

Two-Stage Robust Optimization Under Decision Dependent Uncertainty

Yunfan Zhang, Feng Liu, *Senior Member, IEEE*, Yifan Su, Yue Chen, *Member, IEEE*,
Zhaojian Wang, *Member, IEEE*, and João P. S. Catalão, *Fellow, IEEE*

Abstract—In the conventional robust optimization (RO) context, the uncertainty is regarded as residing in a predetermined and fixed uncertainty set. In many applications, however, uncertainties are affected by decisions, making the current RO framework inapplicable. This paper investigates a class of two-stage RO problems that involve decision-dependent uncertainties. We introduce a class of polyhedral uncertainty sets whose right-hand-side vector has a dependency on the here-and-now decisions and seek to derive the exact optimal wait-and-see decisions for the second-stage problem. A novel iterative algorithm based on the Benders dual decomposition is proposed where advanced optimality cuts and feasibility cuts are designed to incorporate the uncertainty-decision coupling. The computational tractability, robust feasibility and optimality, and convergence performance of the proposed algorithm are guaranteed with theoretical proof. Four motivating application examples that feature the decision-dependent uncertainties are provided. Finally, the proposed solution methodology is verified by conducting case studies on the pre-disaster highway investment problem.

Index Terms—Benders decomposition, decision-dependent uncertainty, endogenous uncertainty, robust optimization (RO).

I. INTRODUCTION

A. Background

UNCERTAINTIES widely exist in real-world decision-making problems and various mathematical programming techniques, including scenario-based or chance-constrained

stochastic programs, robust optimization (RO) [1] and distributionally robust optimization [2], have been developed to fit in different characteristics of the applications [3]. RO seeks a risk-averse solution by explicitly considering the worst-case effect of all possible realizations of the uncertain parameter within a pre-determined uncertainty set. It appeals especially when the decision maker has no knowledge of the probability distributions of the uncertain parameters, or when the feasibility of the system over the entire uncertainty set is prioritized. Due to its salient advantages on modeling capability, feasibility, and computational tractability [4], RO has gained increasing popularity over the recent decades and encompassed a wide variety of applications including process scheduling [4]–[6], power system planning and scheduling [7]–[10], and network optimization [3], [11], [12], etc.

In the context of RO, the uncertainty sets are usually assumed to be a priori and fixed. The underlying assumption is that the decision maker's strategies would not alter the range of uncertainty realization. In many real-world decision-making problems, however, uncertainties can depend on the strategies chosen by the decision makers and are assumed as endogenous. For example, in a batch-process scheduling problem, the processing time or the production yield of a task is endogenously uncertain since it retains a physical meaning only when the optimizer decides to operate the associated task in a given period [4]–[6], [13]. Another example is the demand response program on buildings' electricity consumption [8]. For buildings participating in the program, the reserve demand requested from the system operator is uncertain and of endogenous nature due to its dependency on the reserve capacity provided by the building in the day-ahead market. The uncertainties affected by decisions are called decision-dependent uncertainties (DDUs) or endogenous uncertainties. In this paper, we use these two terms interchangeably without distinction. Also, we refer to decision-independent uncertainties (DIUs) or exogenous uncertainties as those not altered by decisions. Consideration of DDUs in RO problems can provide considerably less conservative solutions, attributed to the fact that DDUs can be proactively controlled by the optimizer. However, differently from DIUs, the presence of DDUs brings great challenges to solving RO problems, mainly due to the mutual influences between uncertainties and decisions. In this paper, we propose a convergence-guaranteed algorithm to efficiently solve the two-stage RO problems with DDUs.

Manuscript received July 6, 2021; revised October 22, 2021; accepted December 26, 2021. This work was supported by the Joint Research Fund in Smart Grid under cooperative agreement between the National Natural Science Foundation of China (NSFC) and State Grid Corporation of China (U19666601). Recommended by Associate Editor Zhen Song. (Corresponding author: Feng Liu.)

Citation: Y. F. Zhang, F. Liu, Y. F. Su, Y. Chen, Z. J. Wang, and J. P. S. Catalão, "Two-stage robust optimization under decision dependent uncertainty," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 7, pp. 1295–1306, Jul. 2022.

Y. F. Zhang, F. Liu, and Y. F. Su are with the State Key Laboratory of Power System and Generation Equipment, the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (e-mail: zhang-yf18@mails.tsinghua.edu.cn; lfeng@tsinghua.edu.cn; suyf19@mails.tsinghua.edu.cn).

Y. Chen is with the Department of Mechanical and Automation Engineering, the Chinese University of Hong Kong, Hong Kong SAR, China (e-mail: yuechen@mae.cuhk.edu.hk).

Z. J. Wang is with the Ministry of Education Key Laboratory of System Control and Information Processing, the Department of Automation, Shanghai Jiao Tong University, and also with Shanghai Engineering Research Center of Intelligent Control and Management, Shanghai 200240, China (e-mail: wangzhaojian@sjtu.edu.cn).

J. P. S. Catalão is with the Faculty of Engineering of the University of Porto and Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Porto 4200-465, Portugal (e-mail: catalao@fe.up.pt).

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

Digital Object Identifier 10.1109/JAS.2022.105512

B. Literature Review

Regarding RO under DDUs, existing works can be categorized into two classes, based on whether there are recourse or wait-and-see decisions after the revelation of uncertainty realization:

1) *Static RO Under DDUs*: In static RO, which is also called single-stage RO, all the decisions are here-and-now, i.e., to be determined before observing the uncertainty realization. In [14] and [15], variable budgeted static RO is studied, where the budget parameter that regulates the conservatism of the uncertainty set is modeled as an affine function with respect to the decision variables. In [16], polyhedral and conic uncertainty sets are introduced where decisions control the upper bounds of the uncertain variables. In [3] and [4], uncertainties are governed by binary-valued materialization indicator variables. The materialization indicator attains the values of 1 if the corresponding uncertain parameter retains a physical meaning in the problem and 0 otherwise. The DDU-involved RO models in the above works maintain the advantage of computational tractability of conventional RO problems, since a robust counterpart of mixed-integer linear programming (MILP) formulation can be derived by applying the strong duality theory and McCormick envelopes convex relaxation.

2) *Adaptive RO Under DDUs*: Adaptive robust optimization (ARO), or called adjustable RO [17], is established in a two-stage or a more general multi-stage setting where recourse or wait-and-see decisions are incorporated in response to actual uncertainty realization. The computational complexity of ARO lies in the fact that multi-level programming problems with more than two decision levels are non-deterministic polynomial-time (NP)-hard [18]. To overcome this intrinsic computational burden as well as the challenges raised by the coupling relationship between the here-and-now decisions and the uncertainties, affine decision rules parameterized in the uncertainty realizations and the here-and-now decisions are postulated on the wait-and-see decisions in [5], [19]–[21]. Differently from the aforementioned approximate affine policies, the decision rules in [22] are generated with the aid of multi-parametric programming technique, rendering the exact solution to the ARO problem; however, computation burden of solving the second-stage problem parametrically as an explicit function of the here-and-now decisions and the uncertainty is transferred off-line but not eliminated. In the line of primal cut algorithms [23], [24], [9] proposes an improved column-and-constraint generation (C&CG) algorithm based on a worst-case scenario mapping technique. Its application, however, is limited to a high-dimensional rectangle decision-dependent uncertainty set (DDUS).

Compared with conventional ARO problems, the challenges of solving DDUs-incorporated ARO problems reside in the fact that the existing cutting plane algorithms [23]–[26], with constraints generated by whether primal or dual information of the second-stage problem, fail to warrant finite convergence to the global optimum when the uncertainty set depends on the here-and-now decision variable which changes with iterations. This is because in that case the optimality cuts and feasibility cuts generated with concrete uncertain

parameter values may become invalid and ruin the optimality of the solution, or even render an infeasible master problem which falsely implies that the original ARO problem is not robust feasible. Moreover, since the vertices of the uncertainty set change with the here-and-now decision, it is difficult to justify the finite convergence of these cutting plane algorithms by completely enumerating the vertices of the polyhedral uncertainty set. To the best of the authors' knowledge, the solution approach for ARO under DDUs with a generic decision-dependency and without any assumption on approximation policies has not been addressed.

C. Contribution and Organization

The main contributions of this paper are twofold:

1) *From a Modeling Perspective*: This paper addresses a generic two-stage robust optimization model with a class of polyhedral DDUSs whose right-hand-side vector has a dependency on the here-and-now decisions. Differently from [3], [4], [6], [14]–[16] that contemplate the static RO involving endogenous uncertainty, our two-stage RO grants more feasibility by the implementation of wait-and-see decisions after the revelation of uncertainty realization. Compared with [5], [20], [21] that are established in a two-stage or multi-stage setting and postulate affine decision rules on the wait-and-see decisions, our model seeks to derive the exact optimal solution to the second-stage problem. In view of the DDUS, we extend the polyhedral uncertainty sets with reduced decision-dependency considered in [6], [8], [9], [15], [16] into a generic form where both the shape and size of the uncertainty set may be altered by decisions. Due to the aforementioned improved modeling capability, our model can considerably eradicate the conservatism effects of RO by proactively controlling the level of uncertainties and inherently capture the trade-off between the robustness and conservatism of the solution.

2) *From a Solution Standpoint*: To solve the two-stage RO model with DDUs, we propose a novel iterative solution algorithm involving one master problem and two sub-problems, based on the Benders dual decomposition. Compared with the C&CG algorithm [23], [24] which is widely applied to conventional two-stage RO problems, the proposed algorithm overcomes the challenges stemming from endogenous worst-case uncertainty scenarios by using only dual information of the second-stage problem. Also differently from classical Benders dual cutting-plane algorithms [25]–[27], advanced optimality cuts and feasibility cuts are designed to accommodate the coupling between uncertain parameters and decision variables. Performance of the proposed algorithm, including convergence and optimality, is guaranteed with a strict proof. To implement the algorithm, we derive the robust counterpart of the master problem and the reformulations of the sub-problems, maintaining the computational tractability of the proposed solution algorithm.

The remainder of the paper is organized as follows. The remainder of this section introduces the notation. The mathematical formulation of the two-stage RO with DDUs is provided in Section II. Section III proposes the solution approach based on modified Benders decomposition. Section IV

addresses some implementation issues of the proposed algorithm, including the derivation of a tractable robust counterpart of the master problem which is a static RO and the reformulation of the bi-level sub-problems into MILPs. In Section V, four examples are provided as the applications of the proposed model and solution methodology. And we present a case study on the pre-disaster highway network investment problem. We conclude the paper in Section VI.

Notation: \mathbb{R}^n is the set of n -dimensional real vectors and $\mathbb{R}^{m \times n}$ is the set of $m \times n$ real matrices. \mathbb{Z} (\mathbb{Z}^+) denotes the set of (positive) integers. $[n] \triangleq \{1, \dots, n\}$ denotes the set of integers from 1 to n and $[0] \triangleq \emptyset$. For a vector $x \in \mathbb{R}^n$ (a matrix $A \in \mathbb{R}^{m \times n}$), x^T (A^T) denotes its transpose. We use $\mathbf{1}$ and $\mathbf{0}$ to denote vectors of ones and zeros, respectively. For $x, y \in \mathbb{R}^n$, we denote the inner product $x^T y = \sum_{i=1}^n x_i y_i$ where x_i, y_i stand for the i -th entries of x and y , respectively. The Hadamard product is defined as $x \circ y = (x_1 y_1, \dots, x_n y_n)^T$, i.e., the element-wise product of two equally-sized vectors or matrices. For a polytope $X \subseteq \mathbb{R}^n$, $\text{vert}(X)$ is the set of the vertices of X . $[a]$ is the largest integer less than or equal to a . We define the difference of set X and Y as $X \setminus Y \triangleq \{x : x \in X, x \notin Y\}$.

II. PROBLEM FORMULATION

A. Two-Stage RO With DDU

The two-stage RO with DDU (TRO-DDU) problem is formulated as

$$\min_x \left\{ f(x) + \max_{w \in W(x)} \min_{y \in Y(x, w)} c^T y \right\} \quad (1a)$$

$$\text{s.t. } x \in X \cap X_R \quad (1b)$$

where

$$W(x) = \{w \in \mathbb{R}^{n_w} | Gw \leq g + h(x)\} \quad (1c)$$

$$Y(x, w) = \{y \in \mathbb{R}^{n_y} | Ax + By + Cw \leq b, y \geq \mathbf{0}\} \quad (1d)$$

$$X_R = \{x \in X | \forall w \in W(x), Y(x, w) \neq \emptyset\}. \quad (1d)$$

The here-and-now decision $x \in \mathbb{R}^{n_R} \times \mathbb{Z}^{n_Z}$ is a mixed-integer variable vector with a dimension of $n_x = n_R + n_Z$. $X \subseteq \mathbb{R}^{n_R} \times \mathbb{Z}^{n_Z}$ is the feasible region of x . $f(x) : \mathbb{R}^{n_R} \times \mathbb{Z}^{n_Z} \rightarrow \mathbb{R}^1$ is the cost function with respect to x . The DDU parameter is denoted by $w \in \mathbb{R}^{n_w}$ and the DDUS of w is $W(x) : \mathbb{R}^{n_R} \times \mathbb{Z}^{n_Z} \mapsto 2^{\mathbb{R}^{n_w}}$, which is essentially a set-valued map parameterized by x where $G \in \mathbb{R}^{r \times n_w}$ is a constant matrix, $g \in \mathbb{R}^r$ is a constant vector and $h(x) : \mathbb{R}^{n_R} \times \mathbb{Z}^{n_Z} \rightarrow \mathbb{R}^r$ is a vector-valued function with respect to x . The wait-and-see decision vector is $y \in \mathbb{R}^{n_y}$. We consider the case that the second-stage decision problem is a linear program (LP) and $c \in \mathbb{R}^{n_y}$ denotes the cost coefficient of y . $Y(x, w) : \mathbb{R}^{n_R} \times \mathbb{Z}^{n_Z} \times \mathbb{R}^{n_w} \mapsto 2^{\mathbb{R}^{n_y}}$, which is essentially a set-valued map parameterized by x and w , denotes the feasible region of y with a specific form in (1d) where $A \in \mathbb{R}^{m \times n_x}$, $B \in \mathbb{R}^{m \times n_y}$, $C \in \mathbb{R}^{m \times n_w}$ and $b \in \mathbb{R}^m$ are constant parameters. The constraint set $X_R \subseteq \mathbb{R}^{n_R} \times \mathbb{Z}^{n_Z}$ defined in (1e) is the set of x that are robust feasible. A here-and-now decision x is called robust feasible [1], [26] if for any realization of the uncertain w within $W(x)$ there exists at least one feasible wait-and-see decision y lying within $Y(x, w)$. The

TRO-DDU problem (1) aims at minimizing the total cost of the two stages, which is $f(x) + c^T y$, under the worst-case uncertainty realization. The following assumptions are made on problem (1).

Assumption 1: 1) $f(x)$ is a convex function; 2) X is a convex set; 3) $X \cap X_R \neq \emptyset$; 4) X is bounded, $W(x)$ is bounded for any $x \in X$ and $Y(x, w)$ is bounded for any $x \in X$ and $w \in W(x)$.

Under Assumption 1-1) and 1-2), the nominal problem of (1) is a convex minimization problem. Assumption 1-3) and 1-4) imply that the TRO-DDU problem (1) is feasible and bounded, thus there exists the optimal solution to (1).

The TRO-DDU problem distinguishes itself from the existing literature by the consideration of the DDU w and the DDUS $W(x)$. Regarding the generality of the DDU modeling in the TRO-DDU problem (1), we have the following remark.

Remark 1 (Decision-Dependent Uncertainty Set): The polyhedral DDUS $W(x)$ in (1c), whose right-hand-side vector has a dependency on x , covers the formulations proposed in the existing literature [8], [9], [14]–[16], [20], [21]. It provides generic modeling capability since both the shape and size of the uncertainty set can be altered by x . Moreover, $W(x)$ in (1c) applies readily with DIUs by setting $h(x) = \mathbf{0}$. Though the coefficient matrix G is fixed in (1c), the solution strategy proposed in Sections III and IV can be easily extended to the cases in [3]–[5] where x affects the coefficient matrix through certain binary-valued functions with respect to x .

Regarding the difficulty in solving TRO-DDU problem (1), we have the following remark.

Remark 2: The well-known C&CG algorithm is no longer applicable to TRO-DDU problem (1) since finite convergence to the global optimum is not warranted in the presence of DDUs.

1) *Failure in Robust Feasibility and Optimality:* The cutting planes in the C&CG algorithm with recourse decision variables in the primal space are generated with concrete values of the uncertain parameter for each identified scenario; and these cuts can become invalid when the uncertainty set depends on the here-and-now decision x which varies with iterations. For example, given a here-and-now decision x^1 , the identified worst-case uncertainty realization $w^1 \in W(x^1)$ would never come true if another here-and-now decision x^2 is adopted and $w^1 \notin W(x^2)$; and in that case, the feasibility and optimality cuts generated according to w^1 no longer necessarily perform a relaxation to problem (1) and may ruin the optimality of the solution, or even render an infeasible master problem.

2) *Failure in Finite Convergence:* The finite convergence of C&CG algorithm is justified by a complete enumeration of the vertices of the uncertainty set. However, in problem (1), the vertices set of polytope $W(x)$ changes with x which varies with iterations, thus the C&CG algorithm no longer necessarily terminates within a finite number of iterations.

III. SOLUTION METHOD

A. Equivalent Transformation

1) *Robust Optimality:* Looking at the inner-level problem in

(1), we denote the following max-min bi-level optimization by $S(x)$:

$$S(x) = \max_{w \in W(x)} \min_{y \in Y(x, w)} c^T y. \quad (2)$$

By noting the duality of the inner problem in $S(x)$, $S(x)$ can be equivalently transformed into the following bi-linear maximization problem:

$$S(x) = \max_{w, u} u^T (b - Ax - Cw) \quad (3a)$$

$$\text{s.t. } w \in W(x), u \in U \triangleq \{u | B^T u \leq c, u \leq 0\} \quad (3b)$$

where $u \in \mathbb{R}^m$ is the dual variable that corresponds with the constraints $Ax + By + Cw \leq b$. Then, problem (1) is equivalent to

$$\min_x f(x) + S(x) \quad \text{s.t. } x \in X \cap X_R. \quad (4)$$

Introducing a supplementary variable $\alpha \in \mathbb{R}^1$, then problem (4) can be further rewritten in an epigraph form as

$$\begin{aligned} \min_{x, \alpha} f(x) + \alpha \\ \text{s.t. } x \in X \cap X_R, S(x) \leq \alpha. \end{aligned} \quad (5)$$

2) *Robust Feasibility*: The robust feasibility of decision x , i.e., whether x lies in X_R , can be examined by solving the following relaxed bi-level problem:

$$R(x) = \max_w \min_{y, s} \mathbf{1}^T s \quad (6a)$$

$$\text{s.t. } w \in W(x) \quad (6b)$$

$$Ax + By + Cw \leq b + s \quad (6c)$$

$$y \geq 0, s \geq 0 \quad (6d)$$

where $s \in \mathbb{R}^m$ is the slack variable vector introduced to relax the feasible region of y defined in (1d). If $R(x) \leq 0$, x is robust feasible, i.e., $x \in X_R$. Else if $R(x) > 0$, there must exist a realization $w^* \in W(x)$ such that no feasible wait-and-see decision y is available. It is useful to write the dual of the inner minimization problem in $R(x)$. Then $R(x)$ is equivalent to the single-level bi-linear optimization problem as follows:

$$R(x) = \max_{w, \pi} \pi^T (b - Ax - Cw) \quad (7a)$$

$$\text{s.t. } w \in W(x) \quad (7b)$$

$$\pi \in \Pi \triangleq \{\pi | B^T \pi \leq 0, -1 \leq \pi \leq 0\} \quad (7c)$$

where $\pi \in \mathbb{R}^m$ is the dual variable vector corresponding with constraints (6c). Since $x \in X_R$ if and only if $R(x) \leq 0$, problem (5) can be further written as

$$\begin{aligned} \min_{x, \alpha} f(x) + \alpha \\ \text{s.t. } x \in X, S(x) \leq \alpha, R(x) \leq 0. \end{aligned} \quad (8)$$

From the above transformation, we derive problem (8) which is the surrogate model of TRO-DDU problem (1).

B. Modified Benders Decomposition Algorithm

Next, we have the overall Algorithm 1 to solve problem (8) where the master problem (MP) involved is formulated as

$$\min_{x, \alpha} f(x) + \alpha \quad (9a)$$

$$\text{s.t. } x \in X, \alpha \geq -M \quad (9b)$$

$$\alpha \geq u_j^{*T} (b - Ax - Cw), \forall w \in W(x), j \in [k_u] \quad (9c)$$

$$0 \geq \pi_j^{*T} (b - Ax - Cw), \forall w \in W(x), j \in [k_\pi]. \quad (9d)$$

Regarding the difference between the proposed modified Benders dual decomposition (Algorithm 1) and the existing algorithms such as the C&CG algorithm and classic Benders dual cutting-plane algorithms, we give a remark as below.

Algorithm 1 Modified Benders Dual Decomposition Algorithm for TRO-DDU Problem (1)

Input: The TRO-DDU problem (1), a sufficiently large positive number M , and the convergence tolerance $\varepsilon > 0$.

Output: The optimal solution x^k and the optimal value $(UB^k + LB^k)/2$.

```

1: Initialize  $k = 0, k_u = 0, k_\pi = 0, UB^0 = M, LB^0 = -M$ .
2: while True do
3:   Update  $k = k + 1$ ;
4:   Solve the MP (9) which is a static RO and let  $(x^k, \alpha^k)$  be the optimum;
5:   Update the lower bound of the optimal objective by  $LB^k = f(x^k) + \alpha^k$ ;
6:   Solve sub-problem  $R(x^k)$  in (7) and denote by  $(w^k, \pi^k) = \arg R(x^k)$ ;
7:   if  $R(x^k) > 0$  then
8:     Update  $k_\pi = k_\pi + 1$  and set  $\pi_{k_\pi}^* = \pi^k$ ;
9:   Update the upper bound of the optimal objective by  $UB^k = UB^{k-1}$ ;
10:  else
11:    Solve sub-problem  $S(x^k)$  in (3) and denote by  $(w^k, u^k) = \arg S(x^k)$ ;
12:    Update the upper bound of the optimal objective by  $UB^k = f(x^k) + S(x^k)$ ;
13:    Update  $k_u = k_u + 1$  and set  $u_{k_u}^* = u^k$ ;
14:    if  $|UB^k - LB^k| < \varepsilon$  then
15:      Break;
16:    end
17:  end
18: end

```

Remark 3 (Advanced Optimality and Feasibility Cuts): Constraints (9c) and (9d), which are appended to the master problem with iterations, are called optimality cuts and feasibility cuts, respectively. They are designed to have the following salient features to adapt to the DDUS: 1) Concrete worst-case uncertainty, i.e., the solution to $R(x^k)$ or $S(x^k)$ denoted by w^k , is not involved in the optimality cuts or the feasibility cuts. This is different from the C&CG algorithm, as it incorporates the coupling relationship between x and w ; 2) Dual information of sub-problems $R(x^k)$ and $S(x^k)$, i.e., π^k and u^k , are utilized to generate the feasibility cuts and optimality cuts, inspired by the Benders dual decomposition. However, these cuts are designed to be no longer hyperplanes, but a set of static robust constraints, to comprise a cluster of endogenous worst-case uncertainty realizations.

Regarding the extensive adaptability of Algorithm 1, we have the following remark.

Remark 4 (Adaptability to DDU Formulations): Algorithm 1 is designed to be applicable to a wide range of DDU formulations. Note that the explicit formulation of DDUS $W(x)$ is not involved in Algorithm 1, nor in the theoretical justification of its performance. Thus, Algorithm 1 does not preclude the implementation of more DDUS structures besides polyhedrons, such as ellipsoidal and conic sets. However, considering the difficulty in solving master problem (9) and sub-problems $R(x)$ and $S(x)$, here we focus on a class of polyhedral uncertainty sets with decision dependency in the right-hand-side vector, as formulated in (1c).

We justify the convergence and optimality of Algorithm 1 by the following theorem.

Theorem 1: Let q be the number of vertices of U defined in (3b) and p be the number of vertices of Π defined in (7c). Let x^* denote the optimal solution to the TRO-DDU problem (1). Then, Algorithm 1 terminates within $O(p+q)$ iterations and outputs the solution x^k such that $x^k \in X \cap X_R$ and $|f(x^k) + S(x^k) - f(x^*) - S(x^*)| < \varepsilon$.

Theorem 1 indicates that the proposed modified Benders dual decomposition algorithm is finitely convergent to the optimum. Proof of Theorem 1 can be found in Appendix.

IV. ROBUST COUNTERPART AND BIG-M REFORMULATION

In this section, we show that the implementation of Algorithm 1 is computationally tractable, by deriving the robust counterpart of the master problem (9) and reformulations of sub-problems $R(x^k)$ and $S(x^k)$.

A. Robust Counterpart of the Master Problem

In Algorithm 1, the master problem (9) involves static robust constraints (9c) and (9d). Without loss of generality, we illustrate how to deal with the robust constraint (9c) in the master problem by substituting it with its robust counterpart.

The following robust constraint with respect to (α, x) with given u_j^* :

$$\alpha \geq u_j^{*T} (b - Ax - Cw), \quad \forall w \in W(x) \quad (10)$$

is equivalent to

$$\alpha \geq u_j^{*T} (b - Ax) + \max_{w_j^u} -u_j^{*T} Cw_j^u \quad (11a)$$

$$\text{s.t. } Gw_j^u \leq g + h(x) \quad (11b)$$

where we substitute w by w_j^u for denoting variable in constraint formed by u_j^* . Next, the following two cases are discussed:

1) *A Bi-Linear Term $\lambda^T h(x)$ can be Precisely Linearized Through the Big-M Method Where $\lambda \in \mathbb{R}^r$ is an r -Dimensional Variable and x is an n_x -Dimensional Variable:* To illustrate this, we take $x \in \{0, 1\}^{n_x}$ and $h(x) = Hx$ as an example. Duality of the inner-level problem in (11) is deployed, thus (11) can be reformulated as

$$\begin{aligned} \alpha &\geq u_j^{*T} (b - Ax) + \min_{\lambda_j} (g + Hx)^T \lambda_j \\ \text{s.t. } G^T \lambda_j &= -C^T u_j^*, \quad \lambda_j \geq \mathbf{0} \end{aligned} \quad (12)$$

where $\lambda_j \in \mathbb{R}^r$ is the dual variable corresponding with constraint (11b). Note that constraint (12) is equivalent to

$$\begin{aligned} \alpha &\geq u_j^{*T} (b - Ax) + g^T \lambda_j + x^T H^T \lambda_j \\ G^T \lambda_j &= -C^T u_j^*, \quad \lambda_j \geq \mathbf{0} \end{aligned} \quad (13)$$

by dropping the minimization operation. Since x is binary, the bi-linear term $x^T H^T \lambda_j$ can be exactly linearized through the big-M method by introducing supplementary variable $\eta_j \in \mathbb{R}^{n_x}$ and a large enough positive number M , where $\eta_j = x \circ (H^T \lambda_j)$. Then, constraint (13) has the following equivalent formulation:

$$\alpha \geq u_j^{*T} (b - Ax) + g^T \lambda_j + \mathbf{1}^T \eta_j \quad (14a)$$

$$-Mx \leq \eta_j \leq Mx \quad (14b)$$

$$H^T \lambda_j - M(\mathbf{1} - x) \leq \eta_j \leq H^T \lambda_j + M(\mathbf{1} - x) \quad (14c)$$

$$G^T \lambda_j = -C^T u_j^*, \quad \lambda_j \geq \mathbf{0} \quad (14d)$$

$$\lambda_j \in \mathbb{R}^r, \quad \eta_j \in \mathbb{R}^{n_x} \quad (14e)$$

which constitutes the robust counterpart of constraint (10). The robust counterpart of feasibility cuts (9d) can be derived in a similar way. Thus the master problem (9) has the following MILP robust counterpart:

$$\min_{x, \alpha, \beta, \lambda, \eta, \gamma, \mu} \beta + \alpha \quad (15a)$$

$$\text{s.t. } x \in X, \quad \alpha \in \mathbb{R}^1, \quad \beta \in \mathbb{R}^1 \quad (15b)$$

$$\beta \geq a_i^1 x + a_i^0, \quad \forall i \in [I] \quad (15c)$$

$$(14a) - (14e), \quad \forall j \in [k_u] \quad (15d)$$

$$\begin{cases} 0 \geq \pi_j^{*T} (b - Ax) + g^T \gamma_j + \mathbf{1}^T \mu_j, \\ G^T \gamma_j = -C^T \pi_j^*, \quad \gamma_j \geq \mathbf{0}, \\ -Mx \leq \mu_j \leq Mx, \\ H^T \gamma_j - M(\mathbf{1} - x) \leq \mu_j \leq H^T \gamma_j + M(\mathbf{1} - x), \\ \gamma_j \in \mathbb{R}^r, \quad \mu_j \in \mathbb{R}^{n_x}, \end{cases} \quad \forall j \in [k_\pi] \quad (15e)$$

where the convex function $f(x)$ is substituted by its piecewise linearization form as in (15c) and the optimality cuts and the feasibility cuts are substituted by their robust counterparts. $a_i^1 \in \mathbb{R}^{n_x}$, $a_i^0 \in \mathbb{R}^1 (i \in [I])$ are constant parameters for the piecewise linearization of $f(x)$.

2) *Otherwise:* If a bi-linear term $\lambda^T h(x)$ cannot be exactly linearized by the big-M method, we deploy the Karush-Kuhn-Tucker (KKT) conditions of the inner-level problem in (11) as follows:

$$G^T \lambda_j = -C^T u_j^* \quad (16a)$$

$$\lambda_j \geq \mathbf{0} \perp Gw_j^u \leq g + h(x) \quad (16b)$$

where $\lambda_j \in \mathbb{R}^r$ is the dual variable corresponding with constraints (11b), and (16b) denotes the complementary relaxation condition. The nonlinearity of complementary condition (16b) can be eliminated through the big-M method by introducing the binary supplementary variable $z_j \in \{0, 1\}^r$.

Thus, the optimality cut (11) has the following equivalent robust counterpart:

$$\alpha \geq u_j^{*T}(b - Ax) - u_j^{*T}Cw_j^u \quad (17a)$$

$$G^T \lambda_j = -C^T u_j^* \quad (17b)$$

$$\mathbf{0} \leq \lambda_j \leq M(\mathbf{1} - z_j) \quad (17c)$$

$$\mathbf{0} \leq g + h(x) - Gw_j^u \leq Mz_j \quad (17d)$$

$$z_j \in \{0, 1\}^r, \lambda_j \in \mathbb{R}^r, w_j^u \in \mathbb{R}^{n_w} \quad (17e)$$

where the complementary relaxation condition (16b) is substituted by (17c) and (17d). The robust counterpart of the feasibility cut (9d) can be derived similarly. Thus the master problem (9) has the following robust counterpart which is a mixed-integer programming problem.

$$\begin{aligned} \min_{x, \alpha, \beta, \lambda, z, w^u, \gamma, v, w^\pi} \quad & \beta + \alpha \\ \text{s.t.} \quad & x \in X, \alpha \in \mathbb{R}^1, \beta \in \mathbb{R}^1 \\ & \beta \geq a_i^{1T}x + a_i^0, \forall i \in [I] \\ & (17a) - (17e), \forall j \in [k_u] \\ & \begin{cases} 0 \geq \pi_j^{*T}(b - Ax) - \pi_j^{*T}Cw_j^\pi, \\ G^T \gamma_j = -C^T \pi_j^*, \\ \mathbf{0} \leq \gamma_j \leq M(\mathbf{1} - v_j), \\ \mathbf{0} \leq g + h(x) - Gw_j^\pi \leq Mv_j, \\ v_j \in \{0, 1\}^r, \gamma_j \in \mathbb{R}^r, w_j^\pi \in \mathbb{R}^{n_w}, \end{cases} \quad \forall j \in [k_\pi]. \end{aligned} \quad (18)$$

B. Reformulation of Sub-Problems

In Algorithm 1, the robust feasibility examination sub-problem $R(x)$ in (7) and the robust optimality sub-problem $S(x)$ in (3) both have bi-linear objective, imposing difficulties on the solving. Next, we provide linear surrogate formulations of $S(x)$ and $R(x)$. Since they have similar structure, without loss of generality we only focus on $R(x)$.

$R(x)$ in (7) can be equivalently rewritten into

$$\max_{\pi \in \Pi} \left\{ \pi^T(b - Ax) + \max_{w \in W(x)} -\pi^T Cw \right\} \quad (19)$$

where the here-and-now decision x is given by solving the master problem. We deploy the KKT condition of the inner-level LP problem in (19) as follows:

$$\zeta \geq \mathbf{0} \perp Gw \leq g + h(x) \quad (20a)$$

$$G^T \zeta = -C^T \pi \quad (20b)$$

where $\zeta \in \mathbb{R}^r$ is the dual variable corresponding with constraint $w \in W(x)$ in (19). The complementary constraints (20a) can be linearized by introducing binary supplementary variable $z \in \{0, 1\}^r$. Then, the complementary constraints (20a) can be substituted with its equivalent linear formulations, like what we have done to (16b). Moreover, since strong duality holds, we substitute the $-\pi^T Cw$ in (19) by $(g + h(x))^T \zeta$. Then, sub-problem $R(x)$ is equivalent to the following MILP problem:

$$\begin{aligned} \max_{\pi, w, z, \zeta} \quad & \pi^T(b - Ax) + (g + h(x))^T \zeta \\ \text{s.t.} \quad & \mathbf{0} \leq \zeta \leq M(\mathbf{1} - z) \\ & \mathbf{0} \leq g + h(x) - Gw \leq Mz \\ & G^T \zeta = -C^T \pi \\ & \pi \in \Pi, w \in \mathbb{R}^{n_w}, z \in \{0, 1\}^r, \zeta \in \mathbb{R}^r. \end{aligned} \quad (21)$$

V. APPLICATIONS

The proposed model and solution algorithm cater for a variety of application problems. In this section, we provide four motivating DDU-featured examples that can be formulated into the proposed TRO-DDU model (1). For the first three applications, we focus on the endogeneity of the uncertainties by constructing the DDUS and do not go into the details of problem formulation. For the last application, pre-disaster network investment, detailed problem statements and numerical case studies are provided with an out-of-sample analysis.

A. Application 1: Batch Scheduling Problem

This case is from [13] (Section 8) where stochastic programming formulations with endogenous uncertainties are established. Here we reform the characterization of DDUs to fall in the scope of RO. Consider a chemical process network as shown in Fig. 1.

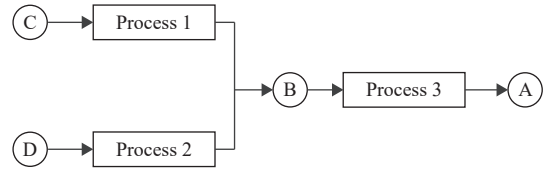


Fig. 1. Illustration of Application 1.

Here, chemical A is produced in Process 3 from chemical B while chemical B can be produced in Process 2 from chemical D or produced in Process 1 from chemical C. If needed, chemicals A, B, C, and D can be purchased from market. Now there is a demand for chemical A that must be satisfied and decisions on which specific processes should be operated are made to maximize the net profit. The per unit yields of the three processes, denoted by $\theta_{i,t}$ for each process $i \in \mathcal{I} = \{1, 2, 3\}$ and time $t \in \mathcal{T}$, are uncertain. The endogeneity stems from the fact that $\theta_{i,t}$ retains a physical meaning only when process i is operated at time period t . Let $x_{i,t} \in \{0, 1\}$ denote the binary decision of operating process i at time t , then the DDUS for θ is constructed as

$$W(x) = \left\{ \begin{array}{l} \theta \in \mathbb{R}^{|\mathcal{T}||\mathcal{I}|} : \\ -x_{i,t}M \leq \theta_{i,t} \leq x_{i,t}M, \forall i \in \mathcal{I}, t \in \mathcal{T} \\ \tilde{\theta}_{i,t} + M(x_{i,t} - 1) \leq \theta_{i,t}, \forall i \in \mathcal{I}, t \in \mathcal{T} \\ \theta_{i,t} \leq \tilde{\theta}_{i,t} + M(1 - x_{i,t}), \forall i \in \mathcal{I}, t \in \mathcal{T} \\ (1 - \delta)\theta_{i,t}^0 \leq \tilde{\theta}_{i,t}, \forall i \in \mathcal{I}, t \in \mathcal{T} \\ \tilde{\theta}_{i,t} \leq (1 + \delta)\theta_{i,t}^0, \forall i \in \mathcal{I}, t \in \mathcal{T} \end{array} \right\} \quad (22)$$

where $\theta_{i,t}^0$ is the nominal value of $\theta_{i,t}$ when $x_{i,t}$ equals 1, $\tilde{\theta}_{i,t}$ is

the supplementary variable and M is a sufficiently large positive constant.

B. Application 2: Shortest Path Over an Uncertain Network

The second case comes from [16] (Example 2), considering to find the shortest path from the origin node A to the destination node B over a network with uncertain arc lengths, as shown in Fig. 2.

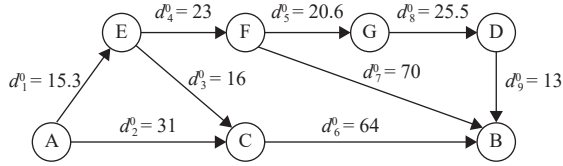


Fig. 2. Illustration of Application 2.

Let \mathcal{E} denote the arcset of the graph and the uncertain length for an arc $e \in \mathcal{E}$ is denoted by d_e . The uncertain arc lengths d lie in the DDUS as follows:

$$W(x) = \left\{ \begin{array}{l} d_e = d_e^0(1 + 0.5\xi_e), \quad \forall e \in \mathcal{E} \\ d \in \mathbb{R}^{|\mathcal{E}|} \mid 0 \leq \xi_e \leq 1 - 0.8x_e, \quad \forall e \in \mathcal{E} \\ \sum_{e \in \mathcal{E}} \xi_e \leq 1 \end{array} \right\} \quad (23)$$

where d_e^0 denotes the nominal value of d_e and ξ_e is the uncertain parameter that controls the value of d_e . For each arc $e \in \mathcal{E}$, $x_e \in \{0, 1\}$ is the binary decision that determines the variation range of ξ_e . If x_e attains the value of 1, the maximum possible uncertainty of ξ_e is reduced to 0.2. Else when $x_e = 0$, the maximum possible uncertainty of ξ_e is left at 1. For simplicity, it is assumed that at most one of the arcs can achieve such uncertainty reduction. In addition, it is required in (23) that the sum of ξ_e over the entire \mathcal{E} is at most 1.

C. Application 3: Frequency Reserve Provision

The third application comes from [8]. Consider a smart building participating in the demand response program on electricity consumption. The building is asked by the grid operator to adjust its electricity consumption within a pre-specified reserve capacity. The building is rewarded for providing this reserve service and the payment depends on both the size of the reserve capacity and the actually deployed reserve energy.

Consider a symmetric reserve capacity $\mathcal{R}_t(X_t) = [-X_t, X_t]$ where $X_t \geq 0$ is the size of reserve capacity at time t . For the smart building, the operator's real-time reserve deployment request r_t is endogenously uncertain at the planning stage. The endogeneity stems from the fact that the range of reserve deployment that the smart building admits is fundamentally affected by its decisions on the reserve capacity. Specifically, $r_t \in \mathcal{R}_t(X_t)$ and X_t is the building's day-ahead decision. It is assumed that providing reserve capacity of size X_t is rewarded by $\lambda_t^{\text{DA}} X_t$ and offering reserve energy of size r_t is compensated by $\lambda_t^{\text{RT}} r_t$. The frequency reserve provision problem is formulated to maximize the building's profit while always satisfying inner technical and comfort constraints.

D. Application 4: Pre-Disaster Network Investment

In this subsection, a pre-disaster highway network investment problem is provided, including its problem formulation and a case study. This application is motivated by the work of [28] where the decision maker aims to strengthen the highway system in advance to prevent link failures due to earthquakes. By contrast, we model the problem into a two-stage RO with DDUs for the case that link failure probabilities are not known.

The goal is to guarantee the existence of a path between a given origin-destination (O-D) pair after the earthquake and concurrently minimize the post-disaster travel cost from the origin node to the destination node and the costs of necessary pre-disaster investment. Endogenous uncertainties arise from the fact that the post-disaster state of each link, either functional or non-functional, is uncertain but can be altered by pre-disaster reinforcement.

1) *Problem Formulation*: The high-way network is modeled into a directed graph $N = (V, E)$ with node set V and arc set E . To characterize the functionality of the highway system after the disaster, two nodes in the graph N are specified: D denotes the destination node which is the district with the highest expected damage in the earthquake, and O denotes the origin node which usually refers to the district with the most support resources.

We use binary-valued variable $x_e \in \{0, 1\}^1$ to denote the investment decision on link $e \in E$. x_e takes the values of 1 if there is an investment on link e and 0 otherwise. Let $w_e \in \{0, 1\}^1$ denote the decision-dependent uncertain post-disaster state of link $e \in E$. w_e attains the value of 1 if the link e is not functional after the occurrence of the disaster and the value of 0 otherwise. Strengthened links have guaranteed functionality after the disaster and the non-strengthened links are subject to random failures. Thus the realization of w_e is restricted by

$$w_e \leq 1 - x_e, \quad \forall e \in E \quad (24a)$$

$$\sum_{e \in E} w_e \leq \lfloor \psi \sum_{e \in E} (1 - x_e) \rfloor \quad (24b)$$

where $\psi \in [0, 1]$ is the robustness budget to reduce conservatism. Note that the constraint matrix of (24) satisfies total unimodularity, thus the binary-valued w_e can be relaxed to be continuous without any compromise [3]. Thus, the budgeted DDUS for w_e is constructed as

$$W(x) = \{w_e \in \mathbb{R}^1, \quad \forall e \in E : (24a) - (24b), \quad 0 \leq w_e \leq 1\}. \quad (25)$$

To verify the post-disaster connectivity from O to D as well as find the path from O to D with minimal traversal costs after the disaster, the following second-stage optimization problem is considered:

$$\begin{aligned} \min_y \quad & \sum_{e \in E} c_e y_e \\ \text{s.t.} \quad & y \in Y(w) \end{aligned} \quad (26)$$

where $Y(w) \subseteq \mathbb{R}^{|\mathcal{E}|}$ is defined through the following constraints:

$$0 \leq y_e \leq 1 - w_e, \quad \forall e \in E \quad (27a)$$

$$\sum_{e=(O,j) \in E} y_e - \sum_{e=(j,O) \in E} y_e = 1 \quad (27b)$$

$$\sum_{e=(D,j) \in E} y_e - \sum_{e=(j,D) \in E} y_e = -1 \quad (27c)$$

$$\sum_{e=(i,j) \in E} y_e - \sum_{e=(j,i) \in E} y_e = 0, \quad \forall i \in V \setminus \{O, D\} \quad (27d)$$

and $c_e \in \mathbb{R}^1$ denotes the length of link e . The objective is to minimize the traversal cost from O to D in the network. The wait-and-see decision variable $y_e \in \{0, 1\}^1$ takes the values of 1 if the path from O to D goes through link e and 0 otherwise. Note that in (27), the binary-valued y_e is substituted by its continuous relaxation (27a) since the constraint matrix of y satisfies total unimodularity property. There exists a path from O to D after the occurrence of the disaster if and only if (26) has at least one feasible solution and the optimal solution to (26) implies the path with least traversal cost from O to D in the network.

Next, the mathematical formulation of the two-stage robust pre-disaster highway network investment problem is given

$$\min_{x \in X_R} \left\{ \sum_{e \in E} a_e x_e + \max_{w \in W(x)} \min_{y \in Y(w)} \sum_{e \in E} c_e y_e \right\} \quad (28a)$$

$$\text{s.t. } x_e \in \{0, 1\}^1, w_e \in \mathbb{R}^1, y_e \in \mathbb{R}^1, \quad \forall e \in E \quad (28b)$$

$$x \in X_R \triangleq \{x : \forall w \in W(x), Y(w) \neq \emptyset\} \quad (28c)$$

where $a_e \in \mathbb{R}^1$ denotes the investment cost of link e . $W(x)$ and $Y(w)$ are defined in (25) and (27), respectively. The objective is to minimize the traversal costs between the O-D pair under the worst-case uncertainty realization as well as the necessary investment cost. Constraint (28c) guarantees the existence of at least one path from O to D after the disaster under any realization of w in the uncertainty set $W(x)$. The proposed modified Benders dual decomposition algorithm is applied to solve problem (28). Note that given a variable $\lambda \in \mathbb{R}^1$, the bi-linear item $\lambda[\psi \sum_{e \in E} (1 - x_e)]$ can be precisely linearized, thus the robust counterpart of the master problem in Algorithm 1 can be formulated into an MILP problem.

2) *Basic Results*: The computational results are based on a network with 8 nodes and 9 links, as demonstrated in Fig. 3. The relevant data are from [28]. The traversal costs (lengths) and investment costs of links are provided in Table I. The O-D pair is chosen as (1, 6), i.e., $O = 1$, $D = 6$. For the nominal 9-link network, there are 4 possible paths from O to D , as listed in Table II. If all links remain functional, the shortest path from O to D is Path 1 (1-3-5-9) with a length of 13.52. The program runs on an Intel Core-i5 1.6-GHz computer and is coded with YALMIP. CPLEX 12.6.0 is utilized as the solver.

The proposed modified Benders dual decomposition algorithm is applied with an initialization of $UB^0 = 4000$, $LB^0 = 0$, $\varepsilon = 0.01$, $\psi = 0.3$. The algorithm reaches convergence of $UB = LB = 1100.65$ after 8 rounds of iterations and the evolution process of UB^k and LB^k is shown in Fig. 4. The optimal robust investment (the here-and-now decision) is on Link 3, Link 8, and Link 9. For all possible network realizations under this pre-disaster investment scheme, the worst-case is the failure on Link 5 after the earthquake. In that

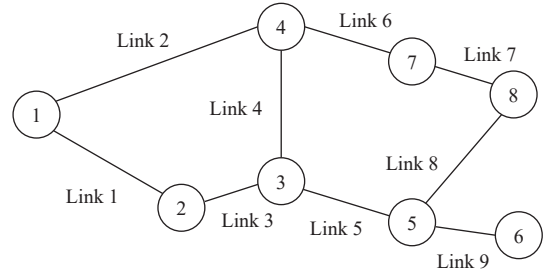


Fig. 3. The 8-node 9-link network.

TABLE I
LINK LENGTHS AND INVESTMENT COSTS

Link e	1	2	3	4	5	6	7	8	9
c_e	6.41	8.09	1.97	6.35	2.87	4.11	2.27	3.91	2.27
a_e	500	620	160	780	260	220	500	120	800

TABLE II
PATHS FOR O-D PAIR (1,6)

	Links	Length
Path 1	1-3-5-9	13.52
Path 2	2-4-5-9	19.58
Path 3	2-6-7-8-9	20.65
Path 4	1-3-4-6-7-8-9	27.29

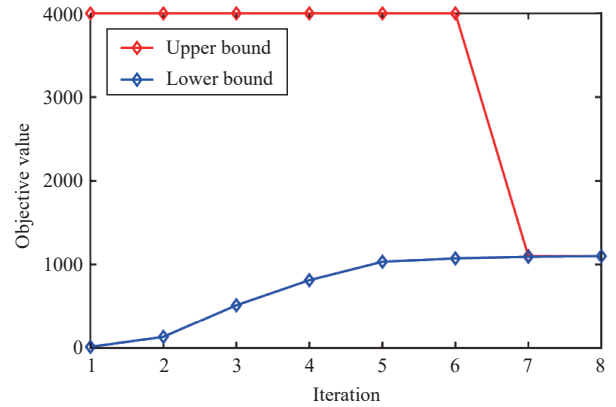


Fig. 4. Evolution of the UB^k and LB^k .

case, the shortest post-disaster path from O to D (the wait-and-see decision) is Path 3 with a length of 20.65. The correctness of the computational results can be verified by enumeration.

3) *Comparative Study*: To emphasize the necessity and superiority of the proposed algorithm for DDU-involved two-stage RO problems, existing cutting plane algorithms that are theoretically applicable to only the case of DIUs are adopted to solve the same problem in (28). In both the classical Benders decomposition algorithm and the C&CG algorithm, the master problem becomes infeasible at the second round, then the solution procedure interrupts without convergence. This is because the worst-case uncertainty realization identified on the first round is a failure in Link 9, under which there is no path from O to D . This situation can be eliminated by strengthening Link 9; however, the invalid feasibility cut

TABLE III
IMPACT OF ROBUSTNESS PARAMETER ψ

ψ	Total cost	Investment cost	Travel cost	Reinforced links	Worst-case non-functional links	Shortest path
0	13.52	0	13.52	–	–	Path 1
0.1	13.52	0	13.52	–	–	Path 1
0.2	820.65	800	20.65	9	5	Path 3
0.3	1100.65	1080	20.65	3,8,9	5	Path 3
0.4	1579.58	1560	19.58	3,5,6,8,9	1	Path 2
0.5	1733.52	1720	13.52	1,3,5,9	7,8	Path 1
0.6	1733.52	1720	13.52	1,3,5,9	6,7,8	Path 1

generated with this concrete scenario is appended to the master problem and the consequent infeasibility falsely implies that there is no feasible solution to the original problem (28). The observations in this comparative case study verify the statements in Remark 2.

4) *Sensitivity Analysis*: We present the impact of robustness budget ψ on the optimal solution to problem (28). In this case, 7 robustness parameters ψ from 0 to 0.6 with a gradient of 0.1 are introduced and the results are displayed in Table III. It is observed that, as ψ increases, the optimal objective value including investment costs and worst-case traversal costs increases accordingly. This is because the decision maker has to enhance more links to hedge against the increasing uncertainties on link failures. However, the optimal pre-disaster investment saturates when ψ is greater than or equal to 0.5. This is due to the fact that no matter how large ψ is, there always exists a path between the O-D pair if links 1,3,5,9 (Path 1) are all reinforced.

5) *Out-of-Sample Analysis*: Finally, an out-of-sample analysis is carried out to verify the performance of the robust investment decision. Steps of the out-of-sample assessment are as follows [29]: a) Considering all possible states of the 9 links, we generate $2^9 = 512$ scenarios, namely $w^o \in \{0,1\}^9$, $o \in [512]$. b) Given x^* derived by solving (28), for each scenario o , the following relaxed second-stage problem is solved:

$$\min_{y^o, s^{o+}, s^{o-}} \sum_{e \in E} c_e y_e^o + P \sum_{i \in V} (s_i^{o+} + s_i^{o-}) \quad (29a)$$

$$\text{s.t. } 0 \leq y_e^o \leq 1 - w_e^o(1 - x_e^*), \quad \forall e \in E \quad (29b)$$

$$s^{o+}, s^{o-} \in \mathbb{R}^8, \quad s^{o+} \geq \mathbf{0}, \quad s^{o-} \geq \mathbf{0} \quad (29c)$$

$$\sum_{e=(O,j) \in E} y_e^o - \sum_{e=(j,O) \in E} y_e^o = 1 + s_O^{o+} - s_O^{o-} \quad (29d)$$

$$\sum_{e=(D,j) \in E} y_e^o - \sum_{e=(j,D) \in E} y_e^o = -1 + s_D^{o+} - s_D^{o-} \quad (29e)$$

$$\sum_{e=(i,j) \in E} y_e^o - \sum_{e=(j,i) \in E} y_e^o = s_i^{o+} - s_i^{o-}, \quad \forall i \in V \setminus \{O, D\} \quad (29f)$$

where non-negative slack variables s^{o+} , s^{o-} are introduced to the flow conservation constraints (29d)–(29f) and penalized in the objective (29a). The penalty cost coefficient $P \in \mathbb{R}^1$ is set as 5000 in this case. Problem (29) is a modified version of the

second-stage problem (26) with a given here-and-now decision x^* and the given uncertainty realization w_o . Denote the solution to (29) by y^{o*} , s^{o+*} , s^{o-*} . c) At last, the average sampled second-stage cost C^{av} and the average sampled infeasibility level s^{av} are computed as follows:

$$\begin{cases} C^{\text{av}} = \frac{1}{512} \sum_{o \in [512]} \left(\sum_{e \in E} c_e y_e^{o*} + P \sum_{i \in V} (s_i^{o+*} + s_i^{o-*}) \right) \\ s^{\text{av}} = \frac{1}{512} \sum_{o \in [512]} \sum_{i \in V} (s_i^{o+*} + s_i^{o-*}). \end{cases} \quad (30)$$

Table IV shows the out-of-sample assessment results for the robust optimal solution under different values of the robustness budget ψ . It is observed that C^{av} and s^{av} decrease with the increasing ψ , indicating that a bigger ψ leads to a more robust solution that can hedge against a higher level of uncertainties, but also accordingly generates higher investment cost in the first stage. Comparison with the deterministic model is also provided in Table IV. We can see that disregarding uncertainty would give rise to significantly high second-stage costs (due to the penalty on the slack variables s_i^{o+} and s_i^{o-}) and infeasibility level. By choosing a proper robustness budget ψ , the decision maker can achieve the trade-off between the first-stage investment costs and the second-stage feasibility level.

TABLE IV
RESULTS FROM THE OUT-OF-SAMPLE ANALYSIS

ψ	Investment cost	Reinforced links	C^{av}	s^{av}
Deterministic	0	–	6775.8122	1.3535
0.1	0	–	6775.8122	1.3535
0.2	800	9	4875.2299	0.9727
0.3	1080	3,8,9	3528.4452	0.7031
0.4	1560	3,5,6,8,9	1575.22	0.3125
0.5	1720	1,3,5,9	13.52	0
0.6	1720	1,3,5,9	13.52	0

VI. CONCLUSIONS

In this paper, a novel two-stage RO model with DDUs is proposed. We introduce a class of polyhedral DDUs whose right-hand-side vectors are in a dependency of the here-and-now decision. Solution methodology for the problem is designed based on modified Benders decomposition, robust

counterpart derivation, and linearization techniques. Computational tractability and convergence performance of the proposed algorithm is guaranteed by a strict proof.

Case studies on the pre-disaster highway network investment problem verify that the proposed DDU-incorporated two-stage RO model is an amenable framework for addressing decision-making problems under endogenous uncertainties. The optimality and feasibility of the robust solution are validated by enumeration in this case. Furthermore, the computational studies elucidate that the DDU-involved two-stage RO model inherently captures the trade-off between uncertainty mitigation (line investment) and the corresponding expenses (line investment costs) and provides less conservative robust results.

Though we focus on polyhedral DDUS whose right-hand-side vector has a general correlation with the here-and-now decision, the proposed solution framework based on the modified Benders decomposition in Section III is also applicable to DDUSs with ellipsoidal or conic structure. The limitation of the work in this paper lies in the fact that the duality-based Benders decomposition is conditional upon the second-stage problem which is an LP. Also, uncertain parameters that are of discrete or binary nature would prohibit us from formulating the robust counterpart of the master problem as described in Section IV. These issues remain to be addressed in our future study. In addition, as an extension of this work, future study could combine the optimality and feasibility cuts (9c) and (9d) with the so-called Pareto cuts [30], or extend the robust counterpart (17) to incorporate primal cuts with recourse decision variables in the primal space, to further improve the convergence rate in practice.

APPENDIX

Before we give the proof of Theorem 1, a crucial lemma is provided as follows.

Lemma 2: Let x^* denote the optimal solution to TRO-DDU problem (1). For any $k \in \mathbb{Z}^+$:

- a) $LB^k \leq f(x^*) + S(x^*) \leq UB^k$;
- b) $\forall j \in [k_\pi]$, $\pi_j^* \in \text{vert}(\Pi)$. Moreover, $\forall j_1, j_2 \in [k_\pi]$ and $j_1 \neq j_2$, there is $\pi_{j_1}^* \neq \pi_{j_2}^*$;
- c) $\forall j \in [k_u]$, $u_j^* \in \text{vert}(U)$. Moreover, $\forall j_1, j_2 \in [k_u - 1]$ and $j_1 \neq j_2$, there is $u_{j_1}^* \neq u_{j_2}^*$. If there exists $j \in [k_u - 1]$ such that $u_{k_u}^* = u_j^*$, then Algorithm 1 terminates at k .

Proof of Lemma 2.

Assertion a): Recalling the formulation of function $R(x)$ in (7) and $S(x)$ in (3), problem (8) can be rewritten as

$$\begin{aligned} \min_{x \in X, \alpha \in \mathbb{R}^1} \quad & f(x) + \alpha \\ \text{s.t.} \quad & \pi^T(b - Ax - Cw) \leq 0, \quad \forall w \in W(x), \quad \forall \pi \in \Pi \\ & u^T(b - Ax - Cw) \leq \alpha, \quad \forall w \in W(x), \quad \forall u \in U. \end{aligned} \quad (31)$$

Since the TRO-DDU problem (1), problem (8) and problem (31) are equivalent, $f(x^*) + S(x^*)$ is also the optimal objective value of (31). Also note that the master problem (9) is a relaxation to minimization problem (31) since $u_j^* \in U$ for all $j \in [k_u]$ and $\pi_j^* \in \Pi$ for all $j \in [k_\pi]$. Thus, the optimal objective value of problem (9), which is denoted by LB^k , is less than or

equal to the optimal objective value of problem (31), i.e., $LB^k \leq f(x^*) + S(x^*)$.

Recall the updating rule of the upper bound which is $UB^k = UB^{k-1}$ when $R(x^k) > 0$ and $UB^k = f(x^k) + S(x^k)$ when $R(x^k) = 0$. Next, we prove $UB^k \geq f(x^*) + S(x^*)$ by induction. First of all, $UB^0 = M > f(x^*) + S(x^*)$. Suppose for the sake of induction that $UB^{k-1} \geq f(x^*) + S(x^*)$ where $k \in \mathbb{Z}^+$. Then, if $UB^k = UB^{k-1}$, we have $UB^k \geq f(x^*) + S(x^*)$ directly. Else if $UB^k = f(x^k) + S(x^k)$ when $R(x^k) = 0$, then we have $UB^k = f(x^k) + S(x^k) \geq f(x^*) + S(x^*)$ since x^k is a feasible solution to the TRO-DDU problem (1) (by recalling that $x^k \in X_R$ if and only if $R(x^k) = 0$) and $f(x^*) + S(x^*)$ is the optimal objective value of TRO-DDU problem (1).

Assertion b): $\pi_j^* \in \text{vert}(\Pi)$ can be easily verified by noting that the optimal solution of a bi-linear program with polyhedral feasible region can be achieved at one of the vertices of the polytopes [31].

Next, we prove that $\forall j_1, j_2 \in [k_\pi]$ and $j_1 \neq j_2$, $\pi_{j_1}^* \neq \pi_{j_2}^*$. Suppose for the sake of contradiction that there exist $j_1, j_2 \in [k_\pi]$ and $j_1 \neq j_2$ such that $\pi_{j_1}^* = \pi_{j_2}^*$. Without loss of generality we assume that $j_1 < j_2$, and thus $j_1 \in [j_2 - 1]$ since $j_1, j_2 \in \mathbb{Z}^+$. Suppose $\pi_{j_2}^*$ is the optimal solution to $R(x^{k_{j_2}})$, there must be $R(x^{k_{j_2}}) > 0$, implying that

$$\max_{w \in W(x^{k_{j_2}})} \pi_{j_2}^{*T}(b - Ax^{k_{j_2}} - Cw) > 0. \quad (32)$$

Since $\pi_{j_1}^* = \pi_{j_2}^*$, we have

$$\max_{w \in W(x^{k_{j_2}})} \pi_{j_1}^{*T}(b - Ax^{k_{j_2}} - Cw) > 0. \quad (33)$$

Recall that $x^{k_{j_2}}$ is the optimal solution to the master problem with the following feasibility cuts:

$$0 \geq \pi_j^{*T}(b - Ax - Cw), \quad \forall w \in W(x), \quad j \in [j_2 - 1]. \quad (34)$$

Since $j_1 \in [j_2 - 1]$, there must be

$$0 \geq \pi_{j_1}^{*T}(b - Ax^{k_{j_2}} - Cw), \quad \forall w \in W(x^{k_{j_2}}) \quad (35)$$

which contradicts with (33).

Assertion c): Similarly to the proof of Lemma 2-b), $u_j^* \in \text{vert}(U)$ can be verified by noting that the optimal solution of bi-linear programming over a polytope can be achieved at one of its vertices.

Suppose there exist $j_1, j_2 \in [k_u]$ and $j_1 \neq j_2$ such that $u_{j_1}^* = u_{j_2}^*$. We assume that $u_{j_1}^*$ is the optimum to $S(x^{k_{j_1}})$ and $u_{j_2}^*$ is the optimum to $S(x^{k_{j_2}})$. Without loss of generality, we assume that $j_1 < j_2$, and thus $j_1 \in [j_2 - 1]$ since $j_1, j_2 \in \mathbb{Z}^+$. Since $(x^{k_{j_2}}, \alpha^{k_{j_2}})$ is the optimal solution to the master problem with the following optimality cuts and $j_1 \in [j_2 - 1]$,

$$\alpha \geq u_{j_1}^{*T}(b - Ax - Cw), \quad \forall w \in W(x), \quad j \in [j_2 - 1] \quad (36)$$

then, we have

$$\alpha^{k_{j_2}} \geq u_{j_1}^{*T}(b - Ax^{k_{j_2}} - Cw), \quad \forall w \in W(x^{k_{j_2}}) \quad (37)$$

which is equivalent to

$$\alpha^{k_{j_2}} \geq \max_{w \in W(x^{k_{j_2}})} u_{j_1}^{*T}(b - Ax^{k_{j_2}} - Cw). \quad (38)$$

Recalling the updating rule of lower bound, we have

$$\begin{cases} LB^{k_{j_2}} = f(x^{k_{j_2}}) + a^{k_{j_2}} \\ (i) \geq f(x^{k_{j_2}}) + \max_{w \in W(x^{k_{j_2}})} u_{j_1}^{*T} (b - Ax^{k_{j_2}} - Cw) \\ (ii) = f(x^{k_{j_2}}) + \max_{w \in W(x^{k_{j_2}})} u_{j_2}^{*T} (b - Ax^{k_{j_2}} - Cw) \\ (iii) = f(x^{k_{j_2}}) + S(x^{k_{j_2}}) \stackrel{(iv)}{=} UB^{k_{j_2}} \end{cases} \quad (39)$$

where *i*) comes from the relationship in (38), *ii*) comes from the fact that $u_{j_1}^* = u_{j_2}^*$, *iii*) comes from the fact that $u_{j_2}^*$ is the optimum to $S(x^{k_{j_2}})$, and *iv*) comes from the updating rule of the upper bound. Thus, we have $UB^{k_{j_2}} \leq LB^{k_{j_2}}$. According to Lemma 2-a), $LB^{k_{j_2}} \leq f(x^*) + S(x^*) \leq UB^{k_{j_2}}$. Thus, we have $LB^{k_{j_2}} = UB^{k_{j_2}} = f(x^*) + S(x^*)$ and the Algorithm 1 terminates at iteration k_{j_2} . This completes the proof of the statement in Lemma 2-c). ■

The proof of Theorem 1 is given as follows.

Proof of Theorem 1: First, we prove that Algorithm 1 terminates within a finite round of iterations through a complete enumeration of the vertices of the polyhedral feasible regions of the dual multipliers. According to Lemmas 2-b) and 2-c), no vertex of Π or U will be appended twice to the master problem. Thus, the termination condition must be met within $O(p+q)$ iterations.

Next, we prove the feasibility of solution x^k . Since x^k is generated through master problem (9), $x^k \in X$. Moreover, since Algorithm 1 terminates with $R(x^k) = 0$, we have $x^k \in X_R$ by recalling that $x \in X_R$ if and only if $R(x) = 0$. Thus, $x^k \in X \cap X_R$.

Finally, we prove the optimality of solution x^k . According to Lemma 2-a), $LB^k \leq f(x^*) + S(x^*) \leq UB^k$. Together with the fact that the Algorithm 1 terminates with $|UB^k - LB^k| < \varepsilon$ and $R(x^k) = 0$, we have $|UB^k - f(x^*) - S(x^*)| < \varepsilon$. Recalling that $UB^k = f(x^k) + S(x^k)$, we have $|f(x^k) + S(x^k) - f(x^*) - S(x^*)| < \varepsilon$, which justifies the optimality of x^k . ■

REFERENCES

- [1] A. Ben-Tal and A. Nemirovski, "Robust solutions of uncertain linear programs," *Oper. Res. Lett.*, vol. 25, no. 1, pp. 1–13, 1999.
- [2] Y. Chen, W. Wei, F. Liu, and S. Mei, "Distributionally robust hydrothermal-wind economic dispatch," *Appl. Energy*, vol. 173, pp. 511–519, 2016.
- [3] N. H. Lappas and C. E. Gounaris, "Robust optimization for decisionmaking under endogenous uncertainty," *Comput. Chem. Eng.*, vol. 111, pp. 252–266, 2018.
- [4] N. H. Lappas and C. E. Gounaris, "The use of decision-dependent uncertainty sets in robust optimization," in *Proc. Foundations of Computer-Aided Process Operations/Chemical Process Control*, 2017.
- [5] N. H. Lappas and C. E. Gounaris, "Multi-stage adjustable robust optimization for process scheduling under uncertainty," *AIChE Journal*, vol. 62, no. 5, pp. 1646–1667, 2016.
- [6] R. Vujanic, P. Goulart, and M. Morari, "Robust optimization of schedules affected by uncertain events," *J. Optim. Theory Appl.*, vol. 171, no. 3, pp. 1033–1054, 2016.
- [7] C. Wang, F. Liu, J. Wang, F. Qiu, W. Wei, S. Mei, and S. Lei, "Robust risk-constrained unit commitment with large-scale wind generation: An adjustable uncertainty set approach," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 723–733, Jan. 2017.
- [8] X. Zhang, M. Kamgarpour, A. Georghiou, P. Goulart, and J. Lygeros, "Robust optimal control with adjustable uncertainty sets," *Automatica*, vol. 75, pp. 249–259, 2017.
- [9] Y. Su, Y. Zhang, F. Liu, S. Feng, Y. Hou, and W. Wang, "Robust dispatch with demand response under decision-dependent uncertainty," in *Proc. IEEE Sustainable Power and Energy Conf.*, 2020, pp. 2122–2127.
- [10] M. Daneshvar, B. Mohammadi-Ivatloo, K. Zare, and S. Asadi, "Twostage robust stochastic model scheduling for transactive energy based renewable microgrids," *IEEE Trans. Ind. Informat.*, vol. 16, no. 11, pp. 6857–6867, Nov. 2020.
- [11] M. Johnston, H. Lee, and E. Modiano, "A robust optimization approach to backup network design with random failures," *IEEE/ACM Trans. Netw.*, vol. 23, no. 4, pp. 1216–1228, Aug. 2015.
- [12] K. Park, J. Kim, H. Lim, and Y. Eun, "Robust path diversity for network quality of service in cyber-physical systems," *IEEE Trans. Ind. Informat.*, vol. 10, no. 4, pp. 2204–2215, Nov. 2014.
- [13] V. Goel and I. E. Grossmann, "A class of stochastic programs with decision dependent uncertainty," *Math. Program.*, vol. 108, no. 2, pp. 355–394, 2006.
- [14] M. Poss, "Robust combinatorial optimization with variable budgeted uncertainty," *4OR – A Quarterly J. Operations Research*, vol. 11, no. 1, pp. 75–92, 2013.
- [15] M. Poss, "Robust combinatorial optimization with variable cost uncertainty," *Eur. J. Oper. Res.*, vol. 237, no. 3, pp. 836–845, 2014.
- [16] O. Nohadani and K. Sharma, "Optimization under decision-dependent uncertainty," *SIAM J. Optim.*, vol. 28, no. 2, pp. 1773–1795, 2018.
- [17] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, "Adjustable robust solutions of uncertain linear programs," *Math. Program.*, vol. 99, no. 2, pp. 351–376, 2004.
- [18] P. Hansen, B. Jaumard, and G. Savard, "New branch-and-bound rules for linear bilevel programming," *SIAM J. Sci. Statist. Comput.*, vol. 13, no. 5, pp. 1194–1217, 1992.
- [19] N. H. Lappas and C. E. Gounaris, "Theoretical and computational comparison of continuous-time process scheduling models for adjustable robust optimization," *AIChE Journal*, vol. 64, no. 8, pp. 3055–3070, 2018.
- [20] W. Feng, Y. Feng, and Q. Zhang, "Multistage robust mixed-integer optimization under endogenous uncertainty," *Eur. J. Oper. Res.*, vol. 294, no. 2, pp. 460–475, 2020.
- [21] Q. Zhang and W. Feng, "A unified framework for adjustable robust optimization with endogenous uncertainty," *AIChE Journal*, vol. 66, p. e.17047, 2020.
- [22] S. Avraamidou and E. N. Pistikopoulos, "Adjustable robust optimization through multi-parametric programming," *Optim. Lett.*, vol. 14, no. 4, pp. 873–887, 2020.
- [23] L. Zhao and B. Zeng, "Robust unit commitment problem with demand response and wind energy," in *Proc. IEEE Power and Energy Soc. Gen. Meeting*, 2012, pp. 1–8.
- [24] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Oper. Res. Lett.*, vol. 41, no. 5, pp. 457–461, 2013.
- [25] R. Jiang, M. Zhang, G. Li, and Y. Guan, "Benders' decomposition for the two-stage security constrained robust unit commitment problem," in *Proc. IIE Annual Conf.*, 2012, pp. 3332–3341.
- [26] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 52–63, Feb. 2013.
- [27] J. Benders, "Partitioning procedures for solving mixed-variables programming problems," *Numer. Math. (Heidelb)*, vol. 4, no. 1, pp. 238–252, 1962.
- [28] S. Peeta, F. Salman, D. Gunec, and K. Viswanath, "Pre-disaster investment decisions for strengthening a highway network," *Comput. Oper. Res.*, vol. 37, pp. 1708–1719, 2010.
- [29] A. Baringo, L. Baringo, and J. M. Arroyo, "Day-ahead self-scheduling of a virtual power plant in energy and reserve electricity markets under uncertainty," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 1881–1894, May 2019.
- [30] T. L. Magnanti and R. T. Wong, "Accelerating benders decomposition:

Algorithmic enhancement and model selection criteria,” *Oper. Res.*, vol. 29, no. 3, pp. 464–484, 1981.

- [31] H. Konno, “A cutting plane algorithm for solving bilinear programs,” *Math. Program.*, vol. 11, no. 1, pp. 14–27, 1976.

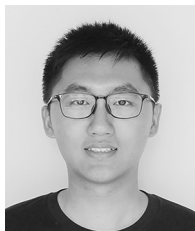


Yunfan Zhang received the B.Sc. degree in electrical engineering from Tsinghua University in 2018, where she is currently a Ph.D. candidate. Her research interests include applied optimization, energy economics, game theory and their applications.



Feng Liu (Senior Member, IEEE) received the B.Sc. and Ph.D. degrees in electrical engineering from Tsinghua University, in 1999 and 2004, respectively. He is currently an Associate Professor of Tsinghua University. From 2015 to 2016, he was a Visiting Associate at California Institute of Technology, USA. His research interests include stability analysis, optimal control, robust dispatch and game theory based decision making in energy and power systems.

He is the author/coauthor of more than 200 peer-reviewed technical papers and two books, and holds more than 20 issued/pending patents. He is an Associated Editor of several international journals including *IEEE Transactions on Smart Grid* and *IEEE PES Letter*. He also served as a Guest Editor of *IEEE Transactions on Energy Conversion*.



Yifan Su received the B.Sc. degree in electrical engineering from Tsinghua University in 2019, where he is currently a Ph.D. candidate. His research interests include distributed optimization and robust dispatch in smart grid.



Yue Chen (Member, IEEE) received the B.E. degree in electrical engineering from Tsinghua University in 2015, the B.S. degree in economics from Peking University in 2017, and the Ph.D. degree in electrical engineering from Tsinghua University in 2020. From 2018 to 2019, she was a Visiting Student with the California Institute of Technology, USA. She is currently an Assistant Professor with the Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China. Her research interests include optimization, game theory, mathematical economics, and their applications in smart grid and integrated energy systems.



Zhaojian Wang (Member, IEEE) received the B.Sc. degree from Tianjin University in electrical engineering in 2013, and the Ph.D. degree from Tsinghua University in electrical engineering in 2018. From 2016 to 2017, he was a Visiting Ph.D. Student at California Institute of Technology, USA. From 2018 to 2020, he was a Post-Doctoral Scholar at Tsinghua University. He is currently an Assistant Professor at Shanghai Jiao Tong University. His research interests include stability analysis, optimal control, microgrid planning and game theory based decision making in energy and power systems. He is an Associate Editor of *IEEE Systems Journal* and *IET Renewable Power Generation*.



João P. S. Catalão (Fellow, IEEE) received the M.Sc. degree from the Instituto Superior Técnico (IST), Portugal in 2003, and the Ph.D. degree and Habilitation for Full Professor (“Agregação”) from the University of Beira Interior (UBI), Portugal, in 2007 and 2013, respectively. Currently, he is a Professor at the Faculty of Engineering of the University of Porto (FEUP), Portugal, and Research Coordinator at INESC TEC. He was the Primary Coordinator of the EU-funded FP7 project SiNGULAR (“Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration”), a 5.2-million-euro project involving 11 industry partners. His research interests include power system operations and planning, power system economics and electricity markets, distributed renewable generation, demand response, smart grid, and multi-energy carriers.