first state two assumptions based on the structure of $\mathcal{H}_k$ such that the RO-CDDU problem is well-defined. Under those assumptions, we can further reformulate model (3.1) as an MILP using the extreme points of $\mathcal{H}_k$. However, the MILP model is still large-scale and hard to solve. Therefore, we propose an alternating direction algorithm (ADA) in Section 3.2.1 and a column generation algorithm (CGA) in Section 3.2.2 to obtain a good feasible solution efficiently. In Section 3.3, we consider an approximation model based on McCormick relaxation to obtain the lower bound.

3.1 MILP Reformulation Based on the Structure of $\mathcal{H}_k$

We note that if the dual polyhedron $\mathcal{H}_k$ is an empty set, the adversary's problem $\max_{\boldsymbol{\xi}_k \in \Xi(\boldsymbol{x})} \left\{\boldsymbol{c}_k^\top \boldsymbol{\xi}_k\right\}$

in constraint (2.3c) of problem (2.3) is either infeasible or unbounded by the linear programming duality theory, subject to which the RO-CDDU problem becomes ill-defined. Thus, we first make the following assumption on $\mathcal{H}_k$.

ASSUMPTION 1. $\mathcal{H}_k \neq \emptyset, \forall k = 1, \ldots, K$.

Since $\mathcal{H}_k$ is contained in $\mathbb{R}^l_+$ and by Assumption 1 it is non-empty, thus it has an extreme point. So, by Minkowski-Weyl Theorem, we can explicitly represent the dual polyhedron $\mathcal{H}_k$ as:

$$\text{(3.2)}\quad \mathcal{H}_k = \left\{ \boldsymbol{\pi} = \sum_{s=1}^{N_k} w^0_s \overline{\boldsymbol{\pi}}_{ks} + \sum_{r=1}^{M_k} w^1_r \overline{\boldsymbol{\lambda}}_{kr} :\ \sum_{s=1}^{N_k} w^0_s = 1,\ \boldsymbol{w}^0 \in \mathbb{R}^{N_k}_+,\ \boldsymbol{w}^1 \in \mathbb{R}^{M_k}_+ \right\},$$

where $\{\overline{\boldsymbol{\pi}}_{ks}\}_{s=1}^{N_k}$ denotes a finite set of points and $\{\overline{\boldsymbol{\lambda}}_{kr}\}_{r=1}^{M_k}$ denotes a finite set of rays in $\mathcal{H}_k$, with finite N_k and M_k for all $k = 1, \ldots, K$. The dual polyhedron $\mathcal{H}_k$ is pointed because $\boldsymbol{\pi} \geq \mathbf{0}$, which suggests that it is without loss of generality to let $\{\overline{\boldsymbol{\pi}}_{ks}\}_{s=1}^{N_k}$ be the set of extreme points and $\{\overline{\boldsymbol{\lambda}}_{kr}\}_{r=1}^{M_k}$ be the set of extreme rays for $\mathcal{H}_k$. With this representation, we further make an assumption on $\mathcal{H}_k$ to make model (1.1) well-defined.

ASSUMPTION 2. *$\overline{\boldsymbol{\lambda}}_{kr}^\top (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}) \geq 0$ for any $k = 1, \ldots, K, r = 1, \ldots, M_k$ and any $(\boldsymbol{x}, \boldsymbol{y}) \in \Omega$,*

For Assumption 2, if for some solution $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}) \in \Omega$ we can find a ray $\overline{\boldsymbol{\lambda}}_{kr}$ such that $\overline{\boldsymbol{\lambda}}_{kr}^\top (\boldsymbol{h} - \boldsymbol{T}\hat{\boldsymbol{x}}) < 0$, the adversary's problem (2.3c) is infeasible because its dual problem is unbounded, i.e., a decision $\hat{\boldsymbol{x}}$ can be made such that $\Xi(\hat{\boldsymbol{x}}) = \emptyset$. For RO-CDDU, though decision-dependent, uncertainty should objectively exist and not be eliminated by the decision. Therefore, we propose Assumption 2 to avoid such an unreasonable situation, which also matches the real-world setups in the demand response management problem introduced in Section 4.

Assumptions 1 and 2 ensure that the adversary's problem $\max_{\boldsymbol{\xi}_k \in \Xi(\boldsymbol{x})} \left\{ \boldsymbol{c}_k^\top \boldsymbol{\xi}_k \right\}$ in constraint (2.3c) is neither unbounded nor infeasible, which are commonly recognized conditions for decision-independent uncertainty sets [13; 29]. To proceed, we consider the following characterization of the RO-CDDU problem.

PROPOSITION 3.1. *Suppose $\{\overline{\boldsymbol{\pi}}_{ks}\}_{s=1}^{N_k}$ and $\{\overline{\boldsymbol{\lambda}}_{kr}\}_{r=1}^{M_k}$ are respectively the extreme points and extreme rays of $\mathcal{H}_k$ given in (3.2) for all $k = 1, \ldots, K$, Assumptions 1 and 2 hold, and Ω is compact. The RO-CDDU problem (1.1) can be reformulated as the following MILP:*

$$\begin{aligned}
&\text{(3.3a)} && \min_{V, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{\mu}} \quad V \\
&\text{(3.3b)} && \quad s.t. \quad V \geq z_k + \boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + d_k, && \forall k = 1, \ldots, K, \\
&\text{(3.3c)} && \qquad\qquad z_k \geq \overline{\boldsymbol{\pi}}_{ks}^\top (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}) - M(1 - \mu_{ks}), && \forall k = 1, \ldots, K, s = 1, \ldots, N_k, \\
&\text{(3.3d)} && \qquad\qquad \sum_{s=1}^{N_k} \mu_{ks} = 1, \quad \boldsymbol{\mu}_k \in \{0,1\}^{N_k}, && \forall k = 1, \ldots, K, \\
&\text{(3.3e)} && \qquad\qquad (\boldsymbol{x}, \boldsymbol{y}) \in \Omega,\ \boldsymbol{x} \in \mathbb{R}^{n_x},\ \boldsymbol{y} \in \mathbb{Z}^{n_y}.
\end{aligned}$$

Proof of Proposition 3.1. For any feasible solution $\boldsymbol{x}$, by LP strong duality, the optimal value of $\max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi}$ on the right hand side of constraint (2.3c) equals to the optimal value of its dual problem $\min_{\boldsymbol{\pi}_k \in \mathcal{H}_k} \boldsymbol{\pi}_k^\top(\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x})$. Thus, model (2.3) is equivalent to the following formulation:

$$\begin{aligned}
&\text{(3.4a)} && \min_{V,\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{\pi}} \quad V \\
&\text{(3.4b)} && \text{s.t.} \quad V \ \geq\ z_k + \boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + d_k, && \forall k = 1, \ldots, K, \\
&\text{(3.4c)} && \qquad\quad z_k \ \geq\ \min_{\boldsymbol{\pi}_k \in \mathcal{H}_k} \boldsymbol{\pi}_k^\top(\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}), && \forall k = 1, \ldots, K, \\
&\text{(3.4d)} && \qquad\quad (\boldsymbol{x}, \boldsymbol{y}) \in \Omega,\ \boldsymbol{x} \in \mathbb{R}^{n_x},\ \boldsymbol{y} \in \mathbb{Z}^{n_y}.
\end{aligned}$$

For each $k = 1, \ldots, K$, by the representation of $\mathcal{H}_k$ in (3.2), we can write the minimization problem in constraint (3.4c) in an equivalent form:

$$\begin{aligned}
&\text{(3.5a)} && \min_{\boldsymbol{w}^0, \boldsymbol{w}^1} \quad \Big(\sum_{s=1}^{N_k} w_s^0 \overline{\boldsymbol{\pi}}_{ks} + \sum_{r=1}^{M_k} w_r^1 \overline{\boldsymbol{\lambda}}_{kr}\Big)^\top (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}) \\
&\text{(3.5b)} && \text{s.t.} \quad \sum_{s=1}^{N_k} w_s^0 = 1,\ \boldsymbol{w}^0 \in \mathbb{R}_+^{N_k},\ \boldsymbol{w}^1 \in \mathbb{R}_+^{M_k}.
\end{aligned}$$

By Assumption 2, at the optimal solution, we have $w_r^1 = 0$ for any $r = 1, \ldots, M_k$ because $\overline{\boldsymbol{\lambda}}_{kr}^\top(\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}) \geq 0$. Thus, constraint (3.4c) can be reformulated as

$$\text{(3.6)} \qquad z_k \ \geq \min_{\boldsymbol{\pi}_k \in \{\overline{\boldsymbol{\pi}}_{ks} : s = 1, \ldots, N_k\}} \boldsymbol{\pi}_k^\top(\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}).$$

Since Ω is compact, there exists a finite M so that we obtain model (3.3). ■

The dual feasible region $\mathcal{H}_k$ can be unbounded, but from Proposition 3.1, the extreme rays $\overline{\boldsymbol{\lambda}}_{ks}$ will not contribute to the objective value given Assumption 2. Therefore, we can focus our reformulation on the extreme points $\overline{\boldsymbol{\pi}}_{ks}$. Admittedly, the number of extreme points N_k can still be exponential in the problem parameters (n_x, n_ξ), leading to an exponential number of binary indicators $\boldsymbol{\mu}$. As a result, solving such a large-scale MILP model is still challenging in general. Therefore, we focus on two widely-used uncertainty sets in the literature: the central-limit-theorem (CLT)-induced uncertainty set in [5] and the budgeted uncertainty set in [12]. Next, we will show that RO-CDDU with either uncertainty set admitspolynomially-solvable reformulations when there are no integer variables.

CLT-induced uncertainty set by [5]. The uncertainty set proposed by [5] is mainly motivated by the central limit theorem. Based on their work, we consider a decision-dependent uncertainty set, $\Xi^{CLT}(\boldsymbol{x})$, in which the mean value and the standard deviation of $\boldsymbol{\xi}$ are affine functions of the decision variables $\boldsymbol{x}$:

$$\text{(3.7)} \quad \Xi^{CLT}(\boldsymbol{x}) \ :=\ \Big\{\boldsymbol{\xi} \in \mathbb{R}^{n_\xi} : \quad \Big|\sum_{i=1}^{n_\xi} \xi_i - \sum_{i=1}^{n_\xi} (\alpha_i^0 + {\boldsymbol{\alpha}_i^1}^\top \boldsymbol{x})\Big| \leq \Gamma\sigma(\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x})\Big\},$$

where constants $\boldsymbol{\alpha}^0 \in \mathbb{R}^{n_\xi}$, $\boldsymbol{\beta}^1 \in \mathbb{R}^{n_x}$, $\boldsymbol{\alpha}^1 \in \mathbb{R}^{n_\xi} \times \mathbb{R}^{n_x}$ and $\beta^0, \Gamma, \sigma \in \mathbb{R}_+$. We use affine functions $\sum_{i=1}^{n_\xi}(\alpha_i^0 + {\boldsymbol{\alpha}_i^1}^\top \boldsymbol{x})$ and $\Gamma\sigma(\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x})$ of decision variables $\boldsymbol{x}$ to replace the random variable's mean and standard deviation. We proceed to characterize the conditions in which the RO-CDDU problem is well-defined in Proposition 3.2, establish a polynomially-sized reformulation, and obtain the optimal value.

PROPOSITION 3.2. *The adversary's problem* $\max_{\boldsymbol{\xi}_k \in \Xi^{CLT}(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi}_k$ *yields a finite value, and correspondingly, the RO-CDDU model is well-defined if and only if* $\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x} \geq 0$ *and* $c_{k1} = \cdots = c_{kn_\xi} = c_k$ *for some* $c_k \in \mathbb{R}$. *Under such conditions, for any feasible solution* $(\boldsymbol{x}, \boldsymbol{y}) \in \Omega$, *the optimal value of the k-th adversary's problem satisfies:*

$$\max_{\boldsymbol{\xi}_k \in \Xi^{CLT}(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi}_k \;=\; |c_k| \cdot \Gamma\sigma(\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x}) + c_k \cdot \sum_{i=1}^{n_\xi} (\alpha_i^0 + {\boldsymbol{\alpha}_i^1}^\top \boldsymbol{x}). \tag{3.8}$$

Proof of Proposition 3.2. Using a standard transformation by introducing two auxiliary variables $u^+, u^- \geq 0$, we first write down a linear program reformulation of the adversary's problem (P^k) and its dual problem (D^k):

$$\begin{aligned}
(P^k) \qquad & \max_{\boldsymbol{\xi} \in \mathbb{R}^{n_\xi}, u^+ \geq 0, u^- \geq 0} && \sum_{i=1}^{n_\xi} c_{ki} \xi_i && \\
& \text{s.t.} && u^+ + u^- \;\leq\; \Gamma\sigma(\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x}), && : \pi_1 \\
& && u^+ - u^- \;=\; \sum_{i=1}^{n_\xi} \xi_i - \sum_{i=1}^{n_\xi} (\alpha_i^0 + {\boldsymbol{\alpha}_i^1}^\top \boldsymbol{x}). && : \pi_2 \\
(D^k) \qquad & \min_{\pi_1 \geq 0, \pi_2 \in \mathbb{R}} && \pi_1 \left[\Gamma\sigma(\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x}) \right] - \pi_2 \left[\sum_{i=1}^{n_\xi} (\alpha_i^0 + {\boldsymbol{\alpha}_i^1}^\top \boldsymbol{x}) \right] && \\
& \text{s.t.} && \pi_1 + \pi_2 \;\geq\; 0, && \\
& && \pi_1 - \pi_2 \;\geq\; 0, && \\
& && -\pi_2 \;=\; c_{ki}, \qquad \forall\, i = 1, \ldots, n_\xi. &&
\end{aligned}$$

The dual polyhedron $\mathcal{H}_k = \{(\pi_1, \pi_2) : \pi_1 \geq 0, \pi_1 \geq \pi_2, \pi_1 \geq -\pi_2, \pi_2 = -c_{ki}, \forall\, i = 1, \ldots, n_\xi\}$ is nonempty only if $c_{k1} = \cdots = c_{kn_\xi}$, which matches Assumption 1 to make sure that the adversary's problem is bounded. Since $\pi_1 \geq |\pi_2| \geq 0$, if $\Gamma\sigma(\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x}) < 0$, π_1 can take infinity to make (D^k) unbounded and (P^k) infeasible. Therefore, $c_{k1} = \cdots = c_{kn_\xi}$ and $\Gamma\sigma(\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x}) \geq 0$ are the conditions for the adversary's problem (P^k), and also the RO-CDDU model, to be well-defined.

On the other hand, if $c_{k1} = \cdots = c_{kn_\xi} = c_k$ and $\Gamma\sigma(\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x}) \geq 0$, we can always find an optimal solution to (D^k) as $\pi_1 = |c_k|$ and $\pi_2 = -c_k$. It is straightforward to see that π_2 has to be fixed at $-c_k$ by the equality constraint. The coefficient for π_1 is nonnegative and thus π_1 should take the minimum value, which is the larger of c_k and $-c_k$, i.e., $|c_k|$. By LP strong duality, the existence of such an optimal solution also suggests that the adversary's problem can achieve an optimal value at $c_k \cdot \sum_{i=1}^{n_\xi} (\alpha_i^0 + {\boldsymbol{\alpha}_i^1}^\top \boldsymbol{x}) + |c_k| \cdot \Gamma\sigma(\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x})$, and RO-CDDU is well-defined. ∎

Based on the characterization above, we only need to consider the unique extreme point of (D^k), $\pi_1 = |c_k|, \pi_2 = c_k$, to develop the following MILP reformulation for the RO-CDDU with a CLT-induced uncertainty set:

$$\min_{V, \boldsymbol{z}, \boldsymbol{x}, \boldsymbol{\mu}} \quad V \tag{3.9a}$$

$$\text{s.t.} \quad V \;\geq\; z_k + \boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + d_k, \qquad \forall k = 1, \ldots, K, \tag{3.9b}$$

$$z_k \;\geq\; c_k \cdot \sum_{i=1}^{n_\xi} (\alpha_i^0 + {\boldsymbol{\alpha}_i^1}^\top \boldsymbol{x}) + |c_k| \left[\Gamma\sigma(\beta^0 + {\boldsymbol{\beta}^1}^\top \boldsymbol{x}) \right], \qquad \forall k = 1, \ldots, K, \tag{3.9c}$$

$$(\boldsymbol{x}, \boldsymbol{y}) \in \Omega, \; \boldsymbol{x} \in \mathbb{R}^n, \; \boldsymbol{y} \in \mathbb{Z}^{n_y}. \tag{3.9d}$$

We note that model (3.9) admits polynomially-sized constraints, which also reduces the computational concerns in the branch-and-bound algorithm in the MILP problem. Without any integer variables $\boldsymbol{y}$, the RO-CDDU problem becomes polynomially solvable under the CLT-induced uncertainty set.

Budgeted uncertainty set by [12]. We consider the widely-studied budgeted uncertainty set in the literature [3; 18; 42], which was first proposed by [12]. Given positive integers T and $n_\xi = n_x = nT$, we define a budgeted uncertainty set $\Xi^B(\boldsymbol{x}) = \times_{t=1}^T \Xi_t^B(\boldsymbol{x}_t)$, in which $\Xi_t^B(\boldsymbol{x}_t)$ is defined as:

$$\Xi_t^B(\boldsymbol{x}_t) := \Bigg\{\boldsymbol{\xi}_t \in \mathbb{R}^n : \quad -\alpha_{it}^0 - \alpha_{it}^1 x_{it} \le \xi_{it} \le \beta_{it}^0 + \beta_{it}^1 x_{it}, \quad \forall i = 1, \ldots, n,$$
$$\sum_{i=1}^n \big|\xi_{it}\big| \le \zeta_t + \big|\boldsymbol{\omega}_t^\top \boldsymbol{x}_t\big|\Bigg\}, \tag{3.10}$$

where $\boldsymbol{\alpha}_t^0, \boldsymbol{\alpha}_t^1, \boldsymbol{\beta}_t^0, \boldsymbol{\beta}_t^1 \in \mathbb{R}_+^n$, $\zeta_t \in \mathbb{R}_+$, $\boldsymbol{\omega}_t \in \mathbb{R}^n$. Parameter $t = 1, \ldots, T$ indexes each piece in the Cartesian product of the uncertainty set. This formulation with $n_\xi = n_x$ is motivated by the multi-period model for the demand response application detailed in Section 4 but can be easily extended to the case where $n_\xi \ne n_x$. Notice that the Cartesian product admits a decomposable structure naturally, which is established in the following lemma.

LEMMA 3.3. *Given the uncertainty set $\Xi^B(\boldsymbol{x}) = \times_{t=1}^T \Xi_t^B(\boldsymbol{x}_t)$ with $\Xi_t^B(\boldsymbol{x}_t)$ defined in (3.10), we have that $\max_{\boldsymbol{\xi}\in\Xi^B(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi} = \sum_{t=1}^T \max_{\boldsymbol{\xi}_t\in\Xi_t^B(\boldsymbol{x}_t)} \boldsymbol{c}_{kt}^\top \boldsymbol{\xi}_t$.*

Proof of Lemma 3.3. We prove the lemma from two sides:

On one side, we first observe that $\max_{\boldsymbol{\xi}\in\Xi^B(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi} = \max_{\boldsymbol{\xi}\in\Xi^B(\boldsymbol{x})} \sum_{t=1}^T \boldsymbol{c}_{kt}^\top \boldsymbol{\xi}_t \le \sum_{t=1}^T \max_{\boldsymbol{\xi}_t\in\Xi_t^B(\boldsymbol{x}_t)} \boldsymbol{c}_{kt}^\top \boldsymbol{\xi}_t$, because an optimal solution $\boldsymbol{\xi}_t$ is chosen for each optimization problem $\max_{\boldsymbol{\xi}_t\in\Xi_t^B(\boldsymbol{x}_t)} \boldsymbol{c}_{kt}^\top \boldsymbol{\xi}_t$ with $t = 1, \ldots, T$.

On the other side, let $\boldsymbol{\xi}_t^*$ be the optimal solution to problem $\max_{\boldsymbol{\xi}_t\in\Xi_t^B(\boldsymbol{x}_t)} \boldsymbol{c}_{kt}^\top \boldsymbol{\xi}_t$ for all $t = 1, \ldots, T$. By definition of $\Xi(\boldsymbol{x})$ in (3.10), solution $\boldsymbol{\xi}^* = (\boldsymbol{\xi}_t^*)_{t=1}^T$ is feasible to problem $\max_{\boldsymbol{\xi}\in\Xi^B(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi}$. So, we obtain that $\max_{\boldsymbol{\xi}\in\Xi^B(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi} \ge \sum_{t=1}^T \max_{\boldsymbol{\xi}_t\in\Xi_t^B(\boldsymbol{x}_t)} \boldsymbol{c}_{kt}^\top \boldsymbol{\xi}_t$.

Combining the two observations above, we conclude that $\max_{\boldsymbol{\xi}\in\Xi^B(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi} = \sum_{t=1}^T \max_{\boldsymbol{\xi}_t\in\Xi_t^B(\boldsymbol{x}_t)} \boldsymbol{c}_{kt}^\top \boldsymbol{\xi}_t$. ∎

Lemma 3.3 suggests that the uncertainty set $\times_{t=1}^T \Xi_t^B(\boldsymbol{x}_t)$ allows us to decompose problem $\max_{\boldsymbol{\xi}\in\Xi^B(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi}$ into T independent adversary's problems. Using the standard linearization technique by letting $\xi_{it} = \xi_{it}^+ - \xi_{it}^-$ where $\xi_{it}^+, \xi_{it}^- \ge 0$, we can write the adversary's problem for a given $t = 1, \ldots, T$ as an equivalent linear program in $\boldsymbol{\xi}_t$:

$$(P^{kt}) \quad \max_{\boldsymbol{\xi}_t^+, \boldsymbol{\xi}_t^- \ge 0} \quad \boldsymbol{c}_{kt}^\top(\boldsymbol{\xi}_t^+ - \boldsymbol{\xi}_t^-) \tag{3.11a}$$
$$\text{s.t.} \quad \xi_{it}^+ - \xi_{it}^- \le \beta_{it}^0 + \beta_{it}^1 x_{it}, \qquad \forall i = 1, \ldots, n, \qquad : \pi_{1i} \tag{3.11b}$$
$$\xi_{it}^+ - \xi_{it}^- \ge -\alpha_{it}^0 - \alpha_{it}^1 x_{it}, \qquad \forall i = 1, \ldots, n, \qquad : \pi_{2i} \tag{3.11c}$$
$$\sum_{i=1}^n \left(\xi_{it}^+ + \xi_{it}^-\right) \le \zeta_t + \big|\boldsymbol{\omega}_t^\top \boldsymbol{x}_t\big|. \qquad : \pi_3 \tag{3.11d}$$

We derive the dual problem of model (3.11) as:

$$
\begin{aligned}
&\text{(3.12a)} \quad (D^{kt}) \min_{\substack{\boldsymbol{\pi}_1 \geq 0, \boldsymbol{\pi}_2 \leq 0 \\ \pi_3 \geq 0}} \sum_{i=1}^{n} \left[\pi_{1i}\left(\beta_{it}^0 + \beta_{it}^1 x_{it}\right) - \pi_{2i}\left(\alpha_{it}^0 + \alpha_{it}^1 x_{it}\right)\right] + \pi_3(\zeta_t + |\boldsymbol{\omega}_t^\top \boldsymbol{x}_t|) \\
&\text{(3.12b)} \quad \text{s.t.} \quad \pi_{1i} + \pi_{2i} + \pi_3 \geq c_{kit}, \qquad \forall i = 1, \ldots, n, \\
&\text{(3.12c)} \quad -\pi_{1i} - \pi_{2i} + \pi_3 \geq -c_{kit} \qquad \forall i = 1, \ldots, n.
\end{aligned}
$$

We show that the feasible region $\mathcal{H}_{kt}$ corresponding to model (3.12) has an exponential number of extreme points in n via the following lemma. Without loss of generality, we can assume that the cost coefficients $\boldsymbol{c}_{kt}$ are different, nonzero, and aligned in ascending order of their absolute values, i.e., $0 < |c_{k1t}| < \cdots < |c_{knt}|$. In addition, we also set $c_{k0t} := 0$ for simplicity of notations.

LEMMA 3.4. *We let the tuple $(\boldsymbol{\pi}_1^*, \boldsymbol{\pi}_2^*, \pi_3^*)$ denote an extreme point for polyhedron $\mathcal{H}_{kt}$. We can obtain a subset of extreme points satisfying the following conditions*

(i) for some $j^* = 1, \ldots, n$, $\pi_3^* = |c_{kj^*t}|$;
(ii) for any $i = 1, \ldots, n$,
- *if $i = j^*$, $\pi_{1i}^* = \pi_{2i}^* = 0$;*
- *if $i > j^*$, $(\pi_{1i}^*, \pi_{2i}^*) = \begin{cases} (c_{kit} - \pi_3^*, 0), & \text{if } c_{kit} > 0, \\ (0, c_{kit} + \pi_3^*), & \text{if } c_{kit} < 0; \end{cases}$*
- *if $i \leq j^*$, $(\pi_{1i}^*, \pi_{2i}^*) \in \{(0,0), (0, c_{kit} - \pi_3^*), (c_{kit} + \pi_3^*, 0)\}$.*

Proof of Lemma 3.4. To verify that the proposed point $(\boldsymbol{\pi}_1^*, \boldsymbol{\pi}_2^*, \pi_3^*)$ is an extreme point, we enumerate the following possibilities of linear independence conditions for the dual linear program for each $i = 1, \ldots, n$:

- for $i = j^*$, by setting up $\pi_{1i}^* = \pi_{2i}^* = 0$, three inequalities $\pi_{1i} \geq 0$, $\pi_{2i} \leq 0$ and $\pi_{1i} + \pi_{2i} + \pi_3 \geq c_{kit}$ (if $c_{kit} > 0$) or $-\pi_{1i} - \pi_{2i} + \pi_3 \geq -c_{kit}$ (if $c_{kit} < 0$) hold as equality;
- for $i > j^*$, if $c_{kit} > 0$, by setting up $\pi_{i1}^* = c_{kit} - \pi_3^*$ and $\pi_2^* = 0$, two inequalities $\pi_{2i} \leq 0$ and $\pi_{1i} + \pi_{2i} + \pi_3 \geq c_{kit}$ hold as equality; if $c_{kit} < 0$, by setting up $\pi_{i2}^* = c_{kit} + \pi_3^*$ and $\pi_1^* = 0$, two inequalities $\pi_{1i} \geq 0$ and $-\pi_{1i} - \pi_{2i} + \pi_3 \geq -c_{kit}$ hold as equality;
- for $i < j^*$:
  - at $\pi_{1i}^* = \pi_{2i}^* = 0$, two inequalities $\pi_{1i}^* \geq 0$ and $\pi_{2i}^* \leq 0$ hold as equality;
  - at $\pi_{1i}^* = 0, \pi_{2i}^* = c_{kit} - \pi_3^*$, $\pi_{1i}^* \geq 0$ and $\pi_{1i} + \pi_{2i} + \pi_3 \geq c_{kit}$ hold as equality;
  - at $\pi_{2i}^* = 0, \pi_{1i}^* = c_{kit} + \pi_3^*$, $\pi_{2i}^* \leq 0$ and $-\pi_{1i} - \pi_{2i} + \pi_3 \geq -c_{kit}$ hold as equality.

The solution $(\boldsymbol{\pi}_1^*, \boldsymbol{\pi}_2^*, \pi_3^*)$ is feasible by construction, at which there are $2n+1$ linearly independent inequality constraints holding as equality. Therefore, the solution $(\boldsymbol{\pi}_1^*, \boldsymbol{\pi}_2^*, \pi_3^*)$ is a basic feasible solution, and thus an extreme point for $\mathcal{H}_{kt}$. For each $j^* = 1, \ldots, n$, we can yield at least 3^{j^*-1} extreme points, three for each $i < j^* - 1$. This makes the total number of constructed extreme points, which is only a subset of all extreme points for $\mathcal{H}_{kt}$, at least $\sum_{j^*=1}^{n} 3^{j^*-1} = \frac{3^n - 1}{2}$. Therefore, the number of extreme points for $\mathcal{H}_{kt}$ satisfies $\Omega(3^n)$ and is exponential in parameter n. ■

We now establish Theorem 3.5 to show that we only need to consider a polynomial subset of extreme points to obtain the optimal solution to model (3.12).

THEOREM 3.5. *For any $k = 1, \dots, K$, $t = 1, \dots, T$, the optimal solution to (D^{kt}) is within a subset of the extreme points:*

$$
\Big\{(\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \pi_3): \quad \pi_3 = |c_{kjt}|,
$$

$$
\pi_{1i} = (|c_{kit}| - \pi_3)^+ \cdot \mathbb{1}_{c_{kit}>0},
$$

$$
\pi_{2i} = -(|c_{kit}| - \pi_3)^+ \cdot \mathbb{1}_{c_{kit}<0}, \quad \forall\, i = 1, \dots, n\Big\}_{j=0}^{n}, \tag{3.13}
$$

where we let $\mathbb{1}_A$ be the indicator function of statement A and $(\psi)^+ := \max\{\psi, 0\}$.

Proof of Theorem 3.5. We show that π_3 can only take $n+1$ possible values $\{|c_{kjt}|\}_{j=0}^n$ by discussing the following two cases will not happen in an optimal solution to (D^{kt}):

(i) $\pi_3 > \max_{j=1,\dots,n}\{|c_{kjt}|\}$: suppose $\pi_3 > \max_{j=1,\dots,n}\{|c_{kjt}|\}$ in the optimal solution. Decision variables π_{1i} and π_{2i} should take value 0 to achieve the minimum value for the objective function of D^{kt}. We can construct $\pi_3 = \max_{j=1,\dots,n}\{|c_{kjt}|\}$, which will yield a strictly better objective value. This contradiction suggests that in the optimal solution, $\pi_3 \le \max_{j=1,\dots,n}\{|c_{kjt}|\}$.

(ii) $|c_{k(j-1)t}| < \pi_3 < |c_{kjt}|$: suppose $\pi_3 \in (|c_{k(j-1)t}|, |c_{kjt}|)$ in the optimal solution. To reach the minimum, we need to have $\pi_1 = \pi_2 = 0$ for $i = 1, \dots, j-1$ and $\pi_1 = |c_{kit}| - \pi_3, \pi_2 = 0$ if $c_{kit} > 0$ or $\pi_1 = 0, \pi_2 = -|c_{kit}| + \pi_3$ for $i = j, \dots, n$. Therefore, we can express the objective value as:

$$
\sum_{i=j}^{n} \left[(|c_{kit}| - \pi_3)\left((\beta^0_{it} + \beta^1_{it}x_{it}) \cdot \mathbb{1}_{c_{kit}>0} + (\alpha^0_{it} + \alpha^1_{it}x_{it}) \cdot \mathbb{1}_{c_{kit}<0}\right)\right]
$$

$$
+ \pi_3(\zeta_t + |\boldsymbol{\omega}_t^\top \boldsymbol{x}_t|)
$$

The objective value is an affine function of π_3. We let $\phi = (\zeta_t + |\boldsymbol{\omega}_t^\top \boldsymbol{x}_t|) - \left((\beta^0_{it} + \beta^1_{it}x_{it}) \cdot \mathbb{1}_{c_{kit}>0} + (\alpha^0_{it} + \alpha^1_{it}x_{it}) \cdot \mathbb{1}_{c_{kit}<0}\right)$ denote the linear coefficient of π_3. If $\phi < 0$, $\tilde{\pi}_3 = |c_{kjt}|$ yields a strictly better objective than the optimal π_3, while if $\phi > 0$, $\tilde{\pi}_3 = |c_{k(j-1)t}|$ yields a strictly better objective. Both cases contradict the assumption that $\pi_3$is part of the optimal solution. When $\phi = 0$, the objective value remains the same with either $\tilde{\pi}_3 = |c_{k(j-1)t}|$ or $\tilde{\pi}_3 = |c_{kjt}|$ and we can equivalently consider $\tilde{\pi}_3$.

By excluding the two cases above, we are left with a finite set of values for π_3, $\{|c_{kjt}|\}_{j=0}^n$. For a candidate solution with $\pi_3 = |c_{kjt}|$ given j, we can realign constraints (3.12b) and (3.12c) as $c_{kit} - |c_{kjt}| \le \pi_{1i} + \pi_{2i} \le c_{kit} + |c_{kjt}|$. We can enumerate the following cases to show that either $\pi_{1i} = 0$ or $\pi_{2i} = 0$:

(i) $c_{kit} < 0, i < j$: here $c_{kit} + |c_{kjt}| > 0$ and $c_{kit} - |c_{kjt}| < 0$, since $\pi_{1i}\left(\beta^0_{it} + \beta^1_{it}x_{it}\right)$ and $-\pi_{2i}\left(\alpha^0_{it} + \alpha^1_{it}x_{it}\right)$ are both nonnegative, we have $\pi_{1i} = \pi_{2i} = 0$ at optimality;

(ii) $c_{kit} < 0, i > j$: here $c_{kit} + |c_{kjt}| < 0$ and $c_{kit} - |c_{kjt}| < 0$, to minimize the objective value, we have $\pi_{1i} = 0, \pi_{2i} = c_{kit} + |c_{kjt}|$ at optimality;

(iii) $c_{kit} > 0, i < j$: here $c_{kit} + |c_{kjt}| > 0$ and $c_{kit} - |c_{kjt}| < 0$, since $\pi_{1i}\left(\beta^0_{it} + \beta^1_{it}x_{it}\right)$ and $-\pi_{2i}\left(\alpha^0_{it} + \alpha^1_{it}x_{it}\right)$ are both nonnegative, we have $\pi_{1i} = \pi_{2i} = 0$ at optimality;

(iv) $c_{kit} > 0, i > j$: here $c_{kit} + |c_{kjt}| > 0$ and $c_{kit} - |c_{kjt}| > 0$, to minimize the objective value, we have $\pi_{1i} = c_{kit} - |c_{kjt}|, \pi_{2i} = 0$ at optimality.

Summarizing the four cases above, we can write the closed-form solution as $\pi_3 = |c_{kjt}|$ and for each $i = 1, \dots, n$, $\pi_{1i} = (|c_{kit}| - |c_{kjt}|)^+ \cdot \mathbb{1}_{c_{kit}>0}$, and $\pi_{2i} = -(|c_{kit}| - |c_{kjt}|)^+ \cdot \mathbb{1}_{c_{kit}<0}$, given a specific $j = 0, \dots, n$. Therefore, we conclude that the dual optimal solution can only come from the finite set stated in Theorem 3.5. ■

Theorem 3.5 establishes that for each piece t, the optimal value of adversary's problem $\max_{\boldsymbol{\xi}_t \in \Xi_t^B(\boldsymbol{x}_t)} \boldsymbol{c}_{kt}^\top \boldsymbol{\xi}_t$ subject to a budgeted uncertainty set $\Xi_t^B(\boldsymbol{x}_t)$ can be expressed as the minimum of $n+1$ linear functions, instead of an exponential number based on Lemma 3.4. By Proposition 3.1, we can establish the following MILP model:

$$\min_{V,\boldsymbol{z},\boldsymbol{x},\boldsymbol{\mu}} \quad V \tag{3.14a}$$

$$\text{s.t.} \quad V \geq \sum_{t=1}^{T} z_{kt} + \boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + d_k, \quad \forall k = 1, \dots, K, \tag{3.14b}$$

$$\begin{aligned} z_{kt} \geq \sum_{i=1}^{n} \Big[& (|c_{kit}| - |c_{kjt}|)^+ \cdot \mathbb{1}_{c_{kit}>0} \left(\beta_{it}^0 + \beta_{it}^1 x_{it}\right). \\ & + (|c_{kit}| - |c_{kjt}|)^+ \cdot \mathbb{1}_{c_{kit}<0} \left(\alpha_{it}^0 + \alpha_{it}^1 x_{it}\right) \Big] \\ & + |c_{kjt}|(\zeta_t + e_\ell \boldsymbol{\omega}_t^\top \boldsymbol{x}_t) - M(1 - \mu_{kjt}), \\ & \forall j = 0, \dots, n, k = 1, \dots, K, t = 1, \dots, T, \ell = 1, 2 \end{aligned} \tag{3.14c}$$

$$\sum_{j=0}^{n} \mu_{kjt} = 1, \qquad \forall k = 1, \dots, K, \ t = 1, \dots, T, \tag{3.14d}$$

$$\mu_{jkt} \in \{0, 1\}, \quad \forall j = 0, \dots, n, k = 1, \dots, K, t = 1, \dots, T, \tag{3.14e}$$

$$(\boldsymbol{x}, \boldsymbol{y}) \in \Omega, \ \boldsymbol{x} \in \mathbb{R}^{n_x}, \ \boldsymbol{y} \in \mathbb{Z}^{n_y}. \tag{3.14f}$$

We use the parameters $e_1 = 1$ and $e_2 = -1$, to linearize the absolute value of $\boldsymbol{\omega}_t^\top \boldsymbol{x}_t$. Constraint (3.14c) is equivalent to constraint (3.3c) with $\overline{\boldsymbol{\pi}}_{ks}$ substituted by the candidate dual solutions in (3.13). Recall that a major computational challenge for the MILP problem in (3.3) is that the number of binary variable $\boldsymbol{\mu}$ in (3.3c) may be exponential in (n_x, n_ξ). Under the budgeted uncertainty set in (3.10), Theorem 3.5 shows that it is without loss of optimality to consider a subset of binary variable $\boldsymbol{\mu}$ with polynomial size given a fixed number of function pieces K in RO-CDDU model (1.1) and a fixed number T of budgeted uncertainty sets, $\Xi^B(\boldsymbol{x}) = \times_{t=1}^T \Xi_t^B(\boldsymbol{x}_t)$, which can reduce the computational burden in the branch-and-bound algorithm when solving for the MILP problem.

Furthermore, we note the polynomial solvability for problem (3.14) with fixed parameters (T, K) and without any integer variables such that $n_y = 0$.

Corollary 3.6. *Under the budgeted uncertainty set* (3.10)*, when there are no integer variables* $\boldsymbol{y}$ *(i.e.,* $n_y = 0$*), for fixed* $K, T \in \mathbb{N}_+$*, model* (3.14) *has a polynomial run-time in parameters* n*.*

Proof of Corollary 3.6. By Theorem 3.5, we only need $n+1$ steps to enumerate all candidate solutions. Therefore, it takes $(n+1)^{KT}$ steps to enumerate all feasible dual solution candidates $\boldsymbol{\mu}$. Given that $n_y = 0$, for a specific feasible candidate $\boldsymbol{\mu}$, model (3.14) is reduced to a linear program, which can be solved by the interior point method in $\mathcal{O}(n_x^{3.5})$ steps [24]. With $n_x = nT$, as a result, model (3.14) could be solved in $\mathcal{O}((nT)^{3.5}(n+1)^{KT})$ steps. ■

Together with Proposition 3.1, Corollary 3.6 provides a sufficient condition for

RO-CDDU with a budgeted uncertainty set to be solved polynomially in input size n: after eliminating the integer variable $\boldsymbol{y}$, adversary's problem $\max_{\boldsymbol{\xi}\in\Xi(\boldsymbol{x})} \boldsymbol{c}_k^\top \boldsymbol{\xi}$ needs to admit a corresponding dual feasible region with effectively a polynomial number of extreme points to consider.

3.2 Algorithms to solve RO-CDDU We consider two algorithms in this section to solve the RO-CDDU reformulation (3.1): an alternating direction algorithm (ADA) and a column generation algorithm (CGA). For the demand response problem in subsequent Section 4, we numerically demonstrate that both ADA and CGA achieve tight optimality gap with a shorter run-time compared to solving the MILP (3.3) directly with the commercial solver.

3.2.1 Alternating Direction Algorithm In ADA, we iteratively search for the feasible solutions to model (3.1) in the subspace of $\boldsymbol{\pi}$ and $\boldsymbol{x}$ in constraint (3.1c). This is equivalent to keeping one vector $\boldsymbol{\pi}$ for constraints (3.3c) in each iteration. We present ADA in Algorithm 3.1.

Algorithm 3.1 Alternating Direction Algorithm (ADA)

1: **Initialization**: $s = 0$ and $(\boldsymbol{x}^0, \boldsymbol{y}^0) \in \Omega$, $\boldsymbol{y}^0 \in \mathbb{Z}^{n_y}$
2: **repeat**
3: **for** $k = 1, \dots, K$ **do**
4: Solve model (3.15) and obtain an optimal solution $\boldsymbol{\pi}_k^{s+1}$:

$$\min \quad \boldsymbol{\pi}^\top (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}^s) \quad \text{s.t.} \quad \boldsymbol{\pi} \in \mathcal{H}_k. \tag{3.15}$$

5: **end for**
6: Let $\overline{z}_k^{s+1} = {\boldsymbol{\pi}_k^{s+1}}^\top (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}^s)$ and $\overline{V}^{s+1} = \max_{k=1,\dots,K} \overline{z}_k^{s+1} + \boldsymbol{a}_k^\top \boldsymbol{x}^s + \boldsymbol{b}_k^\top \boldsymbol{y}^s + d_k$.
7: Solve

$$\begin{aligned}
&\min_{V,\boldsymbol{z},\boldsymbol{x},\boldsymbol{y}} && V && \text{(3.16a)}\\
&\text{s.t.} && V \ \geq \ z_k + \boldsymbol{a}_k^\top \boldsymbol{x} + \boldsymbol{b}_k^\top \boldsymbol{y} + d_k, && \forall k, \quad \text{(3.16b)}\\
& && z_k \ \geq \ (\boldsymbol{\pi}_k^{s+1})^\top (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}), && \forall k, \quad \text{(3.16c)}\\
& && (\boldsymbol{x}, \boldsymbol{y}) \in \Omega, && \text{(3.16d)}\\
& && \boldsymbol{x} \in \mathbb{R}^{n_x}, \ \boldsymbol{y} \in \mathbb{Z}^{n_y}. && \text{(3.16e)}
\end{aligned}$$

8: Obtain an optimal solution $(V^{s+1}, \boldsymbol{z}^{s+1}, \boldsymbol{x}^{s+1}, \boldsymbol{y}^{s+1})$ of (3.16).
9: $s \leftarrow s + 1$
10: **until** convergence criterion is met.

Note that model (3.15) is an LP and (3.16) is an MILP. We can show that the sequence of value functions $\{(\overline{V}^s, V^s) : s = 1, 2, \dots\}$ is convergent due to the monotonicity of the optimal values.

THEOREM 3.7. *Suppose the model (3.1) has a finite global optimal value V^*. The sequence of the objective function values, $\{(\overline{V}^s, V^s) : s = 1, 2, \dots\}$, generated by Algorithm 3.1, is monotonically nonincreasing, i.e. $\overline{V}^{s+1} \geq V^{s+1} \geq \overline{V}^{s+2} \geq V^*$ for all $s \geq 0$. Hence, $\{\overline{V}^s, V^s\}$ converges to a finite value, which is an upper bound on V^*.*

Proof of Theorem 3.7. From the minimization problems (3.15) and (3.16) in iteration s, we obtain a feasible solution vector $(V^s, \boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{z}^s, \boldsymbol{\pi}^s)$ to model (3.1), which implies that the global optimal value V^* of model (3.1) serves as a lower bound of the sequence $\{\overline{V}^1, V^1, \overline{V}^2, V^2, \dots\}$.

Moreover, noting that $(V^s, \boldsymbol{x}^s, \boldsymbol{y}^s, \boldsymbol{z}^s)$ is a feasible solution to problem (3.16) from iteration s given $\boldsymbol{\pi}^s$, in iteration $s+1$, we can establish that $\overline{V}^{s+1} \le V^s$ from minimization problem (3.15). Next, in minimization problem (3.16), solution vector $(\overline{V}^{s+1}, \boldsymbol{x}^s, \boldsymbol{y}^s, \overline{\boldsymbol{z}}^{s+1})$ is a feasible solution given the updated $\boldsymbol{\pi}^{s+1}$ from problem (3.15). Thus, we have $V^{s+1} \le \overline{V}^{s+1}$ and the sequence of $\{\overline{V}^1, V^1, \overline{V}^2, V^2, \dots\}$ is a nonincreasing sequence bounded from below by V^*, and thus is convergent. ■

The algorithm searches for solution $\boldsymbol{\pi}_k$ in a subset of the extreme points for $\mathcal{H}_k$ and the convergent process obtains a feasible solution to model (3.1) in every iteration. The sequence $\{V^s\}$ is possible to converge to a suboptimal value, but we show in Section 4 that ADA can achieve good feasible solutions quickly. We show in the subsequent discussion that ADA could be further improved with the budgeted uncertainty set.

Improved ADA with the budgeted uncertainty set. We consider the budgeted uncertainty set in (3.10). When we solve adversary's problem (3.15), we leverage the special structures in Theorem 3.5, with which solving model (3.15) only requires verifying $n+1$ solution candidates.

COROLLARY 3.8. *Given the budgeted uncertainty set defined in* (3.10)*, the following* $\boldsymbol{\pi}$ *solves model* (3.15)*: for any* $k = 1, \dots, K$, $t = 1, \dots, T$,

$$\begin{aligned}
\boldsymbol{\pi}^*_{kt} \in \operatorname*{arg\,min}_{j=0,\dots,n} \Bigg\{ & \sum_{i=1}^{n} \Big[(|c_{kit}| - \pi_3)^+ \Big((\beta^0_{it} + \beta^1_{it} x_{it}) \cdot \mathbb{1}_{c_{kit}>0} + (\alpha^0_{it} + \alpha^1_{it} x_{it}) \cdot \mathbb{1}_{c_{kit}<0} \Big) \Big] \\
& + \pi_3 (\zeta_t + |\boldsymbol{\omega}_t^\top \boldsymbol{x}_t|) : \\
& \pi_{1i} = (|c_{kit}| - \pi_3)^+ \cdot \mathbb{1}_{c_{kit}>0}, \quad \forall i = 1, \dots, n, \\
& \pi_{2i} = (|c_{kit}| - \pi_3)^+ \cdot \mathbb{1}_{c_{kit}<0}, \quad \forall i = 1, \dots, n, \\
& \pi_3 = |c_{kjt}| \Bigg\}.
\end{aligned} \tag{3.17}$$

Proof of Corollary 3.8. The proof follows directly from the proof of Theorem 3.5 that it is without loss of optimality to only consider the subset of the extreme points in (3.17) of the dual polyhedron for problem (D^{kt}) from (3.13). ■

Based on Corollary 3.8, we simplify the optimization of model (3.15) to a search process. In (3.17), solution π_3 takes one of the values from $\{|c_{k1t}|, \dots, |c_{knt}|, 0\}$ and $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ can be subsequently decided. Since there are only $n+1$ solution candidates for π_3 for each $k = 1, \dots, K$ and $t = 1, \dots, T$, we only need to make nKT comparisons to find the optimal solution $\boldsymbol{\pi}$.

3.2.2 Column Generation Algorithm We propose a Column Generation Algorithm (CGA) for problem (3.1). CGA has been proposed to solve robust optimization problems in the literature [4; 52]. CGA starts from a master problem with an incomplete set of variables and calls certain oracles to compute the next variable to append to the master problem. In many cases, the number of critical variables added to the master problem is small, which makes the algorithm computationally tractable. We adopt the idea of CGA to solve model (3.1) and present CGA in Algorithm 3.2.

In Algorithm 3.2, in contrast to ADA that only preserves the most recent solution $\hat{\boldsymbol{\pi}}_k$ in each iteration, CGA appends $\hat{\boldsymbol{\pi}}_k$ to a solution set Π_k, preserves more elements

Algorithm 3.2 Column Generation Algorithm

Initialization: an initial set of extreme points Π_k of $\mathcal{H}_k$ of cardinality $N_k = |\Pi_k|$
repeat
  Solve model (3.3) with $\boldsymbol{\pi}_{ks} \in \Pi_k, s = 1, \dots, N_k$, obtain the feasible solution $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}$ and the objective value $\hat{V}$;
  for $k = 1, \dots, K$ **do**
    Solve model (3.18) and obtain an optimal extreme point $\hat{\boldsymbol{\pi}}$:

$$\min \quad \boldsymbol{\pi}^\top(\boldsymbol{h} - \boldsymbol{T}\hat{\boldsymbol{x}}) \quad \text{s.t.} \quad \boldsymbol{\pi} \in \mathcal{H}_k. \tag{3.18}$$

    if $\hat{\boldsymbol{\pi}} \notin \Pi_k$ **then**
      Append $\hat{\boldsymbol{\pi}}$ to Π_k, $N_k \leftarrow N_k + 1$, $U_k =$ **false**
    else
      $U_k =$ **true**
    end if
  end for
until $\cap_{k=1}^K U_k =$ **true**.

in the solution set Π_k, and creates more opportunities to find better solutions than ADA does. Similar to Algorithm 3.1, given that Π_k is a subset of the extreme points in polyhedron $\mathcal{H}_k$, CGA terminates with a subset of variables μ_{ks}, which yields the finite convergence result as follows:

Corollary 3.9. *Algorithm 3.2 terminates after finite steps with a feasible solution to model* (3.3).

Proof of Corollary 3.9. The proof follows from the same monotonicity arguments as in Theorem 3.7. In each iteration, for $k = 1, \dots, K$, an extreme point of $\mathcal{H}_k$ is added to Π_k. As the number of elements in Π_k increases monotonically, the objective value of model (3.3) decreases monotonically. Given that the number of extreme points for $\mathcal{H}_k$ is finite, this leads to convergence of CGA. ■

We again leverage the monotonicity property to prove this convergence result. When Algorithm 3.2 terminates, solution $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$ may be suboptimal for problem (3.1). Despite this, the numerical performance for the demand response problem in Section 4 shows that CGA consistently reaches the global optimum. Given the budgeted uncertainty set in (3.10), we can also simplify solving model (3.18) based on Corollary 3.8, which further improves the speed of Algorithm 3.2.

3.3 Lower Bound from McCormick Relaxation To approximate the problem (3.1) from below, we consider the McCormick relaxation proposed in [32] to approximate the bilinear terms. We point to [20; 25; 33] for reference of theory and applications on McCormick approximation. Without loss of generality, we assume that Ω and $\mathcal{H}_k$ are compact, and thus the decision variables $(\boldsymbol{x}, \boldsymbol{\pi})$ are bounded. Suppose the technology matrix $\boldsymbol{T}$ has l rows. We define the lower bound of $(\boldsymbol{x}, \boldsymbol{\pi})$ by $(\underline{\boldsymbol{x}}, \underline{\boldsymbol{\pi}})$ where $\underline{\boldsymbol{x}} = (\underline{x}_1, \dots, \underline{x}_{n_x})$ and $\underline{\boldsymbol{\pi}}_k = (\underline{\pi}_{k1}, \dots, \underline{\pi}_{kl})$. Similarly, we define the upper bound of $(\boldsymbol{x}, \boldsymbol{\pi})$ by $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\pi}})$ where $\overline{\boldsymbol{x}}_i = (\overline{x}_1, \dots, \overline{x}_{n_x})$ and $\overline{\boldsymbol{\pi}}_k = (\overline{\pi}_{k1}, \dots, \overline{\pi}_{kl})$. Constraint (3.1c) can be approximated by the McCormick relaxation below: for any $k = 1, \dots, K$,

$j = 1, \ldots, l$, $i = 1, \ldots, n_x$, we have

$$z_k \;\geq\; \boldsymbol{\pi}_k^\top \boldsymbol{h} - \sum_{i=1}^{n_x}\sum_{j=1}^{l} T_{ji} q_{kji}, \tag{3.19a}$$

$$q_{kji} \;\geq\; \pi_{kj}\underline{x}_i + x_i\underline{\pi}_{kj} - \underline{\pi}_{kj}\underline{x}_i, \tag{3.19b}$$

$$q_{kji} \;\geq\; \pi_{kj}\overline{x}_i + x_i\overline{\pi}_{kj} - \overline{\pi}_{kj}\overline{x}_i, \tag{3.19c}$$

$$q_{kji} \;\leq\; \pi_{kj}\underline{x}_i + x_i\overline{\pi}_{kj} - \overline{\pi}_{kj}\underline{x}_i, \tag{3.19d}$$

$$q_{kji} \;\leq\; \pi_{kj}\overline{x}_i + x_i\underline{\pi}_{kj} - \underline{\pi}_{kj}\overline{x}_i. \tag{3.19e}$$

We can further refine constraints (3.19b) - (3.19e) if we partition the intervals of $\boldsymbol{x}$ and $\boldsymbol{\pi}$ into more pieces. It will result in a disjunctive MILP formulation, in which only one subset of $(\boldsymbol{x}, \boldsymbol{\pi})$ is selected. The McCormick relaxation will be tightened when the number of partitions increases. In this approach, we could approximate the RO-CDDU problem in (1.1) with arbitrary precision. However, in this MILP problem, the size of disjunctive constraints grows in the order of M^2, where M represents the number of partitions. For example, for large-scale instances in Section 4, it would be intractable to solve this MILP problem with multiple partitions. Therefore, we use the formulation in (3.19) with $M = 1$ to generate a lower bound for model (3.1).

4 Application in Demand Response Portfolio Management

4.1 Modeling Background In electricity markets, consumers who can reduce or shift their electricity usage during certain periods are considered DR resources. DR resources have gained more attention in recent years to help power system operators balance supply and demand, lower generation costs, and improve system efficiency [1; 26]. A DR portfolio can have thousands of DR resources of various characteristics, such as the ability to respond to load reduction under the variance of the demands [21]. Proper scheduling is necessary and challenging [34; 41]. For DR scheduling optimization, Reference [47] proposed a deterministic optimization model to solve the automatic load management problem in a smart home. Reference [23] developed a forward market clearing algorithm for the demand flexibility problem with the goal of co-optimizing the scheduling cost and the system security. Reference [40] characterized a novel control approach based on online optimization to manage the operations of responsive electrical appliances. The impact of uncertainty has also been extensively studied. For example, various robust optimization models with exogenous price uncertainty are proposed in [15; 16].

There are three main players in a DR event: the system operator, the DR aggregator, and the DR resources. The DR aggregator gains revenue from the system operator for providing the required demand reduction. At the same time, it offers payment to the participating DR resources in its portfolio [21]. Each DR resource has a set of operational characteristics to be respected during a DR event. Figure 1 illustrates these key characteristics on a scheduled dispatch trajectory of a DR resource. The key constraints include three parts as follows:

(1) Reduction constraints: DR resource i has a capacity $x_i^{\max}$ and minimum commitment requirement $x_i^{\min}$. Since we consider active demand reduction, we assume $x_i^{\min} \geq 0$.
(2) Ramping constraints: DR resource i has ramping limits r_i^+ and r_i^-.

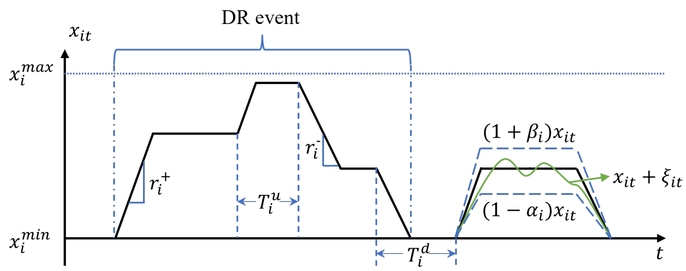


Fig. 1: The dynamics of a DR resource and realization uncertainty.

(3) Smoothness constraints: every time the demand reduction level of DR resource i increases (decreases), it cannot decrease (increase) before at least T_i^u (T_i^d) periods due to DR resource's inertia.

4.2 The Deterministic Model We take the perspective of the DR aggregator, who earns revenue c_i from committing DR resource i for a unit demand reduction. There is a required level of demand reduction for the DR aggregator based on contracts with the system operator. A mismatch of DR amount leads to a penalty at any time t: (i) if the total reduction level is less than the required level, the unit under-commitment cost for the DR aggregator induced by refund, contractual penalty, and loss of market, is s_t; (ii) if the total load reduction level exceeds the required level, the unit over-commitment cost caused by value loss of DR resources, is h_t. The DR aggregator aims to maximize its profit by committing the right portfolio of DR resources.

We let D_t be the required total demand reduction level at time t, which is a deterministic parameter known to the DR aggregator. Let $\boldsymbol{x} = (x_{it})$, in which x_{it} is the demand reduction level for resource i at the beginning of time t. For the DR aggregator, the total cost includes the over-commitment cost, the under-commitment cost, and the commitment revenue, which can be expressed as:

$$\sum_{t=1}^{T}\left[h_t\left(\sum_{i=1}^{n} x_{it} - D_t\right)^+ + s_t\left(D_t - \sum_{i=1}^{n} x_{it}\right)^+ - \sum_{i=1}^{n} c_i x_{it}\right], \tag{4.1}$$

where $(x)^+ := \max(x, 0)$. The objective function is a piecewise-linear convex function. Complicated operational constraints, such as startup, shutdown, and ramping limits, can cause a mismatch in DR scheduling. We let the binary variables u_{it} indicate whether resource i is committed at time t. We also set two binary ramping indicators w_{it} and v_{it} such that $w_{it} = 1$ if $x_{it} - x_{i(t-1)} \geq 0$, and $v_{it} = 1$ if $x_{it} - x_{i(t-1)} \leq 0$. We propose the following novel deterministic model for DR portfolio management:

$$
\begin{aligned}
&(4.2\text{a}) \quad && \min_{\boldsymbol{x},\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}} \quad f(\boldsymbol{x}) && \\
&(4.2\text{b}) && \text{s.t.} \quad x_i^{\min} u_{it} \leq x_{it} \leq x_i^{\max} u_{it}, && \forall i = 1,\dots,n,\ t = 1,\dots,T,\\
&(4.2\text{c}) && \quad -r_i^- u_{it} \leq x_{i(t+1)} - x_{it} \leq r_i^+ u_{i(t+1)}, && \forall i = 1,\dots,n,\ t = 1,\dots,T-1,\\
&(4.2\text{d}) && \quad x_{it} - x_{i(t-1)} \leq M w_{it}, && \forall i = 1,\dots,n,\ t = 2,\dots,T,\\
&(4.2\text{e}) && \quad x_{i\tau} - x_{i(\tau-1)} \geq -M(1 - w_{it}), && \forall i = 1,\dots,n,\ t = 1,\dots,T,\\
& && && \tau = t,\dots,\min(t + T_i^u - 1, T),\\
&(4.2\text{f}) && \quad x_{it} - x_{i(t-1)} \geq -M v_{it}, && \forall i = 1,\dots,n,\ t = 2,\dots,T,\\
&(4.2\text{g}) && \quad x_{i\tau} - x_{i(\tau-1)} \leq M(1 - v_{it}), && \forall i = 1,\dots,n,\ t = 1,\dots,T,
\end{aligned}
$$

$$\tau = t, \dots, \min(t + T_i^d - 1, T),$$

$$\text{(4.2h)} \qquad u_{it}, w_{it}, v_{it} \in \{0, 1\}, \qquad \forall i = 1, \dots, n,\ t = 1, \dots, T.$$

In constraint (4.2b), when a DR resource is committed ($u_{it} = 1$), the reduction amount x_{it} has to be bounded from above and below. Constraint (4.2c) defines the maximum and minimum ramping rates for committed resource i at time t, because the DR aggregator needs to respect the smoothness characteristics in scheduling demand reduction. In constraints (4.2d) and (4.2e), if resource i increases its commitment at any time t, it has to keep the non-decreasing trend for a minimum of T_i^u periods. Similarly, constraints (4.2f) and (4.2g) require that if resource i decreases its commitment at any time t, it has to keep the non-increasing trend for at least T_i^d periods. The big-M parameter M stands for a large positive real number. The proposed model (4.2) is a novel formulation for DR portfolio management that explicitly models the detailed commitment cycle dynamics of DR resources. It also considers a piecewise linear cost function which can balance the over- and under-commitment costs for DR aggregators.

4.3 Robust Demand Response Model In a DR event, the aggregator schedules the reduction level for each resource. However, unlike conventional generators, the demand reduction of DR resources can have significant uncertainty due to unexpected factors in operations and market conditions. The final realized reduction level of a DR resource may be different from the one scheduled.

We model the final realization of demand reduction as $\tilde{x}_{it} = x_{it} + \xi_{it}$, where ξ_{it} represents the implementation noise bounded in the uncertainty set below:

$$\text{(4.3)} \quad \Xi_t(\boldsymbol{x}_t) = \left\{ \boldsymbol{\xi}_t = (\Delta x_{1t}, \dots, \Delta x_{nt}) \ : \ \begin{array}{c} -\alpha_i x_{it} \le \xi_{it} \le \beta_i x_{it}, \ \ \forall i = 1, \dots, n \\ \sum_{i=1}^n |\xi_{it}| \le \Gamma_t \sum_{i=1}^n x_i^{\max}, \ \ \forall t = 1, \dots, T \end{array} \right\},$$

where $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}_+^n$. The proposed uncertainty set captures the positive correlation between the implementation noise and the scheduled commitment, which is pointed out in the demand response literature [46; 53]. Since a resource with a large capacity can sometimes commit a small demand reduction, such an uncertainty model (4.3) can avoid the over-conservativeness caused by decision-independent uncertainty in which the uncertainty range is only proportional to the resource's capacity. Moreover, we use Γ_t to capture the DR aggregator's conservativeness level and the risk preference in uncertainty. It is worth noting that the uncertainty set formulation (4.3) is a special case of the budgeted uncertainty set (3.10) with $\boldsymbol{\alpha}_t^0 = \boldsymbol{\beta}_t^0 = \boldsymbol{\omega}_t = \mathbf{0}$.

Similar to the objective function in (4.1), we define the objective function of the robust DR problem as the following piecewise-linear function:

$$\text{(4.4)} \ f(\boldsymbol{x}, \boldsymbol{\xi}) = \sum_{t=1}^{T} \left\{ h_t \left(\sum_{i=1}^{n} (x_{it} + \xi_{it}) - D_t \right)^+ + s_t \left(D_t - \sum_{i=1}^{n} (x_{it} + \xi_{it}) \right)^+ - \sum_{i=1}^{n} c_i (x_{it} + \xi_{it}) \right\}.$$

Given (4.3) and (4.4), we formulate the robust DR portfolio management problem with the framework of RO-CDDU in (1.1). Notice that the condition $x_i^{\min} \ge 0$ for any $i = 1, \dots, n$ will guarantee that Assumption 2 holds for any feasible $\boldsymbol{x}$ because the dual minimization problem of $\max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{x})} \ f(\boldsymbol{x}, \boldsymbol{\xi})$ is lower bounded by 0.

$$\text{(4.5a)} \qquad \min_{\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}} \ \max_{\boldsymbol{\xi} \in \Xi(\boldsymbol{x})} \quad f(\boldsymbol{x}, \boldsymbol{\xi})$$

$$\text{(4.5b)} \qquad \text{s.t.} \quad (\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) \text{ satisfies (4.2b)-(4.2h)}.$$

The rest of this section covers the computational experiments solving model (4.5). In Section 4.4, we detail the setups for our numerical experiments. In Section 4.5, we demonstrate the performance of ADA and CGA. We numerically benchmark the objective value obtained by ADA and CGA against the lower bound obtained from (i) the McCormick relaxation of formulation (3.1), (ii) the best objective value of (3.1), and (iii) the best objective value of (3.3), with both (ii) and (iii) solved within a fixed time span. In Section 4.6, we investigate how the robust solutions obtained with different uncertainty budgets Γ_t perform under a stochastic setting.

4.4 Experiment Setup We construct two test cases for the numerical experiments, one with simulated DR resources' parameters and the other with real-world data. For both cases, we let the time horizon length $T = 9$ and use a time-invariant parameter Γ for the uncertainty budget such that $\Gamma_t = \Gamma$ for all $t = 1, \ldots, T$.

Table 1: Parameter setups for the simulated DR resources, with $\mathcal{U}\{a, b\}$ representing binary distribution between bounds a and b and $U[a, b]$ representing continuous uniform distribution between bounds a and b

Parameters	Setting		
	Type A	Type B	Type C
c_i ($U_c \sim U[0, 2]$)	$22 + 2U_c$	$20 + 2U_c$	$18 + 2U_c$
$\alpha_i = \beta_i$	0.5	0.3	0.1
$x_i^{\max}$	$15 + \overline{U}_x$, $\overline{U}_x \sim \mathcal{U}\{0, 5\}$		
$x_i^{\min}$	$4 + \underline{U}_x$, $\underline{U}_x \sim \mathcal{U}\{0, 1\}$		
r_i^+	$5 + \overline{U}_r$, $\overline{U}_r \sim \mathcal{U}\{0, 2\}$		
r_i^-	$5 + \underline{U}_r$, $\underline{U}_r \sim \mathcal{U}\{0, 2\}$		
T_i^u	$2 + \overline{U}_e$, $\overline{U}_e \sim \mathcal{U}\{0, 2\}$		
T_i^d	$2 + \underline{U}_e$, $\underline{U}_e \sim \mathcal{U}\{0, 2\}$		

Parameters		Setting
D_t	Low	$\begin{cases} 6n & \text{if } t = 4, 5, 6 \\ 0 & \text{otherwise} \end{cases}$
	High	$\begin{cases} 15n & \text{if } t = 4, 5, 6 \\ 0 & \text{otherwise} \end{cases}$
s_t	Low	$1000 + U_s$, $U_s \sim U[0, 500]$
	High	$1000 + U_s$, $U_s \sim U[0, 500]$
h_t		$h_t = 30 + U_h$, $U_h \sim U[0, 5]$

The detailed parameter setups of the simulated test case are shown in Table 1. We assume variations of resources' power reduction commitments are positively correlated to profitability, which is common in risk-return analysis [14]. We simulate three types of resources: A, B, and C. In the order from A to C, resources have increasing unit revenue c_i, but also bear an increasing operational uncertainty, measured by $\beta_i - \alpha_i$. We set resources' uncertainty bounds homogeneous within each type. The ramping rates and capacity limits are randomly generated from uniform distributions.

Based on the current industry practice, we let $h_t > c_i$ because too much supply impairs DRs' economic value and causes power system instability [21; 43]. We set a substantially higher under-commitment cost $s_t >> c_i$, because a shortage of power supply can lead to severe contractual penalties from system operators who suffer from power outage, credibility damage, and potential loss of market share to competitors. We set two levels of under-commitment costs s_t (high and low) and demands D_t (high and low) to approximate different market conditions and load profiles.

We further group 20 DR resources as a cluster, since they may be correlated in realistic power systems [50]. Within a cluster, we assume that the binary ramping decisions are the same for every resource in all time periods: all resources in a cluster need to increase/decrease their response output together. This reduces the number of binary variables and helps solve the problem computationally.

For the second test case with real data, we obtain the DR resources' parameters from the electricity demand data of 115 buildings on the University of Southern California (USC) campus, which are modeled as DR resources in [2]. The USC data has rather heterogeneous resource capacities compared to the simulated data, where the largest generator has a generation capacity 1000 times larger than the smallest one. We list the detailed parameter setup based on the USC data in a GitHub repository.[1] Since the dataset contains independent buildings, we do not cluster the resources to align our test case with reality.

The optimization models specific to the DR problem follow the constructions in Section 3 and are implemented using JuMP package v0.22.1 [17] in Julia v1.6.2, with bilinear, linear, and mixed-integer programs solved by Gurobi 9.5.0 [19]. All tests are run on a server with 30 Intel Xeon cores at 2.6 GHz and 128 GB of RAM.

4.5 Computational Performance Analysis We discuss the computational performance of the proposed methods to solve the robust DR model in (4.4), which includes directly solving the MINLP model (3.1) with bilinear terms, solving the exact MILP formulation, ADA, CGA, and the McCormick relaxation for a lower bound.

We record the computational performance in Table 2, with a setting of low demand, low under-commitment penalty cost and $\Gamma = 0.05$. We record the negative objective values (row "-Obj"), the run-time (row "Time"), and the optimality gap information. For the bilinear formulation and the exact MILP formulation, we set a run-time limit as 18,000 seconds, and we record the gap information output by Gurobi when the solution process terminates either by reaching optimality or the time limit (row "Gap"). The objective values in ADA/CGA/MILP correspond to some feasible solutions and provide upper bounds for model (4.5). The upper bounds and the lower bound are obtained by solving the McCormick relaxation model to form an optimality gap (row "MC Gap").

Table 2: Different algorithms' time performance to solve model (4.5)

Test Case		Simulated Data						USC
		$n = 5$	$n = 50$	$n = 200$	$n = 400$	$n = 800$	$n = 1200$	$n = 115$
Bilinear	-Obj ($)	1221	12649	52198	95347	202024	279261	386912
	Time (sec.)	6.8	> 18000	> 18000	> 18000	> 18000	> 18000	> 18000
	Gap (%)	0.00	0.70	0.71	0.53	6.69	12.25	0.74
	MC Gap (%)	0.00	0.85	0.79	0.69	1.16	0.74	1.60
MILP	-Obj ($)	1221	12649	52198	95347	202695	279189	387765
	Time (sec.)	4.2	631.1	> 18000	> 18000	> 18000	> 18000	384.1
	Gap (%)	0.00	0.01	1.42	3.09	4.67	14.73	0.01
	MC Gap (%)	0.00	0.85	0.79	0.69	0.84	0.77	1.38
ADA	-Obj ($)	1221	12649	52198	95347	202695	279266	387763
	Time (sec.)	0.6	2.3	18.7	44.0	263.6	165.9	3.7
	MC Gap (%)	0.00	0.85	0.79	0.69	0.84	0.74	1.38
CGA	-Obj ($)	1221	12649	52198	95347	202695	279266	387765
	Time (sec.)	0.6	26.7	255.6	302.4	2187.0	5883.8	855.0
	MC Gap (%)	0.00	0.85	0.79	0.69	0.84	0.74	1.38
MC	-Obj ($)	1221	12758	52615	96011	204405	281346	393203
	Time (sec.)	0.2	3.8	101.1	407.6	7824.4	> 18000	2.1

From Table 2, we observe that the MC gap is less than 2% for all test cases, which indicates that the upper bounds and the lower bounds are close to the true optimum. In addition, for all cases in Table 2, ADA and CGA achieve close solutions, with some minimal differences caused by numerical precision when terminating the optimization process. Since ADA and CGA solve different sequences of the mixed-integer programs

[1] `https://github.com/haoxiangyang89/RO-CDDU`

with the relative termination gap set at 10^{-4}, they can terminate at different solutions within the termination gap. Table 2 shows that ADA is significantly faster than CGA when they achieve the same solution. On the other hand, the exact bilinear/MILP formulation takes an extended period of time to solve for large test cases. For example, in test cases with $n \geq 200$, no experiment terminates at the default tolerance level within five hours. When the bilinear/MILP solution process terminates due to the time limit, it can achieve the same solution quality as the ones obtained by ADA/CGA in smaller problem instances ($n \leq 800$), but it does not yield solutions as good as from ADA and CGA in the largest case ($n = 1200$). We observe from the Gurobi log file that the lower bound increases slowly and many nodes are generated towards the end of the branch-and-bound process. This effect is more apparent when n is large because each resource only contributes to a small share of demand. Many resources with similar unit profits can be considered substitutes, which leads to different solutions with similar objective values. The similarity of DR resources makes it difficult to prune nodes in the branch-and-bound tree. This is also reflected in the USC test case: the MILP solves much faster even with a larger n and no clustering, because the resources are heterogeneous in both capacity and unit profit. Note that this issue can be alleviated by further clustering the resources: instead of only clustering resources' ramping decisions, a DR operator can ask the clustered resources to output the same percentage of their capacity. This is equivalent to reducing the number of resources, which is shown to be computationally effective in Table 2.

A natural question from the results in Table 2 is that since ADA obtains the same solution at a faster speed compared to CGA, should we always prefer ADA to CGA? In Table 3, we show that in many simulated test cases with $n = 20$ and all resources in one cluster, ADA can end at a suboptimal point with an optimality gap as large as 19%, but CGA consistently reaches optimality for the same cases. Although CGA takes a longer time to converge, its optimality is validated in all test cases we run as it obtains the same optimal values as the MILP model in (3.3). The gap between CGA and ADA is more prominent in the high demand cases. Since more resources need to be utilized in those instances, solution structures can be more complicated with more non-zero commitment, which makes ADA more likely to land in a suboptimal solution.

Table 3: Impact of uncertainty set budget parameter Γ on computational performance ($n = 20$). Notation V denotes the objective value of model (4.5). Superscripts A and C stand for "ADA" and "CGA". We omit the optimal values of MILP since they are identical to CGA's. All tests (ADA/CGA/MILP) take less than 10 seconds to finish.

Cost-Demand Setting	Obj. ($)	Γ									
		0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
Low-Low	$-V^A$	5746	5438	5131	4826	4531	4241	3958	3676	3439	3206
	$-V^C$	5746	5438	5131	4826	4531	4241	3958	3676	3439	3206
Low-High	$-V^A$	11572	10972	10386	9809	9238	8668	8106	8658	8117	7608
	$-V^C$	12562	11995	11436	10883	10335	9777	9217	8658	8117	7608
High-Low	$-V^A$	5231	4883	4539	4201	3875	3551	3231	2914	2665	2415
	$-V^C$	5231	4883	4539	4201	3875	3551	3231	2914	2665	2415
High-High	$-V^A$	9534	8872	8221	7587	6965	6342	5714	6449	5880	5318
	$-V^C$	11799	10091	9487	8888	8287	7680	7061	6449	5880	5318

From Table 3, we observe that the negative optimal value decreases almost linearly as the conservativeness level Γ increases. This result implies that the optimal dual variable for the budget constraint is approximately a constant. The slope of the linear relationship characterizes the value of the budget limit. Such a value increases when the under-commitment penalty becomes higher. We also notice that the increasing demand increases the profit with a diminishing marginal benefit. While in the high demand case, in which the total demand is 2.33 times larger than in the low demand case, the optimal profit ratio is less than 2.33. This is mainly caused by the following two factors: (i) as the demand increases, we need to schedule less-profitable resources, which brings down the marginal profit; (ii) more resource commitment comes with a larger magnitude of uncertainty, which negatively affects the marginal profit.

4.6 Solution Analysis In this section, we examine the property of the solution to model (4.5), generated by CGA, to understand how the robust solution improves the demand response performance under uncertainty.

Favoring DR resources with less uncertainty. Figure 2 shows the percentage of commitment from each type of resources in the test cases with $n = 800$ and $\Gamma \in [0, 0.1]$. We observe that type-A resources are generally favored in the deterministic solution due to their high unit profit. However, as the uncertainty budget Γ increases, the utilization rate of less uncertain resources increases. The robust optimization model returns more conservative solutions by committing more type-B and type-C resources. This demonstrates the ability of the robust DR model to balance between the nominal profit and the operational uncertainty. In the low demand setting, since the demand is only 30% of the total capacity, there is relatively more freedom to choose from different types of DR resources, which leads to a more diverse portfolio of DR resources. On the other hand, the higher demand setting requires more participation of all types of resources, which brings the commitment percentages closer.

Increasing total reduction. Since the under-commitment cost is significantly higher than the over-commitment cost, strategically committing resources above the required reduction level substantially reduces the likelihood of the under-commitment penalty in actual operations. We observe that the robust DR model is able to do so to avoid the negative impact of the worst-case scenarios. As shown in Figure 3, the total scheduled DR level of the robust solutions is higher during the peak time, while the deterministic solution satisfies the demand exactly.

Table 4: Reduction comparison between resources with different uncertainty levels using real-world data

Resource ID	α_i	Total reduction for $\Gamma = 0.01$ vs. $\Gamma = 0.05$
13	0.235	-42.9
16	0.225	-21.2
20	0.045	667.1

We also observe the same properties with the USC dataset, as illustrated by the following example. The total demand response level is 2859.0 for time period $t \in \{4, 5, 6\}$ when $\Gamma = 0.01$, but it increases to 3095.3 for $\Gamma = 0.05$. Resources 13 and 16 have the largest and the second largest α_i and β_i, which means they have the largest operational uncertainty. Their commitment decreases as Γ increases, while the total demand response level increases. This gap is filled by deploying more stable resources such as resource 20. The numerical results are displayed in Table 4.

Next, we study the profit performance of the robust DR solutions obtained with uncertainty budget Γ in a stochastic setting. For such setting, we assume that the

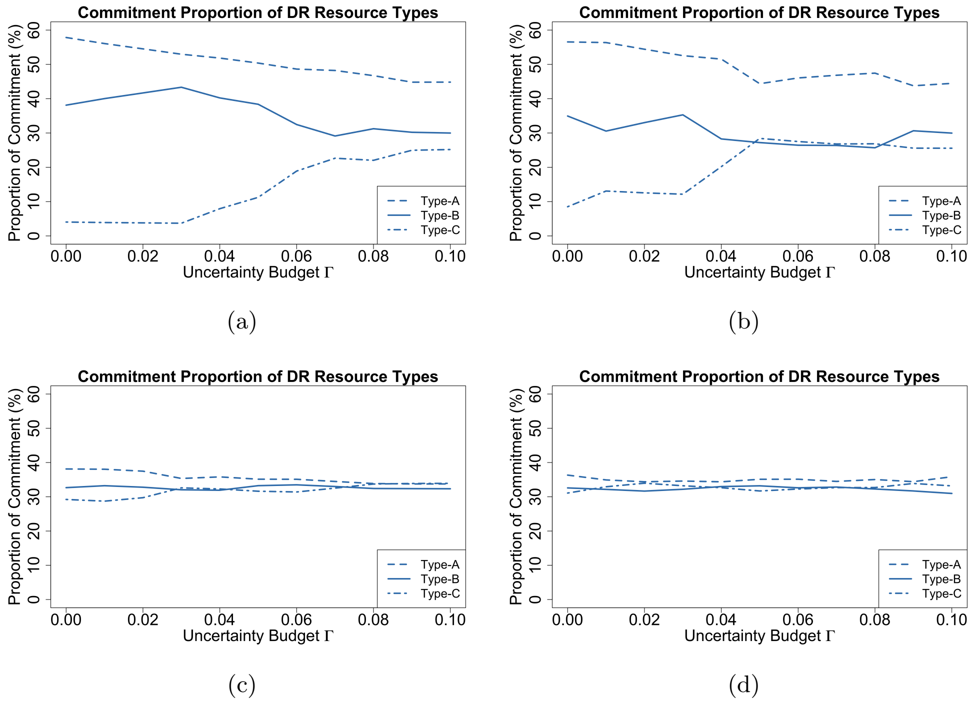
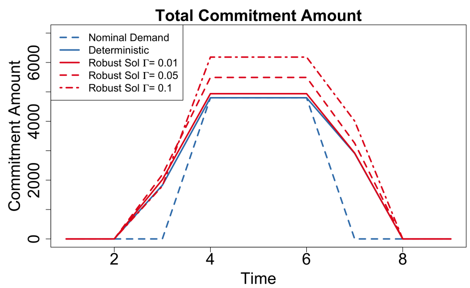
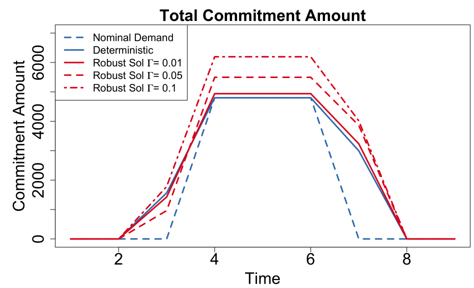


Fig. 2: Allocation proportion of three types of DR resources vs. different Γ under the setting of: (a) low demand level and low shortage cost; (b) low demand level and high shortage cost; (c) high demand level and low shortage cost; (d) high demand level and high shortage cost. The point $\Gamma = 0$ corresponds to the deterministic DR solution.

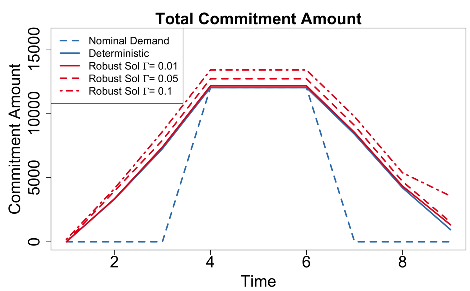
uncertain load $\xi_{it} = \rho_{it} x_{it}$, given a demand response solution $\boldsymbol{x}$, where the uncertainty coefficient ρ_{it} follows a uniform distribution within the interval $[-\alpha_i, \beta_i]$ for every $i = 1, \dots, n$ and $t = 1, \dots, T$. We generate 5,000 samples of $\boldsymbol{\rho}$ to create a load profile using Monte Carlo simulation, with which we evaluate the cost obtained for the given solution $\boldsymbol{x}$. The experiment serves the purpose of an out-of-sample test as the load scenario may lie outside of the uncertainty set proposed in (4.3). Figure 4 shows the mean out-of-sample costs of the robust solutions in four demand-cost settings. The x-axis captures different uncertainty budgets Γ. Figure 4 shows that because of the severe shortage penalty, the robust DR solutions display better results than the deterministic solution. The mean out-of-sample profit improves significantly even when we consider a small uncertainty budget $\Gamma = 0.01$. Combined with the results from Figure 2 and 3, Figure 4 shows that the solution with $\Gamma = 0.01$ does not increase the total commitment by much, but it slightly changes the proportion of DR resource types. This significantly improves the out-of-sample expected profit. The result further shows that achieving robustness may not necessarily always require a large reserve of resources. A smart commitment allocation can improve the overall robustness with a lean operation. As the uncertainty budget increases, the solution becomes more conservative, and thus the profit peaks at a certain level and then decreases. Only under the "high demand, high shortage cost" setting, such peak is at $\Gamma = 0.02$ and in every other case, the robust solution with $\Gamma = 0.01$ achieves the best out-of-sample performance.



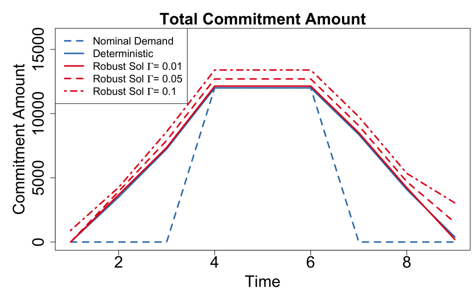
(a)



(b)



(c)



(d)

Fig. 3: Comparison of demand response amount between the deterministic solution and robust solutions with three different Γ under the setting of: (a) low demand level and low shortage cost; (b) low demand level and high shortage cost; (c) high demand level and low shortage cost; (d) high demand level and high shortage cost. Nominal demand is represented using a blue dashed Line.

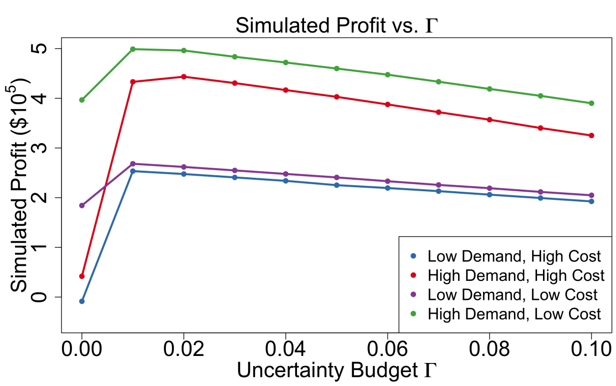


Fig. 4: Mean out-of-sample profit vs. Γ under the setting of: (i) low demand level and low shortage cost; (ii) low demand level and high shortage cost; (iii) high demand level and low shortage cost; (iv) high demand level and high shortage cost. The point $\Gamma = 0$ corresponds to the deterministic DR solution.

5 Conclusions In this paper, we propose the RO-CDDU model and show that it is strongly $\mathcal{NP}$-hard. The original RO-CDDU model can be formulated as an MINLP and we investigate the structure of the dual polyhedron for the adversary's problem such that RO-CDDU is well-defined. Meanwhile, we develop an equivalent

MILP reformulation using extreme points of the dual polyhedron and show two special uncertainty sets with polynomial solvability. We develop an alternating direction algorithm and a column generation algorithm to obtain feasible solutions and upper bounds for RO-CDDU. We compare the upper bounds with the lower bound obtained by solving a McCormick relaxation. Then, we propose a novel RO-CDDU model for portfolio management of demand response resources in electricity markets, where the realization of demand response is uncertain and depends on the demand response decision. The proposed ADA algorithm can obtain good solutions efficiently in most test cases. The proposed CGA algorithm further improves on the solution quality of ADA and obtains global optimal solutions in all test cases.

Acknowledgements. Hongfan (Kevin) Chen's research was supported by research grants from the Chinese University of Hong Kong Business School. Andy Sun's research was partially supported by the National Science Foundation under grant 1751747. Haoxiang Yang's research was supported by the National Natural Science Foundation of China, under the grant numbers 72231008 and 72201232, and the "AI for Low-Carbon Initiative" of Shenzhen Institute of Artificial Intelligence and Robotics for Society (AIRS). We thank the Information Technology Services Office (ITSO) at the Chinese University of Hong Kong, Shenzhen, for the generous allotment of computing time on its clusters. The authors thank Dr. David P. Morton and Dr. John R. Birge for helpful suggestions in modeling and theoretical analysis. We appreciate two anonymous referees and the associate editor, Dr. Wolfram Wiesemann, for comments and suggestions which improved the paper substantially.

References

[1] M. Albadi and E. El-Saadany, *A summary of demand response in electricity markets*, Electr. Power Syst. Res., 78 (2008), pp. 1989–1996.

[2] S. Aman, M. Frincu, C. Chelmis, U. M. Noor, Y. Simmhan, and V. Prasanna, *Empirical comparison of prediction methods for electricity consumption forecasting*. Technical Report 14-942, University of Southern California, 2014.

[3] A. Ardestani-Jaafari and E. Delage, *Robust optimization of sums of piecewise linear functions with application to inventory problems*, Oper. Res., 64 (2016), pp. 474—494.

[4] S. Babaei, R. Jiang, and C. Zhao, *Distributionally robust distribution network configuration under random contingency*, IEEE Trans. Power Syst., 35 (2020), pp. 3332–3341.

[5] C. Bandi and D. Bertsimas, *Tractable stochastic analysis in high dimensions via robust optimization*, Math. Prog., 134 (2012), p. 23–70.

[6] C. Bandi and D. Gupta, *Operating room staffing and scheduling*, Manuf. Serv. Oper. Manag., 22 (2020), p. 958–974.

[7] C. Bandi, E. Han, and O. Nohadani, *Sustainable inventory with robust periodic-affine policies and application to medical supply chains*, Manage. Sci., 65 (2019), p. 4636–4655.

[8] C. Bandi, N. Trichakis, and P. Vayanos, *Robust multiclass queuing theory for wait time estimation in resource allocation systems*, Manage. Sci., 65 (2019), p. 152–187.

[9] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*, Princeton University Press, 2009.

[10] D. Bertsimas and V. Goyal, *On the power and limitations of affine policies in two-stage adaptive optimization*, Math. Prog., 134 (2012), pp. 491–531.

[11] 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., 28 (2013), pp. 52–63.

[12] D. Bertsimas and M. Sim, *The price of robustness*, Oper. Res., 52 (2004), pp. 35–53.

[13] C. Buchheim and J. Kurtz, *Robust combinatorial optimization under convex and discrete cost uncertainty*, EURO J. Comput. Optim., 6 (2018), pp. 211–238.

[14] J. Y. Campbell, *Understanding risk and return*, J. Polit. Econ., 104 (1996), pp. 298–345.

[15] Z. Chen, L. Wu, and Y. Fu, *Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization*, IEEE Trans. Smart Grid, 3 (2012), pp. 1822–1831.

[16] A. Conejo, J. Morales, and L. Baringo, *Real-time demand response model*, IEEE Trans. Smart Grid, 1 (2010), pp. 236–242.

[17] I. Dunning, J. Huchette, and M. Lubin, *JuMP: A modeling language for mathematical optimization*, SIAM Review, 59 (2017), pp. 295–320.

[18] M. D. Filabadi and H. Mahmoudzadeh, *Effective budget of uncertainty for classes of robust optimization*, INFORMS J. Opt., (2022).

[19] Gurobi Optimization, Inc., *Gurobi Reference Manual*, 2016.

[20] M. Hasan and I. Karimi, *Piecewise linear relaxation of bilinear programs using bivariate partitioning*, AIChE J., 56 (2010), pp. 1880–1893.

[21] T. Healy, *EnerNoc 2013 annual report.* URL: http://investor.enernoc.com/annual-proxy.cfm, 2013.

[22] O. H. Ibarra and C. E. Kim, *Fast approximation algorithms for the knapsack and sum of subset problems*, J. ACM, 22 (1975), pp. 463—-468.

[23] E. Karangelos and F. Bouffard, *Towards full integration of demand-side resources in joint forward energy/reserve electricity markets*, IEEE Trans. Power Syst., 27 (2012), pp. 280–289.

[24] N. Karmakar, *A new polynomial-time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373–395.

[25] R. Karuppiah and I. Grossmann, *Global optimization for the synthesis of integrated water systems in chemical processes*, Comput. Chem. Eng., 30 (2006), pp. 650–673.

[26] D. Kirschen, *Demand-side view of electricity markets*, IEEE Trans. Power Syst., 18 (2003), pp. 520–527.

[27] Q. Kong, S. Li, N. Liu, C.-P. Teo, and Z. Yan, *Appointment scheduling under time-dependent patient no-show behavior*, Manage. Sci., 66 (2020), pp. 3480–3500.

[28] N. Lappas and C. Gounaris, *Robust optimization for decision-making under endogenous uncertainty*, Comput. Chem. Eng., 111 (2018), pp. 252–266.

[29] S. Leyffer, M. Menickelly, T. Munson, C. Vanaret, and S. M. Wild, *A survey of nonlinear robust optimization*, INFOR, 58 (2020), pp. 342–373.

[30] F. Luo and S. Mehrotra, *Distributionally robust optimization with decision dependent ambiguity sets*, Optim. Lett., 14 (2020), pp. 2565–2594.

[31] H. Mamani, S. Nassiri, and M. Wagner, *Closed-form solutions for robust inventory management*, Manage. Sci., 63 (2017), p. 1625–1643.

[32] G. McCormick, *Computability of global solutions to factorable nonconvex programs: Part I convex underestimating problems*, Math. Prog., 10 (1976), pp. 147–175.

[33] H. Nagarajan, M. Lu, S. Wang, R. Bent, and K. Sundar, *An adaptive, multivariate partitioning algorithm for global optimization of nonconvex programs*, J. Glob. Optim., 74 (2019), pp. 639–675.

[34] S. Nan, M. Zhou, and G. Li, *Optimal residential community demand response scheduling in smart grid*, Appl. Energy, 210 (2018), pp. 1280 – 1289.

[35] O. Nohadani and K. Sharma, *Optimization under decision-dependent uncertainty*, SIAM J. Optim., 28 (2018), pp. 1773–1795.

[36] N. Noyan, G. Rudolf, and M. Lejeune, *Distributionally robust optimization under a decision-dependent ambiguity set with applications to machine scheduling and humanitarian logistics*, INFORMS J. Comput., (2021), pp. 1–23.

[37] Ö. Okur, N. Voulis, P. Heijnen, and Z. Lukszo, *Aggregator-mediated demand response: Minimizing imbalances caused by uncertainty of solar generation*, Appl. Energy, 247 (2019), pp. 426–437.

[38] M. Poss, *Robust combinatorial optimization with variable budgeted uncertainty*, 4OR Q. J. Oper. Res., 11 (2013), pp. 75–92.

[39] M. Poss, *Robust combinatorial optimization with knapsack uncertainty*, Discret. Optim., 27 (2018), pp. 88–102.

[40] M. Rastegar and M. Fotuhi-Firuzabad, *Outage management in residential demand response programs*, IEEE Trans. Smart Grid, 6 (2014), pp. 1453–1462.

[41] H. Roh and J. Lee, *Residential demand response scheduling with multiclass appliances in the smart grid*, IEEE Trans. Smart Grid, 7 (2015), pp. 94–104.

[42] E. Roos and D. den Hertog, *Reducing conservatism in robust optimization*, INFORMS J. Comput., 32 (2020), pp. 855–1186.

[43] J. Shair, H. Li, J. Hu, and X. Xie, *Power system stability issues, classifications and research prospects in the context of high-penetration of renewables and power electronics*, Renew. Sust. Energ. Rev., 145 (2021), p. 111111.

[44] S. Spacey, W. Wiesemann, D. Kuhn, and W. Luk, *Robust software partitioning with multiple instantiation*, INFORMS J. Comput., 24 (2012), pp. 500–515.

[45] J. Sun and J. A. Van Mieghem, *Robust dual sourcing inventory management: Optimality of capped dual index policies and smoothing*, Manuf. Serv. Oper. Manag., 21 (2019), p. 912–931.

[46] J.-H. Teng and C.-H. Hsieh, *Modeling and investigation of demand response uncertainty on reliability assessment*, Energies, 14 (2021), p. 1104.

[47] K. Tsui and S. Chan, *Demand response optimization for smart home scheduling under real-time pricing*, IEEE Trans. Smart Grid, 3 (2012), pp. 1812–1821.

[48] R. Vujanic, P. Goulart, and M. Morari, *Robust optimization of schedules affected by uncertain events*, J. Optim. Theory Appl., 171 (2016), pp. 1033–1054.

[49] W. Whitt and W. You, *Using robust queueing to expose the impact of dependence in single-server queues*, Oper. Res., 66 (2018), pp. 184–199.

[50] H. Yang, D. P. Morton, C. Bandi, and K. Dvijotham, *Robust optimization for electricity generation*, INFORMS J. Comput., 33 (2021), pp. 336–351.

[51] X. Yu and S. Shen, *Multistage distributionally robust mixed-integer programming with decision-dependent moment-based ambiguity sets*, Math. Prog., (2020), pp. 1–40.

[52] B. Zeng and L. Zhao, *Solving two-stage robust optimization problems using a column-and-constraint generation method*, Oper. Res. Lett., 41 (2013), pp. 457–461.

[53] J. Zhang and A. D. Domínguez-García, *Evaluation of demand response resource aggregation system capacity under uncertainty*, IEEE Trans. Smart Grid, 9 (2017), pp. 4577–4586.

[54] S. Zhao and F. You, *Resilient supply chain design and operations with decision-dependent uncertainty using a data-driven robust optimization approach*, AIChE J., 65 (2019), pp. 1006–1021.