Two-stage Robust Energy Storage Planning with Probabilistic Guarantees: A Data-driven Approach

Chao Yan[a,1], Xinbo Geng[b,1], Zhaohong Bie[a,2], Le Xie[b,3,*]

[a]Department of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, 710049.
[b]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, United States, 77843.

Abstract

In conventional planning decision making, shorter-term (e.g., hourly) variations are not explicitly accounted for. However, given the deepening penetration of variable resources, it is becoming imperative to consider such shorter-term variation in the longer-term planning exercise. This paper addresses a central challenge of jointly considering such shorter-term and longer-term uncertainties in power system planning with increasing penetration of renewable and storage resources. By leveraging the abundant operational observation data, we propose a scenario-based robust planning framework that provides rigorous guarantees on the future operation risk of planning decisions considering a broad range of operational conditions, such as renewable generation fluctuations and load variations. By connecting two-stage robust optimization with the scenario approach theory, we show that with a carefully chosen number of scenarios, the operational risk level of the robust solution can be adaptive to the risk preference set by planners. The theoretical guarantees hold true for any distributions, and the proposed approach is scalable towards real-world power grids. Furthermore, the column-and-constraint generation algorithm is used to solve the two-stage robust planning problem and tighten theoretical guarantees. We substantiate this framework through a planning problem of energy storage in a power grid with deep renewable penetration. Case studies are performed on large-scale test systems (modified IEEE 118-bus system) to illustrate the theoretical bounds as well as the scalability of the proposed algorithm.

Keywords: Power system planning, energy storage, robust optimization, the scenario approach, column-and-constraint generation, short-term uncertainty, operation risk.

1. Introduction

Power system planning refers to a decision making process typically involves a time span of multiple years, and is uncertain by nature [4]. Conventionally, power system planning mainly incorporates *longer-term* uncertainties such as load growth and fuel prices. With the rapid growth of renewable resources, *shorter-term* uncertainties such as renewable fluctuations and load variations become imperative in the longer-term planning exercise [5]. If not accounted for, the short-term uncertainties will likely render an overly expensive or unreliable planning outcome in the long-term.

To facilitate the integration of rapidly growing renewable resources, energy storage is being deployed at an accelerated pace in power systems. From 2014 to 2019, the installed capacity of energy storage increased by 35.7% from 24.6 GW to 33.4 GW in the United States [1, 2]. According to China Energy Storage Alliance (CNESA), 32.3 GW of energy storage has been installed in China as of 2019. Furthermore, Wood Mackenzie recently predicted that energy storage is poised for a decade-defining boom, with capacity set to grow by almost 33% worldwide every year in the 2020s to reach around 741 GWh by 2030 [3]. Consequently, the optimal siting and sizing of energy storage, i.e., storage planning, becomes pivotal to build carbon-neutral and reliable power systems.

To consider shorter-term uncertainties in longer-term power system planning, stochastic optimization (SO) and robust optimization (RO) are the most commonly adopted approaches. SO is built on probabilistic modeling of uncertainties. Since accurate probabilistic models are often unavailable or expensive to obtain, a massive number of samples are needed [9]. This, however, will significantly increase computational burden [7]. RO, on the other hand, relies on a set-based (often deterministic) modeling of uncertainties. *Robust* planning models uncertainties by a pre-defined uncertainty set, which describes a range. Therefore, reducing the computation burden of planning problem. Common choices of uncertainty sets include intervals [10] and polyhedral uncertainty sets [11]. However, this choice of these uncertainties is often scrutinized for returning overly conservative solutions. How to construct the reasonable uncertainty set of these short-term uncertainties such as load variation and renewable fluctuation in the long-term planning decision is the core difficulty of current robust power system planning (RSP) problems.

Some recent studies have been proposed to tackle this chal-

---

lenge in RSP, most of them rely on historical data to construct the uncertainty set. For instances, [12] extracted representative scenarios from historical operation data (HOD) to model short-term operation uncertainties; [14] constructed adaptive uncertainty sets to consider the risk from wind uncertainties in RSP. In most of the proposed approaches (e.g., [14]), short-term uncertainties are represented by a single operating point, thus temporal correlations among uncertainties are neglected. Since the principal value of energy storage devices come from offering temporal flexibilities, it is pivotal to model temporal correlations in the context of storage planning, Reference [38] constructed uncertainty sets as the convex hull of historical wind and load (time-varying) profiles. Similarly, [13] built polyhedral uncertainty sets for representative days by clustering HOD thus captures the temporal dynamics of uncertainties.

In this paper, we demonstrate that a simple construction of uncertainty sets, i.e., the collection of a carefully chosen number of i.i.d. scenarios, possesses many desirable features and advantages. By tuning a risk parameter (violation probability $\bar{\epsilon}$) and the associated number of scenarios (K), planning decisions can be adapted to risk preferences. Meanwhile, the scenarios can be daily operating conditions considering the temporal couplings. This simple construction of uncertainty set is closely related with the scenario approach theory [15, 16, 17, 18]. The scenario approach has been successfully applied to many power system problems, e.g., resource adequacy and security assessment [19], economic dispatch [20], demand response scheduling [21], and unit commitment [22, 23].

The scenario approach theory is a classical data-driven mathematical program theory with rigorous probabilistic guarantees [15, 16, 17, 18] and also has been applied to power system field [19, 20, 22]. In this paper, it is developed with the two-stage robust program to address the above short-term uncertainty modeling issues in the power system planning. It is the first paper applying the scenario approach to two-stage robust power system planning. Specifically, we show that a simple construction of uncertainty sets, i.e., the collection of a carefully chosen number of i.i.d. operational scenarios, could resolve the issue of conservative solutions of the robust optimization . By tuning a risk parameter (violation probability ϵ) and the associated number of scenarios (K), planning decisions can be adapted to risk preferences. Meanwhile, the scenarios can be daily operational conditions considering temporal operational conditions. Finally, we addressed two very important research problems that have never been successfully resolved in the existing research: (i) Why and how can we leverage the historical operational data to represent the operational risk in the long-term power system planning problem based on a rigorous mathematical theory? (ii) How can we quantitatively control the operational risk considered in the robust power system planning decision through simply adjusting the data-based uncertainty set of short-term uncertainties, and make a trade-off between the robustness, investment costs and operational risks for the planning of storage system needing the consideration of temporal operational conditions? After resolving these problems, we makes the following contributions:

(1) We propose a novel two-stage robust storage planning framework to facilitate the integration of renewables based on the scenario approach. Distribution-free theoretical guarantees are provided on the operational risks of the planning solutions of convex and non-convex robust storage planning models for future short-term uncertainties. These operational risks include (a) future operation cost risk; and (b) the load loss risk.

(2) Although the scenario approach is commonly used for single-stage decision making, we establish the connection between two-stage RO and the scenario approach theory; the theoretical guarantees on operational risk are obtained via this connection by examining the cardinality of invariant sets. We further proposes to use C&CG algorithm to find invariant set numerically thus improve the theoretical risk guarantees.

(3) We leverage the theoretical risk guarantees and randomized property from data to make a balance between the robustness, investment costs and operational risk in the robust storage planning problem.

(4) The proposed storage planning approach is tested on the IEEE 118-bus system with realistic wind and load data obtained from Electric Reliability Council of Texas (ERCOT). Numerical results show that the cardinality of invariant sets of convex or non-convex two-stage robust storage planning (RSP) problems is always small, regardless of problem size or power system scale. Consequently, we can achieve the same guarantee on future risk using a moderate number of scenarios, making the proposed approach computationally scalable.

The proposed RSP framework in this paper lies at the intersection among four areas (Figure 1) and makes unique contributions in both methodology and domain application perspectives. The proposed framework is computationally efficient, exploits the value from pervasive operation data, and provides generic and rigorous guarantees on planning solutions.
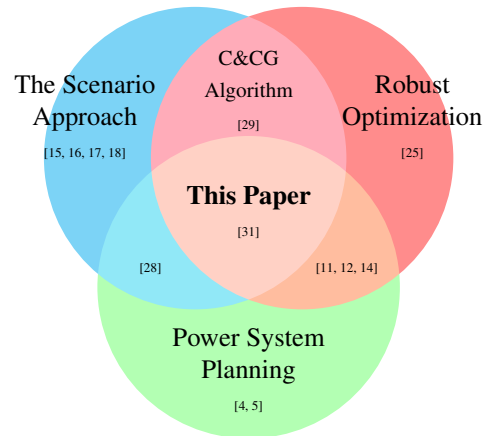

Figure 1: This paper connects power system planning, the scenario approach, robust optimization, and C&CG algorithm. A more detailed review is in [47].

The notations in this paper are standard. All matrices and vectors are in the real field $\mathbb{R}$. Matrices and vectors are in bold fonts, e.g., $\boldsymbol{A}$ and $\boldsymbol{b}$. The transpose of a vector $\boldsymbol{a}$ is $\boldsymbol{a}^{\mathsf{T}}$, and the jth entry of vector $\boldsymbol{a}$ is a_j. Sets are in calligraphy fonts, e.g. the set of transmission lines $\mathcal{L}$. A set consisting of K elements $\boldsymbol{\delta}^{(k)}, \cdots, \boldsymbol{\delta}^{(K)}$ is denoted by $\{\boldsymbol{\delta}^{(k)}, \cdots, \boldsymbol{\delta}^{(K)}\}$ or $\{\boldsymbol{\delta}^{(k)}\}_{k=1}^{|\mathcal{K}|}$ in short. The cardinality of a set $\mathcal{S}$ is $|\mathcal{S}|$. Related variables are repre-

sented by the same letter but distinct superscripts. For instance, $p^{\mathcal{L}}_{l,t}$ denotes the line flows, and $p^{\mathcal{G}}_{i,n,t}$ is the generation output. The upper and lower bounds on variable a are denoted by $\bar{a}$ and $\underline{a}$, respectively.

2. Storage Planning

Energy storage is a key resource to facilitate the integration of renewable energy resources by providing operational flexibility and ancillary services [32]. As of 2019, PJM has deployed approximately 300 MW of energy storage [33]; about 20 MW grid-scale battery-storage projects have been online in ISO New England since 2015, and nearly 2300 MW of grid-scale stand-alone energy-storage projects are in the queue to be interconnected [34]. Appropriate placements of energy storage systems could maximize the value of storage on the secure and economic operation of power systems. Storage planning, decising the optimal siting and sizing of energy storage[35, 36], becomes a topic of increasing importance in power system planning. Similar with other planning problems, various approaches have been proposed to deal with uncertainties, e.g., stochastic storage planning [37], robust optimization planning [38], and chance-constrained planning [39]. As storage is a special type of resources providing temporal flexibility, the consideration of temporal operational dynamics of short-term uncertainties is crucial in the planning exercise to effectively reduce future risk.

2.1. Nomenclature

For notation simplicity, we use $(\cdots)$s to denote the collection of variables, e.g., $\{(E_n, P_n, z_n)\}_{n \in \mathcal{N}}$ is denoted by (E_n, P_n, z_n)s. Unless specified, all decision variables are continuous. For simplicity, we assume that storage systems can be installed at every bus, i.e., $\mathcal{S} = \mathcal{N}$, thus all variables related with storage systems is indexed by the bus index n, e.g., $e^{\text{SOC}}_{n,t}$.

2.2. Deterministic Storage Planning

We first introduce the deterministic storage planning problem (1), which is a two-stage optimization problem.

$$\min_{x \in \Xi_1} C(x) + K_d \min_{y \in \Xi_2(x;\delta)} c(y) \tag{1}$$

2.2.1. Objective

There are two common storage planning formulations: cost-minimizing formulation and curtailment-minimizing formulation. They only differ in objectives.

Cost-minimizing Storage Planning. The objective of the cost-minimizing formulation is (2). The *first stage* determines the optimal energy and power ratings $x = (E_n, P_n, z_n)$ of the storage units at bus n by minimizing total investment cost (2a). The *second stage* optimizes the daily operation cost (2b) of the planned storages $(p^{\text{ch}}_{n,t}, p^{\text{dis}}_{n,t})$, generation $p^{\mathcal{G}}_{i,n,t}$, and load curtailment $p^{\text{shed}}_{n,t}$, i.e., $y = (p^{\text{ch}}_{n,t}, p^{\text{dis}}_{n,t}, p^{\mathcal{G}}_{i,n,t}, p^{\text{shed}}_{n,t}, v_{n,t})$.

$$C(x) := \sum_{n \in \mathcal{N}} (C^{\text{E}}_n E_n + C^{\text{P}}_n P_n), \tag{2a}$$

$$c(y) := \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} \sum_{i \in \mathcal{G}(n)} c^{\mathcal{G}}_{i,n} p^{\mathcal{G}}_{i,n,t} + \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} (c^{\text{ch}}_n p^{\text{ch}}_{n,t}$$
$$+ c^{\text{dis}}_n p^{\text{dis}}_{n,t}) + \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} c^{\text{shed}}_n p^{\text{shed}}_{n,t}. \tag{2b}$$

The first term $C(x)$ in the objective represents the annualized investment of storage systems. Cost coefficients for the energy and power rating investments are denoted by C^{E}_n (unit:\$ / MWh) and C^{P}_n (unit:\$ / MW), which are converted to present values[4]. For conceptual simplicity, the second stage uses a typical day to represent the operation of a whole year, i.e. $K_d = 365$. It is straightforward to extend towards employing D typical days to represent the annual operation cost $\sum_{i=1}^{D} K_i c_i(y)$.

This paper studies large-scale energy storage investment at the transmission level, and assumes that storage investment cost (including the land and construction cost), scales linearly with storage power and energy ratings. The marginal production cost coefficients of energy storage (e.g., from battery efficiency loss or degradation) c^{ch}_n and c^{dis}_n are assumed to be constants. Note that different storage technologies usually have different values for those cost parameters. Compressed air energy storage (CAES) has high investment costs for power ratings but low investment cost for capacity, the marginal production cost is negligible. For lithium batteries (LiBES) batteries, their degradation largely depends on daily operations, the aging of LiBES can be modeled as marginal production cost. Without loss of generality, this paper assumes that battery degradation leads to constant marginal costs in each charging and discharging cycle.

Curtailment-minimizing Storage Planning. Equation (4) is the objective function of the curtailment-minimizing storage planning formulation, which minimizes total load curtailment in the worst-case scenario by investing on storages in the first stage. More specifically, there will be no first-stage cost, i.e., $C(x) = 0$ in (1). The storage planning formulation using (4) is termed curtailment-minimizing storage planning.

$$c(y) := \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} p^{\text{shed}}_{n,t}. \tag{4}$$

---

[4]Let σ be the annual interest and Γ be the investment period, e.g. 20 years. C^{E}_n and C^{P}_n are computed as (3), where $C^{E,\text{total}}_n$ and $C^{P,\text{total}}_n$ represent the total storage energy and power annual investment cost over Γ years.

$$C^{\text{E}}_n = C^{\text{E,total}}_n \frac{\sigma(1+\sigma)^\Gamma}{(1+\sigma)^\Gamma - 1}, \; C^{\text{P}}_n = C^{\text{P,total}}_n \frac{\sigma(1+\sigma)^\Gamma}{(1+\sigma)^\Gamma - 1}. \tag{3}$$

Nomenclature

$\alpha_{n,t}^d$ load factor at bus n at time t.

$\alpha_{n,t}^w$ wind capacity factor at bus n at time t;

δ_s uncertainties in storage planning, $\delta_s := (\alpha_{n,t}^w, \alpha_{n,t}^d)$s;

$\eta_n^{\mathrm{ch}}, \eta_n^{\mathrm{dis}}$ storage charge/discharge efficiency;

$\mathcal{G}(n)$ the set of generators at bus n;

$\mathcal{L}$ the set of transmission lines;

$\mathcal{N}$ the set of buses;

$\mathcal{S}$ potential locations of storage systems;

$\mathcal{T}$ the set of time;

$\mathcal{W}$ the set of wind farms.

$\bar{\rho}, \underline{\rho}$ maximum/minimum power/energy ratio;

$\overline{d_n}$ peak load at node n.

$\overline{R}_i, \underline{R}_i$ generator ramp up/down limit;

$\overline{w_n}$ maximum wind capacity at node n;

$\theta_{n,t}$ nodal voltage angle;

$\underline{p_i^{\mathcal{G}}}, \overline{p_i^{\mathcal{G}}}$ generator output lower/upper limit;

$\underline{p_l^{\mathcal{L}}}, \overline{p_l^{\mathcal{L}}}$ transmission power flow lower/upper limit;

C^{Budget} storage investment budget;

$c_i^{\mathcal{G}}$ hourly incremental generation cost;

c_n^{shed} cost of load curtailment;

$c_n^{\mathrm{dis}}, c_n^{\mathrm{ch}}$ hourly incremental discharging/charging cost, related with the degradation cost of storage units;

C_n^{E} annualized storage energy investment cost;

C_n^{P} annualized storage power investment cost;

$e_{n,t}^{\mathrm{SOC}}$ state of charge (SOC) of a storage;

E_n the energy capacity (MWh) of the storage system;

i, j, k indices for general purposes;

K_d the weight of operational cost;

$o(l), r(l)$ sending and receiving nodes of a line, i.e., $l = (o(l), r(l))$;

$p_{i,n,t}^{\mathcal{G}}$ generator output;

$p_{l,t}^{\mathcal{L}}$ line power flow;

$p_{n,t}^{\mathrm{ch}}, p_{n,t}^{\mathrm{dis}}$ the charging and discharging power (MW) of the storage system at node n;

$p_{n,t}^{\mathrm{shed}}$ nodal load curtailment/shedding.

P_n the power capacity (MW) of the storage system;

q^E quantized energy capacity;

q^P quantized power capacity;

s, t, l, n indices for storages, time, lines and nodes

$v_{n,t}$ binary variable, indicating charging/discharging status of storage devices;

x_l transmission line reactance;

z_n integer variables, number of storage units.

2.2.2. First-stage Constraints

Constraints for the *first stage* $x \in \Xi_1$ is defined below:

$$\Xi_1 := \left\{ (E_n, P_n, z_n)\text{s that satisfy constraints (5)(6)(7)} \right\}$$

We consider the case where the capacities of storage units are quantized (instead of being continuous), parameters q^{E} and q^{P} denote the smallest quantized energy and power ratings for one energy storage unit.

$$z_n \in \{0, 1, 2, \cdots, \bar{z}_n\}, \ n \in \mathcal{N}. \tag{5}$$

Integer variable z_n is the number of storage units to be installed at bus n, which is constrained by the total quantity limit of storage investment (6a). As computed in (6b) and (6c), E_n and P_n represent the energy and power capacities of the storage system at node n. For some energy storage, the size of power and energy rating of storage unit is fixed due to physical limitations. Their investments need to be modeled by (6). For other energy storage without physical limitations on size, their investments can be directly represented by E_n, P_n.

$$\sum_{n \in \mathcal{N}} z_n \leq \bar{z}, \tag{6a}$$

$$q^{\mathrm{E}} z_n = E_n, \ n \in \mathcal{N}, \tag{6b}$$

$$q^{\mathrm{P}} z_n = P_n, \ n \in \mathcal{N}. \tag{6c}$$

Constraint (7) limits total investment within budget.

$$\sum_{n \in \mathcal{N}} (C_n^{\mathrm{E}} E_n + C_n^{\mathrm{P}} P_n) \leq C^{\mathrm{budget}}. \tag{7}$$

2.2.3. Second-stage Constraints

Constraints for the *second stage* $y \in \Xi_2(x; \delta)$ are the set of daily operational constraints. Note the feasible region Ξ_2 depends on the first-stage decision $x = (E_n, P_n)$ and uncertainties $\delta = (\alpha_{n,t}^w, \alpha_{n,t}^d)$.

$$\Xi_2(E_n, P_n; \delta) := \Big\{ (p_{n,t}^{\mathrm{ch}}, p_{n,t}^{\mathrm{dis}}, p_{i,n,t}^{\mathcal{G}}, p_{n,t}^{\mathrm{shed}}, v_{n,t})\text{s}$$
$$\text{that satisfy constraints (11)(9)(10)} \Big\}. \tag{8}$$

The inter-temporal operations of energy storage are (9) and (10), e.g., the state of charge (SOC) update from $t-1$ to t in (9a). Since (9) focuses on short-term operations, the degradation of storage systems in each charging/discharging cycle is

neglected. The charging and discharging power are limited by (9b)-(9c). We assume that the power limits P_n for charging and discharging are the same. SOC is limited to be within energy capacity (9d).

$$e_{n,t}^{\text{SOC}} - e_{n,t-1}^{\text{SOC}} = p_{n,t}^{\text{ch}}\eta_n^{\text{ch}} - p_{n,t}^{\text{dis}}/\eta_n^{\text{dis}}, \tag{9a}$$

$$0 \le p_{n,t}^{\text{ch}} \le P_n \cdot v_{n,t}, \tag{9b}$$

$$0 \le p_{n,t}^{\text{dis}} \le P_n \cdot (1 - v_{n,t}), \tag{9c}$$

$$0 \le e_{n,t}^{\text{SOC}} \le E_n, \tag{9d}$$

$$n \in \mathcal{N}, \ t \in \mathcal{T}.$$

Binary variables $v_{n,t}$ are introduced to avoid charging and discharging at the same time. $v_{n,t} = 1$ indicates the storage system at node n is charging at t; $v_{n,t} = 0$ when discharging:

$$v_{n,t} \in \{0, 1\}, \ n \in \mathcal{N}, \ t \in \mathcal{T}. \tag{10}$$

For the storage system like CAES, it can charge and discharge at the same time, the binary variable $v_{n,t}$ can be relaxed to be continuous.

This paper studies the storage investment in transmission system. Constraint (11) models the secure operation of the transmission power system. (11a) is the power balance at every node; (11b)-(11c) are the capacity and ramping limits of generators. Because this paper studies storage planning at the transmission level, DC power flow equations (11f) are used. All transmission lines, transformers and phase shifters are modeled with a common branch model, consisting of a standard (AC) Π transmission line model. DC power flow equations are obtained via linearizing the AC transmission line model. Line flow limits are in (11d). Constraints on load shedding are (11e). The actual wind generation in (11) is the product of wind capacity $\overline{w_n}$ and the wind capacity factor $\alpha_{n,t}^w$. The actual load is the product of the peak load $\overline{d_n}$ and load factor $\alpha_{n,t}^d$.

$$\sum_{i \in \mathcal{G}(n)} p_{i,n,t}^{\mathcal{G}} + p_{n,t}^{\text{dis}} - p_{n,t}^{\text{ch}} + \alpha_{n,t}^w \overline{w_n} + \sum_{l|n \in r(l)} p_{l,t}^f -$$
$$\sum_{l|n \in o(l)} p_{l,t}^f + p_{n,t}^{\text{shed}} = \alpha_{n,t}^d \overline{d_n}, \tag{11a}$$

$$\underline{p_i^{\mathcal{G}}} \le p_{i,n,t}^{\mathcal{G}} \le \overline{p_i^{\mathcal{G}}}, \tag{11b}$$

$$-\underline{R_i} \le p_{i,n,t}^{\mathcal{G}} - p_{i,n,t-1}^{\mathcal{G}} \le \overline{R_i}, \tag{11c}$$

$$\underline{p_l^{\mathcal{L}}} \le p_{l,t}^{\mathcal{L}} \le \overline{p_l^{\mathcal{L}}}, \tag{11d}$$

$$0 \le p_{n,t}^{\text{shed}} \le \alpha_{n,t}^d \overline{d_n}, \tag{11e}$$

$$p_{l,t}^{\mathcal{L}} = \frac{1}{x_l}(\theta_{o(l),t} - \theta_{r(l),t}), \tag{11f}$$

$$i \in \mathcal{G}(n), \ n \in \mathcal{N}, \ l \in \mathcal{L}, \ t \in \mathcal{T}.$$

3. Background

3.1. Two-stage Robust Optimization

A standard two-stage robust optimization problem is (12).

$$\min_{x \in \mathcal{X}}\left(c^\intercal x + \max_{\delta \in \Delta} \min_{y \in \mathcal{Y}(x;\delta)} d^\intercal y\right) \tag{12}$$

It seeks the objective-minimizing solution (x^*, y^*) for the worst scenario $\delta \in \Delta$ in a pre-defined *uncertainty set* Δ. Without loss of generality (see Appendix Appendix A.1.1), we assume deterministic objectives, i.e., no randomness associated with c and d. Sets $\mathcal{X} \subseteq \mathbb{R}^{n_x}$ and $\mathcal{Y} \subseteq \mathbb{R}^{n_y}$ denote the constraints for the first and second stages, respectively. Note that the second stage constraint $\mathcal{Y}(x;\delta)$ is determined by the first stage decision x and uncertainty δ. We follow the convention that $d^\intercal y = +\infty$ if the second-stage problem $\min_{y \in \mathcal{Y}(x;\delta)} d^\intercal y$ is infeasible.

Definition 1 (Feasible Solution)**.** A tuple (x, y) is a *feasible solution* to (12) if (i) $x \in \mathcal{X}$; and (ii) $y \in \mathcal{Y}(x;\delta)$ for all $\delta \in \Delta$. Equivalently, (x, y) is feasible if it has a finite objective value.

Remark 1 (Optimal and Infeasible Solutions)**.** We say that (x^*, y^*) is an (globally) *optimal solution* to (12), if $c^\intercal x^* + d^\intercal y^* \le c^\intercal x + d^\intercal y$ for all feasible solutions (x, y). For any potentially better solution $(x^\diamond, y^\diamond)$ with $c^\intercal x^\diamond + d^\intercal y^\diamond < c^\intercal x^* + d^\intercal y^*$, there always exists $\delta^\diamond \in \Delta$ such that $(x^\diamond, y^\diamond)$ is infeasible, i.e., $y^\diamond \notin \mathcal{Y}(x^\diamond; \delta^\diamond)$.

Throughout this paper, we only consider the case in which all constraints are linear. Matrix $Q(\delta)$ is the *recourse* matrix.

$$\mathcal{X} := \{x \in \mathbb{R}^{n_x} : Ax \le b\} \tag{13a}$$

$$\mathcal{Y}(x;\delta) := \{y \in \mathbb{R}^{n_y} : Gy \le h, T(\delta)x + Q(\delta)y \le r(\delta)\} \tag{13b}$$

In addition, we focus on the cases where the second-stage problem is feasible, which is formally defined as *relatively complete recourse*. This is mainly for the purpose of simplifying theorems and algorithms. Most results in Sections 3.3 and 5 can be easily generalized towards situations *without* relatively complete recourse. Detailed discussions in different contexts are provided accordingly.

Definition 2 (Relatively Complete Recourse [29])**.** Two-stage RO problem (12) is said to have *relatively complete recourse*, if for any $x \in \mathcal{X}$ and $\delta \in \Delta$, the second stage problem $\min_{y \in \mathcal{Y}(x;\delta)} d^\intercal y$ is feasible, i.e., $\mathcal{Y}(x;\delta) \ne \emptyset$.

One critical observation is that (12) can be formulated as a *single-stage* robust problem[5]. The following proposition lays the foundation of the main theoretical results (Theorems 1, 2 and 3) of this paper.

Proposition 1. *The two-stage robust optimization problem* (12) *is equivalent to the single-stage problem* (14)*:*

$$\min_{x \in \mathcal{X}, \gamma} \ c^\intercal x + \gamma \tag{14a}$$

$$s.t. \ (x, \gamma) \in \mathcal{Z}(\delta), \ \forall \delta \in \Delta. \tag{14b}$$

in which $\mathcal{Z}(\delta) := \{(x, \gamma) : \exists y \in \mathcal{Y}(x;\delta) \text{ and } d^\intercal y \le \gamma\}$. *We say* (12) *is equivalent with* (14) *in the sense that they share the same optimal first-stage solutions and optimal objective values.*

---

[5]This observation may not be true for adjustable robust optimization problems, see Chapter 14 of [25] and Section 6 of [30] for in-depth discussions.

The seemingly simple formulation (14) may not be solved directly, the main reason is that set $\mathcal{Z}(\boldsymbol{\delta})$ could be complicated (e.g., the intersection of an exponential number of half-spaces), sometimes an analytical form of $\mathcal{Z}(\boldsymbol{\delta})$ may not even exist. The only known property of $\mathcal{Z}(\boldsymbol{\delta})$ is its convexity if the original problem (12) is convex[6]. Proposition 1 only aims at connecting two-stage RO problems with the single-stage scenario approach theory in Section 3.3.

In the remainder of this paper, we make the choice that the uncertainty set Δ is a collection of K i.i.d. realizations of random variables $\boldsymbol{\delta}$, i.e., $\Delta = \mathcal{K} := \{\boldsymbol{\delta}^{(1)}, \boldsymbol{\delta}^{(2)}, \cdots, \boldsymbol{\delta}^{(K)}\}$. In Section 3.3, we show that this simple construction of uncertainty sets possesses rigorous theoretical guarantees. In addition, uncertainty set $\Delta = \mathcal{K}$ does not introduce any additional computational complexity (unlike ellipsoidal or conic uncertainty sets in [24], which sometimes render solving RO problems intractable), and the resulting two-stage RO problem can be efficiently solved by C&CG algorithm.

3.2. Column-and-Constraint Generation (C&CG) Algorithm

One popular choice to solve the two-stage RO problem is the column-and-constraint generation (C&CG) algorithm [29]. When the uncertainty set is a collection of K discrete scenarios $\mathcal{K} := \{\boldsymbol{\delta}^{(1)}, \boldsymbol{\delta}^{(2)}, \cdots, \boldsymbol{\delta}^{(K)}\}$,

$$\min_{\boldsymbol{x} \in \mathcal{X}} \left(\boldsymbol{c}^{\mathsf{T}} \boldsymbol{x} + \max_{\boldsymbol{\delta} \in \mathcal{K}} \min_{\boldsymbol{y} \in \mathcal{Y}(\boldsymbol{x}; \boldsymbol{\delta})} \boldsymbol{d}^{\mathsf{T}} \boldsymbol{y} \right) \quad (15)$$

Proposition 1 shows that (15) is equivalent with

$$\min_{\boldsymbol{x}, \gamma} \quad \boldsymbol{c}^{\mathsf{T}} \boldsymbol{x} + \gamma \quad (16a)$$

$$\text{s.t. } \boldsymbol{x} \in \mathcal{X} \text{ and } (\boldsymbol{x}, \gamma) \in \cap_{i=1}^{K} \mathcal{Z}(\boldsymbol{\delta}^{(i)}). \quad (16b)$$

Constraint $(\boldsymbol{x}, \gamma) \in \cap_{i=1}^{K} \mathcal{Z}(\boldsymbol{\delta}^{(i)})$ can be explicitly written as (17) by introducing additional variables $\{\boldsymbol{y}^{(1)}, \cdots, \boldsymbol{y}^{(K)}\}$. Variable $\boldsymbol{y}^{(k)}$ is the recourse decision variable for the kth scenario $\boldsymbol{\delta}^{(k)}$.

$$\boldsymbol{d}^{\mathsf{T}} \boldsymbol{y}^{(k)} \leq \gamma, \ k = 1, 2, \cdots, K. \quad (17a)$$

$$\boldsymbol{y}^{(k)} \in \mathcal{Y}(\boldsymbol{x}, \boldsymbol{\delta}^{(k)}), \ k = 1, 2, \cdots, K. \quad (17b)$$

Therefore, the two-stage RO problem (12) is equivalent to a single-stage optimization problem. It is worth pointing out that constraint (17) is simply enumerating *all* K scenarios. When seeking risk-averse solutions, which is common for power system applications, the number of scenarios K could be colossal. Thus (17) might consist of a gigantic number of decision variables and constraints, which is extremely inefficient or even impossible to solve. A *partial* enumeration such as the C&CG algorithm could significantly outperform the approach of solving (17) in one shot.

The intuition behind the C&CG algorithm is quite simple: only a small portion O of the uncertainty set Δ matters, e.g., extreme points of Δ along the optimization direction. C&CG algorithm is essentially an iterative procedure to identify critical scenarios O. The C&CG algorithm iteratively adds constraints $(\boldsymbol{x}, \gamma) \in \cap_{i \in O} \mathcal{Z}(\boldsymbol{\delta}^{(l)})$ to the problem (*constraint generation*). Since constraint $(\boldsymbol{x}, \gamma) \in \mathcal{Z}(\boldsymbol{\delta}^{(l)})$ guarantees the existence of feasible recourse variable $\boldsymbol{y}^{(l)} \in \mathcal{Y}(\boldsymbol{x}; \boldsymbol{\delta})$, additional second-stage variables are introduced (*column generation*). Algorithm 2 in Appendix B.2 formally defines the C&CG algorithm. More details and theoretical analysis on C&CG can be found in Appendix B.2.

3.3. Probabilistic Guarantees via the Scenario Approach

Throughout this paper, we construct the uncertainty set Δ using K i.i.d. scenarios $\Delta = \mathcal{K} := \{\boldsymbol{\delta}^{(1)}, \boldsymbol{\delta}^{(2)}, \cdots, \boldsymbol{\delta}^{(K)}\}$. The resulting optimization problems are presented in (16) and (17). Let $(\boldsymbol{x}_{\mathcal{K}}^*, \gamma_{\mathcal{K}}^*)$ denote the optimal solution to (16), e.g., returned by the C&CG algorithm. The main results of the scenario approach theory connects the number of scenarios K with the violation probability of a candidate solution.

Definition 3 (Violation Probability)**.** The violation probability of a candidate solution $(\boldsymbol{x}^\diamond, \gamma^\diamond)$ to (16) is defined as $\mathbb{V}(\boldsymbol{x}^\diamond, \gamma^\diamond) := \mathbb{P}_{\boldsymbol{\delta}}((\boldsymbol{x}^\diamond, \gamma^\diamond) \notin \mathcal{Z}(\boldsymbol{\delta}))$.

Remark 2. Mathmatically speaking, the violation probability $\mathbb{V}(\boldsymbol{x}_{\mathcal{K}}^*, \gamma_{\mathcal{K}}^*)$ quantifies the quality of the robust solution $(\boldsymbol{x}_{\mathcal{K}}^*, \gamma_{\mathcal{K}}^*)$. Specifically, $\mathbb{V}(\boldsymbol{x}_{\mathcal{K}}^*, \gamma_{\mathcal{K}}^*)$ is the probability of the following two events happening:

1. $\mathcal{Y}(\boldsymbol{x}_{\mathcal{K}}^*; \boldsymbol{\delta})$ is empty (infeasible second-stage problem);

2. there exists a feasible $\boldsymbol{y} \in \mathcal{Y}(\boldsymbol{x}_{\mathcal{K}}^*; \boldsymbol{\delta})$ but $\boldsymbol{d}^{\mathsf{T}} \boldsymbol{y} > \gamma_{\mathcal{K}}^*$.

In the context of power system planning, the two events above represent two potential risks in operation, and violation probability $\mathbb{V}(\boldsymbol{x}_{\mathcal{K}}^*, \gamma_{\mathcal{K}}^*)$ has clear physical interpretations, see Section 5.1.1 for more discussions.

The main theorems of the scenario approach theory are based on the key definitions of invariant set and essential set[7].

Definition 4 (Invariant Set $\mathcal{I}$ and Essential Set $\mathcal{E}$ [23])**.** Let $(\boldsymbol{x}_{\mathcal{I}}^*, \gamma_{\mathcal{I}}^*)$ denote the optimal solution to (16) using a subset of scenarios $\mathcal{I} \subseteq \mathcal{K}$. Set $\mathcal{I}$ is an Invariant Set if $\boldsymbol{c}^{\mathsf{T}} \boldsymbol{x}_{\mathcal{I}}^* + \gamma_{\mathcal{I}}^* = \boldsymbol{c}^{\mathsf{T}} \boldsymbol{x}_{\mathcal{K}}^* + \gamma_{\mathcal{K}}^*$. An *essential set* $\mathcal{E}$ is an invariant set with minimum cardinality.

The original scenario approach theory [15, 17, 18, 23] only applies to single-stage optimization problems. There is a lack of known results of multi-stage scenario approach. Proposition 1 first shows that the two-stage robust optimization problem (12) can be converted to an equivalent single-stage formulation (14), which enables us to extend the classical scenario

---

[6]Essentially $\mathcal{Z}(\boldsymbol{\delta})$ is obtained by (1) lifting the original feasible region in $\mathbb{R}^{n_x} \times \mathbb{R}^{n_y}$ by introducing γ then (2) projecting the feasible region onto $\mathbb{R}^{n_x}$. If the original feasible region is convex, then the affine projection of a convex set remains convex.

[7]For the convex case in Theorem 1, [23] shows that the essential set is the set of support scenarios as in [15]. For non-convex case, an essential set is the minimal support-subsample in [18].

approach theory towards two-stage decision making problems. Theorems 1, 2 and 3, which provide guarantees on the risk of the robust solution $(\boldsymbol{x}_{\mathcal{K}}^*, \gamma_{\mathcal{K}}^*)$, are essentially applying the key theorems of the classical scenario approach to the converted single-stage scenario problem[8]. The relatively completely recourse assumption is necessary to meet the feasibility assumption in [15, 17, 18, 23] (see Assumption 2 in Subsection Appendix A.1).

Theorem 1 (Prior Guarantees [15]). *Suppose* (15) *is convex and has relatively complete recourse. Given an acceptable risk level* $\epsilon \in (0, 1)$, *a confidence parameter* $\beta \in (0, 1)$, *and the number of first-stage decision variables d of* (16). *Let K be the smallest integer such that*

$$\sum_{i=0}^{d-1} \binom{K}{i} \epsilon^i (1 - \epsilon)^{K-i} \le \beta, \tag{18}$$

then $\mathbb{P}^K(\mathbb{V}(\boldsymbol{x}_{\mathcal{K}}^*, \gamma^*) > \epsilon) \le \beta$, *where* $(\boldsymbol{x}^*, \gamma^*)$ *is the optimal solution to* (16) *with K i.i.d. scenarios.*

Theorem 2 (Posterior Guarantees for Convex Problems [17]). *Suppose* (15) *is convex and has relatively complete recourse. Given a confidence parameter* $\beta \in (0, 1)$ *and* $(\boldsymbol{x}_{\mathcal{K}}^*, \gamma^*)$ *is the optimal solution to* (16). *Let* $|\mathcal{I}|$ *be the cardinality of an invariant set, then the following probabilistic guarantee holds:*

$$\mathbb{P}^K\big(\mathbb{V}(\boldsymbol{x}_{\mathcal{K}}^*, \gamma^*) \ge \epsilon(|\mathcal{I}|)\big) \le \beta, \tag{19}$$

where $0 < \epsilon(k) < 1$ *is the (unique) solution to the polynomial equation given an integer* $k = 0, 1, \cdots, K$,

$$\frac{\beta}{K+1} \sum_{i=k}^{K} \binom{i}{k} (1 - \epsilon)^{i-k} - \binom{K}{k} (1 - \epsilon)^{K-k} = 0 \tag{20}$$

Theorem 3 (Probabilistic Guarantees for Non-convex Problems [18]). *Suppose* (15) *has relatively complete recourse, and* $(\boldsymbol{x}_{\mathcal{K}}^*, \gamma^*)$ *is the optimal solution to* (16), *let* $\mathcal{I}_{\mathcal{K}}$ *be an invariant set of* (16) *using K scenarios, then the following probabilistic guarantee holds:*

$$\mathbb{P}^K(\mathbb{V}(\boldsymbol{x}_{\mathcal{K}}^*, \gamma^*) > \epsilon(|\mathcal{I}_{\mathcal{K}}|, \beta, K)) \le \beta, \ where \tag{21}$$

$$\epsilon(k, \beta, K) := \begin{cases} 1 & \text{if } k = K; \\ 1 - \left(\frac{\beta}{K\binom{K}{k}}\right)^{\frac{1}{K-k}} & \text{otherwise.} \end{cases} \tag{22}$$

Although Theorems 1, 2 and 3 hold for *any* invariant set, the tightest guarantee is achieved with the essential set, i.e., the invariant set with minimal cardinality [23]. In general, finding essential sets is a combinatorial problem, which could be computationally intractable for non-convex scenario problems (15). Proposition 5 shows that the C&CG algorithm can effectively narrow down the range of searching.

---

[8]Theorem 1 is essentially Theorem 1 of [15], Theorem 2 is from Theorem 2 of [17], and Theorem 3 is a direct corollary of Theorem 1 in [18].

Proposition 2. *The set O returned by the C&CG algorithm (Algorithm 2) is an invariant set.*

Proposition 5 is almost self-evident. The last step of C&CG is to solve a two-stage RO problem with all scenarios in O. [29] shows that C&CG algorithm converges to an optimal solution to (17). By definition, O is an invariant set (not necessarily an essential one). In practice, C&CG algorithm usually converges after only a few iterations so that the cardinality of O is small, then we can use Algorithm 3 to identify the essential set from O.

4. Robust Storage Planning

4.1. Compact Formulation

To determine the best location and size of energy storage systems, storage planning must account for short-term operational uncertainties, as the main benefits of energy storage are smoothing out the fluctuations of renewable generation and facilitating the integration of renewables. In the deterministic storage planning model (1), wind generation $w_{n,t} = \alpha_{n,t}^w \overline{w_n}$ and load $d_{n,t} = \alpha_{n,t}^d \overline{d_n}$ are considered as deterministic trajectories. The deterministic approach (1) fails to take the short-term operational risk into consideration. To account for the significant benefits of energy storage in reducing operation risk, we propose a two-stage robust storage planning model. Through constructing a scenario-based uncertainty set using wind and load data, we show that the operation risk is guaranteed to be within acceptable ranges.

The proposed two-stage robust storage planning framework is (23):

$$(\text{nc-RSP}): \ \min_{\boldsymbol{x} \in \Xi_1} C(\boldsymbol{x}) + K_d \max_{\delta \in \Delta} \min_{\boldsymbol{y} \in \Xi_2(\boldsymbol{x}; \delta)} c(\boldsymbol{y}). \tag{23}$$

The uncertainties are modeled by a pre-defined uncertainty set Δ. There is only one difference, i.e., $\max_{\delta \in \Delta}$, between the deterministic formulation (1) and its robust counterpart (23). By optimizing the decision for the worst-scenario, the solution to (23) is immune against all possible realizations of uncertainties in the uncertainty set Δ.

The choice of the uncertainty set Δ lies at the heart of robust optimization. Throughout this paper, we construct the scenario-based uncertainty set $\mathcal{K}$ to model short-term uncertainties from renewables and loads in RSP.

$$\Delta = \mathcal{K} := \{\boldsymbol{\delta}^{(k)}\}_{k=1}^{|\mathcal{K}|} = \left\{ \{\alpha_{n,t}^{d,(k)}\}_{n\in\mathcal{N},t\in\mathcal{T}}, \{\alpha_{n,t}^{w,(k)}\}_{n\in\mathcal{N},t\in\mathcal{T}} \right\}_{k=1}^{|\mathcal{K}|} \tag{24}$$

More specifically, the scenario-based uncertainty set $\mathcal{K}$ is the set of K i.i.d. scenarios. It consists of K daily profiles of load factors $\{\alpha_{n,t}^{d,(k)}\}_{n\in\mathcal{N},t\in\mathcal{T}}$ and renewable capacity factors $\{\alpha_{n,t}^{w,(k)}\}_{n\in\mathcal{N},t\in\mathcal{T}}$. These scenarios could come from historical data [38] or scenario generating algorithms [40]. One direct benefit of using scenarios is to capture spatial and temporal correlations of uncertainties. More discussions on the benefits of constructing the scenario-based uncertainty set for storage planning are in Section 5.

In addition, these operational scenarios could come from historical data [38] and also can come from data-driven scenario generation method [40]. One direct benefit of using scenarios is to capture spatial and temporal correlations of uncertainties. More discussions on the benefits of constructing the scenario-based uncertainty set for storage planning are in Section 5.

We mainly focus on the impacts of short-term uncertainties in this paper. The major long-term uncertainties considered are load growth and increasing penetrations of renewables, which are modeled by predicted peak loads $\overline{d_n}$ and renewable capacities $\overline{w_n}$. The values of peak loads and renewable capacities could come from human experts or state-of-the-art prediction algorithms. Furthermore, the proposed framework can be easily extended towards the joint planning of energy storage, (renewable) generation, transmission lines, and other critical infrastructures.

The proposed RSP model constructs the uncertainty set from HOD to represent the future short-term uncertainties. However, the uncertainty set only includes a small part of all possible renewable outputs and loads. It is uncertain that whether employing some scenarios from HOD is adequate to generate a robust planning result which can deal with all possible short-term uncertainties? The next section will provide a theoretical analysis for this problem.

4.2. Convex and Non-convex Formulations

Due to discrete variables $z_{n,t}$ and $v_{n,t}$, the robust storage planning problem (23) is non-convex. To reduce computational burden and obtain better theoretical results, we introduce the convexified version of (23) in (27). We first remove the discrete planning variable z_n, thus delete constraint (6) and make P_n and E_n planning decision variables.

$$\underline{\rho} E_n \le P_n \le \overline{\rho} E_n, \ n \in \mathcal{N}. \tag{25}$$

Next we relax the binary variable $v_{n,t}$ (charging/discharging) to continuous variable in (26).

$$0 \le v_{n,t} \le 1, \ n \in \mathcal{N}, \ t \in \mathcal{T}. \tag{26}$$

This relaxation is commonly adopted when studying storage system operations. Several sufficient conditions were derived to guarantee the exactness of this relaxation, and many simulation results reported that this relaxation was usually exact in practice, e.g., [41, 32].

After the modifications above, we denote the convexified feasible region as $\breve{\Xi}_1$.

$$\breve{\Xi}_1 := \left\{ (E_n, P_n, z_n) \text{s that satisfy constraints (6)(7)(25)} \right\}$$

The feasible region of the second stage $\breve{\Xi}_2(E_n, P_n; \boldsymbol{\delta})$ is

$$\breve{\Xi}_2(E_n, P_n; \boldsymbol{\delta}) := \left\{ (p_{n,t}^{\text{ch}}, p_{n,t}^{\text{dis}}, p_{i,n,t}^{\mathcal{G}}, p_{n,t}^{\text{shed}}) \text{s} \right.$$
$$\left. \text{that satisfy constraints (11)(9)(26)} \right\}.$$

The convexified robust storage planning problem is in (27). Notice that both $\breve{\Xi}_1$ and $\breve{\Xi}_2(E_n, P_n; \boldsymbol{\delta})$ are convex.

$$\text{(c-RSP)} : \min_{\boldsymbol{x} \in \breve{\Xi}_1} C(\boldsymbol{x}) + K_d \max_{\boldsymbol{\delta} \in \Delta} \min_{\boldsymbol{y} \in \breve{\Xi}_2(\boldsymbol{x}; \boldsymbol{\delta})} c(\boldsymbol{y}) \tag{27}$$

In fact, the convex formulation (27) is also widely adopted in storage planning studies, e.g., [42, 43]. The convex storage planning model also has two formulations: cost-minimizing with objective (2) and curtailment-minimizing with objective (4). In the remainder of this paper, we refer to (23) as non-convex robust storage planning (nc-RSP), and (27) as convex robust storage planning (c-RSP). Both (c-RSP) and (nc-RSP) will be solved using the C&CG algorithm introduced in Section 3.2. There are four different formulations being studied in this paper. Table 1 provides a detailed comparison of different formulations.

Table 1: Four Different Energy Storage Formulations

	cost-minimizing		curtailment-minimizing	
	(c-RSP)	(nc-RSP)	(c-RSP)	(nc-RSP)
Objective	(2)	(2)	(4)	(4)
1st stage	(6)(7)(25)	(5)(6)(7)	(6)(7)(25)	(5)(6)(7)
2nd stage	(11)(9)(26)	(11)(9)(10)	(11)(9)(26)	(11)(9)(10)

5. Theoretical Analysis

A careful choice of uncertainty set Δ is critical to get meaningful storage planning results. Throughout this paper, we adopt a data-driven approach to constructing uncertainty set $\Delta = \mathcal{K} = \{\boldsymbol{\delta}^{(k)}\}_{k=1}^{|\mathcal{K}|}$ using K i.i.d. scenarios. This simple yet powerful choice of uncertainty set originates from the scenario approach [15, 17, 18]. All theoretical results in this paper are based on two critical definitions: violation probability and invariant set; their mathematical definitions and interpretations in the context of RSP are discussed in Section 5.1. Sections 5.2-5.3 present the main theorems of this paper.

5.1. Basic Definitions

5.1.1. Violation Probability $\mathbb{V}(\boldsymbol{x}^\diamond, \gamma^\diamond)$

Definition 5 (Violation Probability). The violation probability of a candidate solution $(\boldsymbol{x}^\diamond, \gamma^\diamond)$ to (14) is defined as $\mathbb{V}(\boldsymbol{x}^\diamond, \gamma^\diamond) := \mathbb{P}_{\boldsymbol{\delta}}((\boldsymbol{x}^\diamond, \gamma^\diamond) \notin \mathcal{Z}(\boldsymbol{\delta}))$.

Similar with Proposition 1, we introduce an auxiliary variable γ to denote the worst-case cost. Let $(\boldsymbol{x}^*, \gamma^*)$ be the optimal (first-stage) solution to (c-RSP) or (nc-RSP). In the context of RSP, $\mathbb{V}(\boldsymbol{x}^*, \gamma^*)$ depicts the probability of the following two events happening in the future:

1. infeasible second-stage problem (infeasible DCOPF);

2. there exists a feasible operation, but its operation cost γ is greater than the planning solution γ^*.

Detailed interpretations of $\mathbb{V}(\boldsymbol{x}^*, \gamma^*)$ differ in the cost-minimizing and curtailment-minimizing formulations.

Remark 3 (Operation Cost Risk). The cost-minimizing formulation minimizes total investment and operation costs (2). The solution γ^* is our estimation of worst-case operation costs. As stated in Remark 5, when allowing load curtailment, the second stage is always feasible, thus $\mathbb{V}(\boldsymbol{x}^*, \gamma^*)$ only depicts event (2),

i.e., the actual worst-case operation cost $d^\mathsf{T} y(\delta; x^*)$ in the future is greater than γ^*:

$$\mathbb{V}(x^*, \gamma^*) = \mathbb{P}_\delta(d^\mathsf{T} y(\delta; x^*) > \gamma^*) \qquad (28)$$

$\mathbb{V}(x^*, \gamma^*)$ in the cost-minimizing formulation is referred as *operation cost risk*.

Remark 4 (Load Curtailment Risk). The curtailment-minimizing formulation minimizes total curtailment (4). The solution γ^* is our estimate on the worst-case load curtailment. $\mathbb{V}(x^*, \gamma^*)$ is the probability that load curtailment is greater than our worst-case estimate γ^*.

$$\mathbb{V}(x^*, \gamma^*) = \mathbb{P}_\delta\Big(\sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} p_{n,t}^{\text{shed}} > \gamma^*\Big) \qquad (29)$$

$\mathbb{V}(x^*, \gamma^*)$ in the curtailment-minimizing formulation is referred as *load curtailment risk*.

We would like to point out the close relationship between $\mathbb{V}(x^*, \gamma^*)$ and power system reliability. When $\gamma^* = 0$, $\mathbb{V}(x^*, \gamma^*)$ is the (daily) loss of load probability (LOLP). If the sole objective of storage planning is to guarantee the LOLP $\leq \bar{\epsilon}$, then we can add the following constraint to (c-RSP) or (nc-RSP).

$$\sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}} p_{n,t}^{\text{shed}} = 0. \qquad (30)$$

5.1.2. Invariant and Essential Sets

Definition 6 (Invariant Set $\mathcal{I}$ and Essential Set $\mathcal{E}$ [23]). Let $(x_\mathcal{I}^*, \gamma_\mathcal{I}^*)$ denote the optimal solution to (14) using a subset of scenarios $\mathcal{I} \subseteq \mathcal{K}$. Set $\mathcal{I}$ is an *invariant set* if $c^\mathsf{T} x_\mathcal{I}^* + \gamma_\mathcal{I}^* = c^\mathsf{T} x_\mathcal{K}^* + \gamma_\mathcal{K}^*$. An *essential set* $\mathcal{E}$ is an invariant set with minimum cardinality.

An invariant set is basically the set of important scenarios that determine the optimal solution. An essential set is essentially the minimal representation of those important scenarios. The number of those important scenarios play a critical role in the theoretical analysis.

Proposition 3. *The cardinality of any essential set of (c-RSP) is no greater than $2|\mathcal{S}| + 1$, where $\mathcal{S}$ denotes the set of candidate locations for storage planning.*

Since (c-RSP) is convex, the cardinality of the essential set is bounded by the total number of decision variables of both first and second stages (Theorem 2 in [44]). In most cases, the number of the second-stage variables is much more than the first-stage variables. For example, the 118-bus case in Section 6.2 has $2|\mathcal{S}| + 1 = 237$ first stage variables and $\sim 2 \times 10^4$ second-stage variables. If directly applying Theorem 2 in [44], we obtain a loose upper bound $|\mathcal{E}| \lesssim 2 \times 10^4$, which leads to an astronomical number of scenarios K per Theorem 4. Proposition 3 significantly tightens the bound as $|\mathcal{E}| \leq 237 \ll 2 \times 10^4$.

Another attractive feature of Proposition 3 is its independence of system size. As long as there are not too many candidate locations for storage planning, the proposed approach will not require too many scenarios when being applied on large-scale real-world systems.

5.1.3. Relatively Complete Recourse

Remark 5. When allowing load curtailment, both (c-RSP) and (nc-RSP) have relatively complete recourse.

Note that all theorems in Sections 5.2-5.3 requires the relatively complete recourse assumption. Remark 5 states that our RSP formulations satisfy this assumption. This remark follows common sense of power system operations. When a system is in severe situations, a common control action is to shed load (e.g., rotating outages). Remark 5 essentially states that system operators can shed load to a lower level to maintain minimum generation, thus the second stage problem (DCOPF) is feasible at the high cost of load curtailment. We also want to point out that Remark 5 does not hold true if the following two types of constraints are included: (1) no load curtailment is allowed, i.e., $p_{n,t}^{\text{shed}} = 0$; or (2) load curtailment is limited, i.e., $\sum_{n\in\mathcal{N}} p_{n,t}^{\text{shed}} \leq \overline{p}^{\text{shed}}$. All results in Sections 5.2-5.3 are for the cases with load curtailment, thus meets the relatively complete recourse assumption. Similar theoretical results *without* the relatively complete recourse assumption can be easily derived, e.g., using Theorem 4.1 in [16].

5.2. A-Priori Guarantees for (c-RSP)

Theorem 4 (A-Priori Guarantees for (c-RSP)). *For the convex RSP formulation (c-RSP) with $|\mathcal{S}|$ candidate storage locations, given an acceptable risk level $\bar{\epsilon} \in (0, 1)$ and a confidence parameter $\beta \in (0, 1)$, let (x^*, γ^*) be the optimal solution to (c-RSP) with K i.i.d. scenarios, and K be the smallest integer such that*

$$\sum_{i=0}^{2|\mathcal{S}|} \binom{N}{i} \bar{\epsilon}^i (1-\bar{\epsilon})^{N-i} \leq \beta, \qquad (31)$$

then $\mathbb{P}^K(\mathbb{V}(x_\mathcal{K}^*, \gamma^*) \leq \bar{\epsilon}) \geq 1 - \beta$.

In Section 6, after solving hundreds of instances of (c-RSP) on different systems, we surprisingly found that the essential set of every instance of (c-RSP) was always one, i.e., $|\mathcal{E}| = 1$. This feature of $|\mathcal{E}| = 1$ is very appealing, *as $|\mathcal{E}| = 1$ is the best non-trivial case in the sense that it requires the least amount of scenarios to achieve a given risk level $\bar{\epsilon}$.* One first attempt to explain the reason that $|\mathcal{I}| = 1$ for (c-RSP) is Proposition 4, which improves Proposition 3 for special cases.

Definition 7 (The Worst Scenario). The *worst scenario* of (12) is defined as $\delta^* := \arg\max_{\delta\in\Delta} f(\delta)$, where $f(\delta) := \min_{x\in\mathcal{X}, (x,\gamma)\in\mathcal{Z}(\delta)} c^\mathsf{T} x + \gamma$.

Proposition 4. *If the worst scenario δ^* of (15) is unique, then $\mathcal{I} = \Omega = \{\delta^*\}$, thus $|\mathcal{I}| = 1$.*

Proposition 4 can be applied on both convex and non-convex problems. However, Proposition 4 needs global optimal solutions, which is challenging for non-convex problems. Table 4 provides an example that suboptimal solution leads to more than one element in the essential set.

5.3. A-Posteriori Guarantees for (c-RSP) and (nc-RSP)

Although we cannot prove $|\mathcal{E}| = 1$ with less restricted assumptions, we can still improve the a-priori guarantees (Theorem 4) using the a-posteriori guarantees (Theorems 5 and 6). Theorem 5 extends the Theorem 2 in [17] towards two-stage robust optimization problems; and Theorem (6) is a direct corollary of Theorem 1 in [18].

Theorem 5 (A-Posteriori Guarantees for (c-RSP) [17, 31]). *Let $(\boldsymbol{x}_{\mathcal{K}}^*, \gamma^*)$ be the optimal solution to (c-RSP). Let $|\mathcal{I}_{\mathcal{K}}|$ be the cardinality of an invariant set and $\beta \in (0, 1)$ be a confidence parameter chosen beforehand, then the following probabilistic guarantee holds:*

$$\mathbb{P}^K\Big(\mathbb{V}(\boldsymbol{x}_{\mathcal{K}}^*, \gamma^*) \geq \epsilon(|\mathcal{I}_{\mathcal{K}}|, \beta, K)\Big) \leq \beta, \qquad (32)$$

where $0 < \epsilon(k, \beta, K) < 1$ is the (unique) solution to the polynomial equation given an integer $k = 0, 1, \cdots, K$,

$$\frac{\beta}{K+1} \sum_{i=k}^{K} \binom{i}{k}(1-\epsilon)^{i-k} - \binom{K}{k}(1-\epsilon)^{K-k} = 0. \qquad (33)$$

Theorem 6 (A-Posteriori Guarantees for (nc-RSP) [18, 31]). *Let $(\boldsymbol{x}_{\mathcal{K}}^*, \gamma^*)$ denote the optimal solution to (nc-RSP), let $\mathcal{I}_{\mathcal{K}}$ be an invariant set of (nc-RSP) using K scenarios, then the following probabilistic guarantee holds:*

$$\mathbb{P}^K(\mathbb{V}(\boldsymbol{x}_{\mathcal{K}}^*, \gamma^*) > \epsilon(|\mathcal{I}_{\mathcal{K}}|, \beta, K)) \leq \beta, \; where \qquad (34)$$

$$\epsilon(k, \beta, K) := \begin{cases} 1 & if \; k = K; \\ 1 - \left(\frac{\beta}{K\binom{K}{k}}\right)^{\frac{1}{K-k}} & otherwise. \end{cases} \qquad (35)$$

As the name suggests, a-priori guarantees (Theorem 4) hold true *before* solving the (c-RSP) problem. In contrast, a-posteriori guarantees (Theorems 5 and 6) become valid *after* obtaining the optimal solution $(\boldsymbol{x}_{\mathcal{K}}^*, \gamma^*)$ and calculating the cardinality of an invariant set $\mathcal{I}_{\mathcal{K}}$. Although Theorems 5 and 6 hold true for *any* invariant set, the tightest guarantee is achieved with the essential set, i.e., the invariant set with minimal cardinality [23]. In general, finding essential sets is a combinatorial problem, which could be computationally intractable for non-convex scenario problems (15). Proposition 5 shows that the C&CG algorithm can effectively narrow down the range of searching.

Proposition 5. *The set O returned by the C&CG algorithm (Algorithm 2) is an invariant set.*

Proposition 5 is almost self-evident. The last step of C&CG is to solve a two-stage RO problem with all scenarios in O. [29] shows that C&CG algorithm converges to an optimal solution to (17). By definition, O is an invariant set (not necessarily an essential one). In practice, C&CG algorithm usually converges after only a few iterations so that the cardinality of O is small, then we can use Algorithm 3 to identify the essential set from O.

The main procedures to solve RSP problems and calculate theoretical guarantees are summarized in Algorithm 1.

Algorithm 1 Calculating Theoretical Guarantees for RSP

1: Choose acceptable risk level $\bar{\epsilon}$ and confidence parameter β;
2: Guess the cardinality of invariant set k;
3: Compute the smallest integer K such that $\epsilon(k, \beta, K) \leq \bar{\epsilon}$; *use (33) in Theorem 5 for (c-RSP); use (35) in Theorem 6 for (nc-RSP)*;
4: Construct the scenario-based uncertainty set $\mathcal{K} \leftarrow \{\boldsymbol{\delta}^{(k)}\}_{k=1}^{K}$ using K i.i.d. scenarios;
5: Solve the RSP problem with the uncertainty set $\mathcal{K}$ using the C&CG algorithm, obtain the optimal storage planning solution $(\boldsymbol{x}^*, \gamma^*)$ and an invariant set O;
6: Compute an invariant set $\mathcal{I}_{\mathcal{K}} \subseteq O$ as small as possible using an algorithm 3 ;
7: Calculate $\epsilon(|\mathcal{I}_{\mathcal{K}}|, \beta, K)$; *use (33) in Theorem 5 for (c-RSP); use (35) in Theorem 6 for (nc-RSP)*;
8: **if** $\epsilon(|\mathcal{I}_{\mathcal{K}}|, \beta, K) > \bar{\epsilon}$ **then**
9:     Go to step 2 with $k \leftarrow |\mathcal{I}_{\mathcal{K}}|$;
10: **else**
11:     Output uncertainty set $\mathcal{K}$, invariant set $\mathcal{I}_{\mathcal{K}}$, optimal solution $(\boldsymbol{x}^*, \gamma^*)$, and theoretical guarantee $\mathbb{P}^K(\mathbb{V}(\boldsymbol{x}^*, \gamma^*) \leq \bar{\epsilon}) \geq 1 - \beta$.
12: **end if**

Proposition 5 states that C&CG algorithm can identify an invariant set O while solving a two-stage RO problem. After solving many instances of (c-RSP) and (nc-RSP) problems, we observed that the C&CG algorithm usually converged within very few iterations, thus the calculated invariant sets O often have small cardinalities. The next step is to apply a greedy algorithm (Algorithm 3 in Appendix B.2), checking if the removal any scenario from O changes optimal solution) to further pinpoint an essential set. Since O only consists of a few scenarios, Algorithm 3 will stop after a few iterations. This indicates that it is computationally inexpensive to identify an essential set.

5.4. Reduce Conservativeness and Find Better Solutions

The procedures described in Sections 5.2 and 5.3 are essentially a randomized algorithm. More specifically, solutions to (c-RSP) or (nc-RSP) with different sets of scenarios of the same size K could be different, but all of them possess the same theoretical guarantees. We can exploit the randomness of the algorithm to get better storage planning decisions and reduce conservativeness. For example, in Section 6, we solved (c-RSP) or (nc-RSP) 10 times using 10 different datasets of same size, then chose the storage planning solutions with the least investment cost. This approach effectively avoids overly conservative solutions and achieves a trade-off between investment cost and risk in the robust planning model.

6. Case Study

Two case studies are presented in this section. The first one investigates the performance of the proposed approach in an environment with *abundant data*. Uncertainties are depicted using probability distributions, from which we can sample as
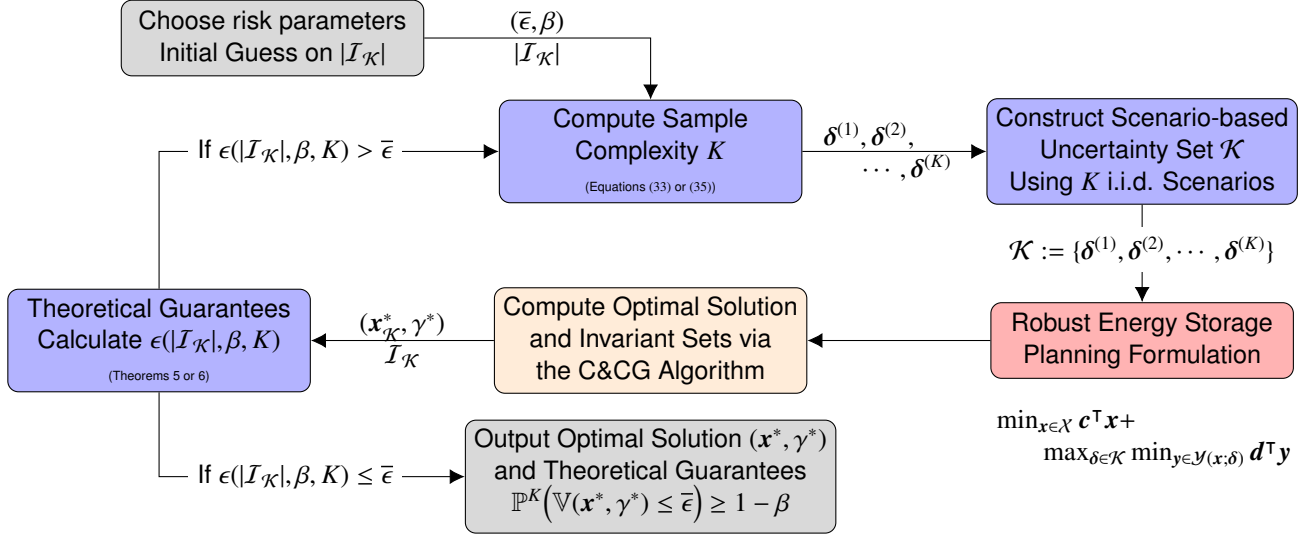
Figure 2: Main Procedures (Algorithm 1) to solve RSP problems and Calculate Theoretical Guarantees.

many scenarios as possible. The second case study is a thorough investigation of the proposed approach in a more realistic setting, i.e., a large-scale system with a limited number of available historical scenarios. Both case studies share the same settings of energy storage systems, critical parameters are summarized in Table 2.

Since the scenario approach is a *randomized* algorithm [15], it is necessary to quantify the randomness of the solution returned by the scenario approach. Specifically, for a given sample complexity K, we repetitively solved RSP using 10 independent sets of K scenarios. Each one of those 10 problems solved is referred as *an experiment* in this section.

6.1. IEEE 6-bus system

6.1.1. System configuration

The first case study is based on a modified IEEE 6-bus system [45]. We added two 100MW wind farms at buses 5 and 6, and increased the cost coefficients of coal and natural gas plants 20% higher than the original case. Key parameters of storage investment are in Table 2.

The uncertainties of load and wind are modeled by probability distributions. We first generated wind speed data from Weibull distributions[9]. The output of wind farms was computed based on the generated wind speed data using equation (44) in [38]. The load scenarios were generated by adding generated forecast errors on 365 days of historic load profiles in [38]. We assume the load forecast errors follow a Gaussian distribution $\mathcal{N}(0, 1\%)$.

6.1.2. Planning Results with Different Risk Parameters ϵ

In this case, we study the impact of sample complexity K on the robust planning solution considering the randomness of the algorithm. We examine the costs and variances of 10 solutions

of 10 experiments at the same sample complexity, i.e. the same risk level. There are 6 candidate locations in the test system, i.e. $|\mathcal{S}| = 6$. When adopting the prior risk guarantees, i.e. Theorem 1, given different risk parameters ϵ, we computed the sample complexity K with the same confidence parameter $\beta = 10^{-3}$ in the Corollary **??** . Results are in Table 3. Given K i.i.d. scenarios, we solved (c-RSP) and obtained optimal storage planning decisions.
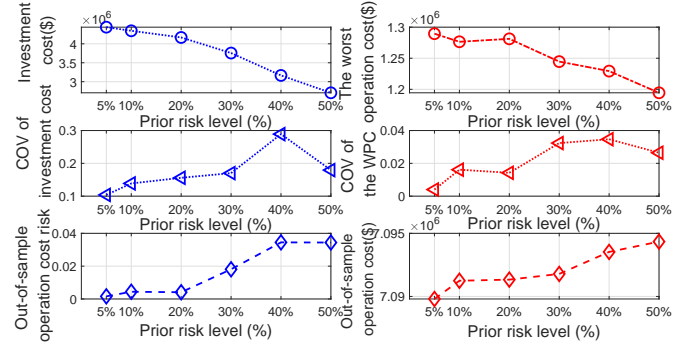

Figure 3: Storage Planning Results with Different A-priori Risk Guarantees ϵ. All results were computed based on 10 independent experiments (Section 6.1.2).

The first row of Fig. 3 shows two components in the objective of (c-RSP): investment cost $C(x)$ and worst-case operational cost $c(y)$. Clearly planning decisions of larger risk parameters ϵ require less planning investments. Although the top-right panel of Fig. 3 also shows worst-case operational cost $c(y)$ decreases with increasing risk parameters ϵ, we would like to emphasize that the worst-case operational cost $c(y)$ of (c-RSP) was evaluated on the sampled K scenarios, which *does not represent true operational costs*.

The true operational cost was estimated using another independent 1.6×10^4 test scenarios, it is shown as the "out-of-sample" operation cost $c(y)$ in the bottom-right panel of Fig. 3. Clearly, the true operation cost increases with risk parameters ϵ,

---

[9]Key parameters: scale factor 11.0086m/s, shape factor 1.9622m/s, cut-in speed V_{ci} = 4m/s, rated speed V_{rated} = 13.61m/s, V_{co} = 25m/s.

Table 2: Critical Parameters of Energy Storage Systems

Parameter	C_n^{P}	C_n^{E}	Γ	σ	η_n^{ch}	η_n^{dis}	$\overline{\rho}$	$\underline{\rho}$
Value	500\$/kW	20\$/kWh	10 Years	10%	0.9	0.9	0.8	0.2
Parameter	K_d	c_n^{dis}	c_n^{ch}	q^{E}	q^{P}	$\overline{z}$	$\overline{z}_n$	
Value	365	18\$/MW	1\$/MW	32MWh	8MW	20	4	

Table 3: The sample complexity for the different a-priori risk guarantees, using Theorem

A-priori risk level ϵ	5%	10%	20%	30%	40%	50%
Sample complexity K	796	398	199	133	100	80
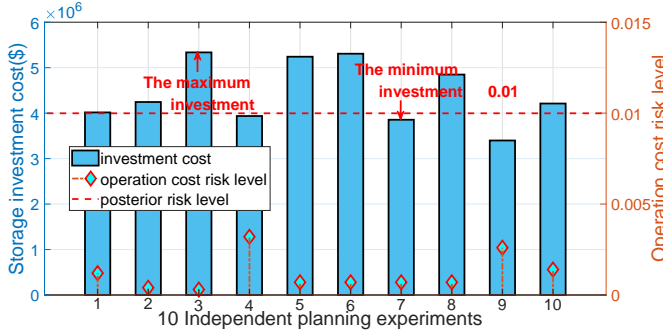
Figure 4: Storage planning results with adequate investment budget ($C(\boldsymbol{x}^*) < C^{\mathrm{budget}}$)

Figure 5: Investment cost and operation costs with adequate investment budget ($C(\boldsymbol{x}^*) < C^{\mathrm{budget}}$)
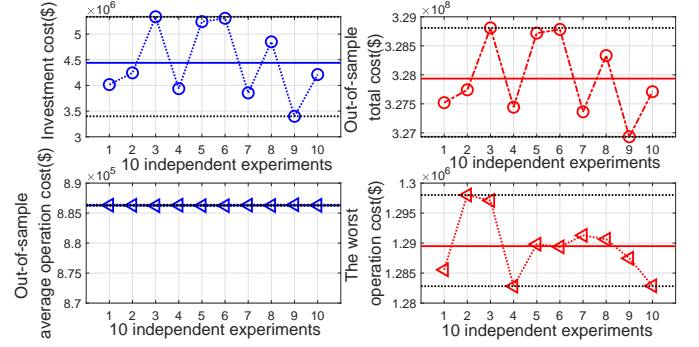
since more load curtailment will arise. The optimal solution γ^* of (c-RSP) provides an estimate of worst-case operation cost. The bottom-left figure of Fig. 3 shows $\mathbb{V}(\boldsymbol{x}^*, \gamma^*)$, i.e., the probability that the true operation cost is higher than γ^* (see Remark 2 and Section 5.1.1).

The second row of Fig. 3 examines the randomness in the planning decisions. The main metric being used here is the coefficient of variance (COV), which is defined by the ratio of standard deviation of a random variable to its mean. Fig. 3 illustrates the positive correlation between ϵ and COV values, more risk-averse (smaller ϵ) solutions are more stable (smaller COV). "WPC" means the worst-case operation cost.

6.1.3. Planning Results with Different Investment Budgets C^{budget}

Results of 10 experiments (c-RSP) with the the same risk guarantee parameters ($\epsilon = 0.01$, $\beta = 10^{-3}$, $K = 920$) but varying budgets C^{budget} are presented below. We take a closer look at these two interesting cases in the following subsections. The case studies adopt the posterior risk guarantee for *c-RSP* and correspond to Theorem 2 and the computation procedures in Section **??**. For each result of planning experiment, we compute its out-of-sample risk level through sampling another $1.6 * 10^4$ scenarios from the uncertainty distribution as test data.

Adequate Investment Budget ($C(\boldsymbol{x}^*) < C^{budget}$). The purpose of robust planning is to avoid future risk by optimizing the system in the worst case. When the investment budget is adequate, i.e., inactive budget constraint $C(\boldsymbol{x}^*) < C^{\mathrm{budget}} = \$6 \times 10^6$, the stor-

age installment will *eliminate* load curtailment in all scenarios in the uncertainty set $\mathcal{K}$. Planning results are reported in Fig. 4 show the storage investment costs and out-of-sample operation cost risk levels of all planning experiments. The cardinality of Invariant Set of all planning experiments are found to be 1. Since the out-of-sample violation probabilities (risk levels) are all below the risk parameter $\epsilon = 0.01$ (red dotted line), we can choose the planning result with the lowest cost. This could effectively reduce the randomness and conservativeness of the scenario approach.

Despite the fact that the storage planning decisions of 10 experiments have different investment costs, they would not lead to the different actual operation cost. Actual operation cost is represented by the *average* value of the operation costs of $1.6 * 10^4$ out-of-sample test scenarios. The *average* operation cost are almost identical, as shown in the bottom-left panel of Fig. 5. Hence, in the first row of Fig. 5, the investment costs and the total out-of-sample costs of the 10 experiments have the nearly same variation trend and amplitude. However, the bottom-right panel of Fig. 5 shows that the worst-case operation cost in (c-RSP) varies due to drawing extreme scenarios. It is worth noting that the worst-case operation cost, instead of the *average* operation cost, is embedded in the robust storage planning model. It is a good choice to avoid risks, but it cannot reflect the actual operation costs. Our method can trade-off the cost and risk in the robust planning model.

Inadequate Investment Budget $C(\boldsymbol{x}^*) = C^{budget}$. When the budget is inadequate, the budget constraint is active $C(\boldsymbol{x}^*) = C^{\mathrm{budget}}$, load curtailment are necessary in extreme scenarios. To ensure the load curtailment risk of the planning solution, we resort to the curtailment-minimizing formulation. In the first row of Fig. 6 , the investment budget is gradually increased and the load curtailment is decreased to zero. According to the

above theoretical analysis, the investment solution at 2.4×10^6\$ (0 load curtailment) can provide the needed load curtailment risk guarantee 0.01. Even though, at the investment 1.8×10^6\$, the out-of-sample load curtailment risk requirement seems to be satisfied. Considering the randomized property of the planning solutions, we adopt the planning solution in the investment of 2.4×10^6\$. Besides, all operation cost risks have been satisfied shown in the second row.
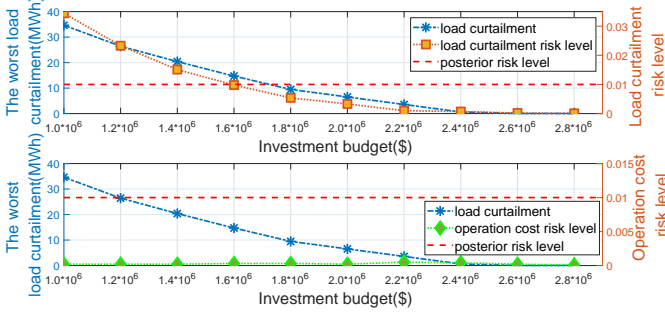

Figure 6: Storage planning with gradual increasing investments.

6.2. IEEE 118-bus system

6.2.1. System Configuration

We conducted numerical simulations on a modified 118-bus system[10] with five 200MW wind farms added to buses 16, 37, 48, 75 and 83. All 118 buses are candidate sites for storage installation. The wind and load profiles are from the Electric Reliability Council of Texas (ERCOT)[11]. The dataset consists of about 7300 days (20 years) of hourly wind and load profiles. It is worth noting that we converted the wind generation profiles to wind capacity factor $\{\alpha_{n,t}^{w,(k)}\}_{n \in \mathcal{N}, t \in \mathcal{T}}$ (0 ~ 100% of full capacities), which depict the spatial and temporal patterns of wind speeds. Load profiles were converted to load factors $\{\alpha_{n,t}^{d,(k)}\}_{n \in \mathcal{N}, t \in \mathcal{T}}$ (0 ~ 100%), which mainly model the temporal variations of system demands. The actual load and wind generation were scaled up according to the predicted peak loads and wind capacities. Wind capacity factors and load factors of 10 consecutive days are plotted in Figure 7.

6.2.2. Storage Planning Results of Four Different Formulations

Similar to the 6-bus case in Section 6.1, 10 experiments on the 118-bus system were conducted. In each experiment, the 7300-day dataset was divided into two non-overlapping subsets. The first subset contained the K scenarios as input to the scenario problem. The second subset consisted of all remaining scenarios, which served as the test dataset to evaluate out-of-sample results. If the remaining scenarios have more than 5000 scenarios, the test dataset includes 5000 scenarios; otherwise it

---

[10]The original system is available at `http://motor.ece.iit.edu/data`. Additional changes include generation costs are set 20% higher than the original case. The hourly ramp rates of generators are set to 45% of the maximum generation for the largest units in the system.

[11]Historical load data of ERCOT is from `http://www.ercot.com/mktinfo/loadprofile/alp`. Wind data is from `http://www.ercot.com/gridinfo/resource`.
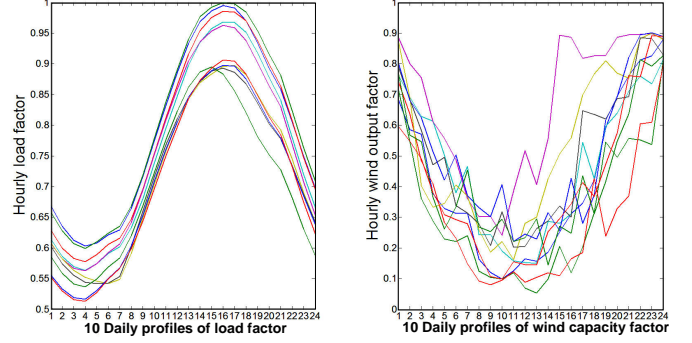

Figure 7: 10 Daily profiles of load factor and wind capacity factor in ECORT historical data

only includes 4000 scenarios. Storage planning results using four different formulations (see Table 1) are reported in Figure 8. In all experiments, we used the same parameters $\epsilon = 0.01$ and $\beta = 0.001$. For non-convex formulations (nc-RSP), we used different parameters $\overline{z}_n = 10$ and $\overline{z} = 1000$ from Table 2. We follow the procedures in Algorithm 1 to solve (c-RSP) and (nc-RSP) problems and obtain theoretical guarantees.

Using Cost-minimizing Formulation. Fig. 8a shows the storage planning results with adequate investment budget $C(\boldsymbol{x}) < C^{\text{budget}} = \$12 \times 10^7$. No experiment reached the investment budget C^{budget} and all experiments successfully maintained out-of-sample risk $\hat{\epsilon}$ within acceptable ranges ($\hat{\epsilon} < 0.01$). The largest out-of-sample risk level $\hat{\epsilon} = 0.0068 \leq 0.01$ happened in the 10th planning experiment. This verifies the probabilistic guarantees in Theorem 2.

Figure 8b is the case that employs the non-convex storage planning model (nc-RSP), in which the capacities of energy storages are quantized. The y-axis in Figure 8b is the total number of energy storage units to be installed. We started by assuming $|\mathcal{I}| = 1$ based on our observations after solving hundreds of (c-RSP) and (nc-RSP), and computed $K = 2250$ using (34) to guarantee $\epsilon = 0.01$ and $\beta = 0.001$. The out-of-sample violation probabilities in all 10 experiments are smaller than $\epsilon = 0.01$, this verifies Theorem 3. Notice that the smallest total investment in storage units happened at the 9th experiment, in which the largest out-of-sample $\hat{\epsilon} = 0.0078 < 0.01$ is within acceptable ranges. Notice that the 1st and 10th experiments have the same total storage units. This does not indicate the same storage planning solution, since they can be installed in different locations.

Using Curtailment-minimizing Formulation. Fig. 8c shows the convex planning result pursuing the posterior load-curtailment risk guarantee by the curtailment-minimizing formulation. To better study the impact of investment on load curtailment, in this case, the peak load level is 1.2 times of that in the above cost-minimizing planning experiment. With the increase of the investment budget, at the investment of 6×10^7\$, the out-of-sample risk level $\hat{\epsilon}$ is approaching the load curtailment risk requirement $\epsilon = 0.01$. But considering the randomized property of the planning solution, the planning solution at the invest-
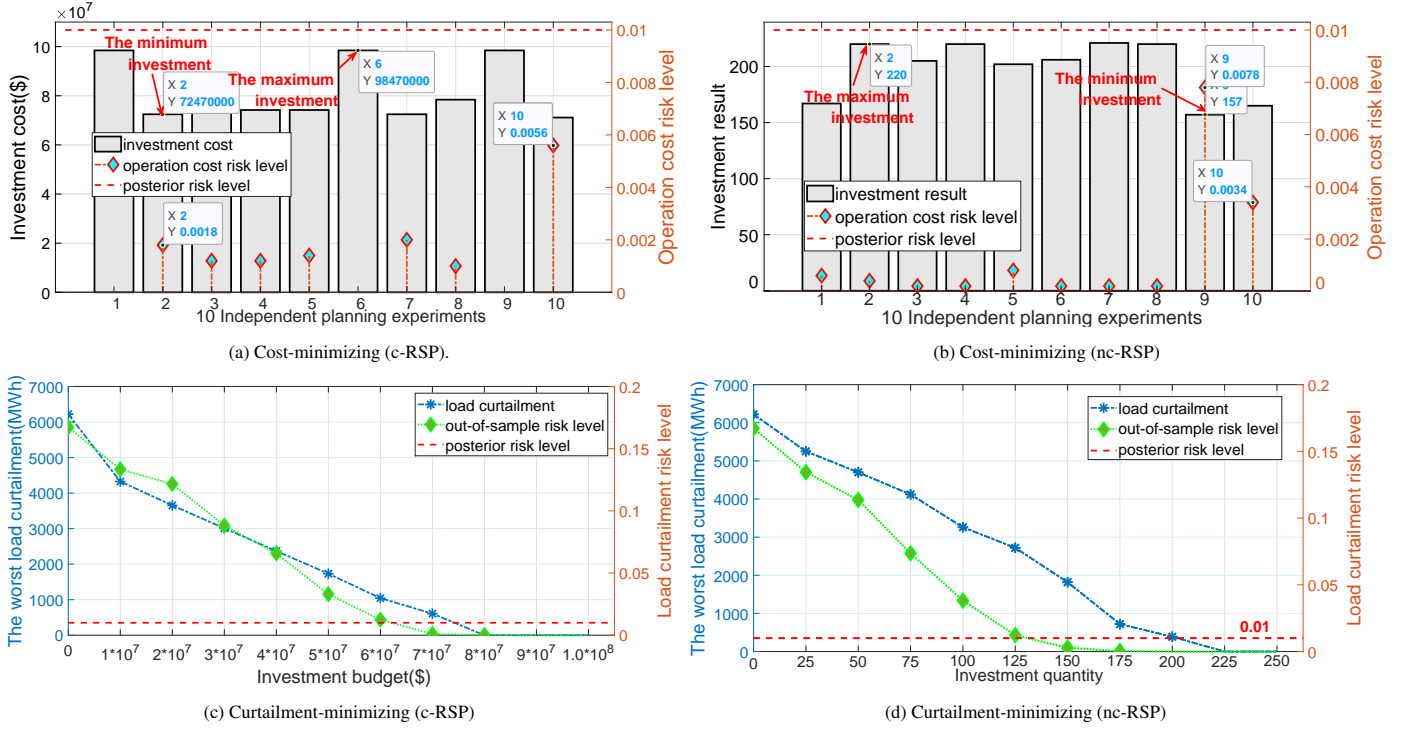
Figure 8: Storage planning Results of the IEEE 118-bus system, using four different formulations

ment of 8×10^7\$ should be employed because of the 0 load curtailment is reached. The blue star is the objective function of the curtailment-minimizing formulation, i.e. the maximum load curtailment. The 0-curtailment means the risk requirement $\epsilon = 0.01$ is ensured according to the above theoretical analysis. The out-of-sample risk level (green line) based on historical data also verified the risk guarantee.

When the investment quantity is limited, the load-curtailment risk is important. Fig. 8c shows the non-convex planning result pursuing the posterior load-curtailment risk guarantee by the curtailment-minimizing formulation. This experiment corresponds to the above curtailment-minimizing convex planning experiment, their peak load levels and other system parameters are the same. With the increase of the investment quantity, at the investment of 150 units, the out-of-sample risk level $\hat{\epsilon}$ is below the load curtailment risk requirement ϵ 0.01. The theoretical guarantee is hold at 225 units. The blue star indicates that the objective of the curtailment-minimizing model reaches 0 at 225 units. When we compared the non-convex and convex planning results by converting the quantity to investment cost, it is easy to find that the non-convex planning model needs more cost to reach the needed load curtailment risk guarantee. That is because the investment of quantized storage units have more physical limitations than the continuous storage units.

The out-of-sample experiments of these above simulations illustrate the robust planning result has the probability guarantees for short-term uncertainties. Hence, we can employ a part of HOD to represent short-term uncertainties in the robust planning model.

6.2.3. Convex Formulation versus Non-convex Formulation

Although this is a much bigger system than the 6-bus case, the cardinality of essential sets of (c-RSP) is always $|\mathcal{E}| = 1$. With the same parameters $\epsilon = 0.01$ and $\beta = 10^{-3}$, the 118-bus system requires the same number of scenarios $K = 920$ as the 6-bus system. Besides the 20 experiments reported in Figures 8a and 8c, we solved many additional (c-RSP) problems using different number of scenarios. We never found an exception of (c-RSP) with $|\mathcal{E}| > 1$. For non-convex problems, the observation $|\mathcal{E}| \le 1$ is not always true. Most of (nc-RSP) problems have $|\mathcal{E}| \le 2$, with a few exceptions in which $|\mathcal{E}| = 2$.

6.2.4. Obtaining Theoretical Guarantees

We illustrate the process of searching for essential set $\mathcal{E}$ and obtaining theoretical guarantees using one instance of (nc-RSP). Based on our experiences in Section 6.2.3, we first assume $|\mathcal{E}| = 1$. Given $\epsilon = 0.01$ and $\beta = 0.001$, we computed $K = 2220$ using equation (34). We first solved (nc-RSP) with 2220 scenarios via C&CG algorithm. Besides the optimal storage planning solution and optimal objective value 2.4253×10^9\$, C&CG algorithm also returned an invariant set $\mathcal{I} = \{589, 164, 1732\}$ (Proposition 5). Algorithm 3 started with $\mathcal{I} = \{589, 164, 1732\}$ and removes scenarios one by one to check if the invariant set can be future reduced. Table 4 illustrates this process. The first row of Table 4 shows the initialization of Algorithm 3 . In the first iteration, scenario 164 was removed and the optimal objective values remained unchanged, thus the updated invariant set is $\{589, 1732\}$. In the second and third iterations, removing scenario 589 and 1732 changed optimal solutions, therefore the invariant set $\{589, 1732\}$ cannot be further reduced.

14

Table 4: Finding Essential Set for (nc-RSP)

Iter No.	Scenario Removed	Invariant Set After Removal	Objective Value After Removal	Invariant Set Updated
0	-	{164, 589, 1732}	2.4253×10^9	{164, 589, 1732}
1	164	{589, 1732}	2.4253×10^9	{589, 1732}
2	589	{1732}	2.2765×10^9	{589, 1732}
3	1732	{589}	2.3224×10^9	{589, 1732}

However, notice that the invariant set $\mathcal{I} = \{589, 1732\}$ has cardinality 2, which is greater than our initial guess 1. There are two possibilities of following steps. First, we can stop here if the relaxed guarantee $\epsilon(2) =$ is acceptable. The second choice is to recompute $K = 3012$ using $\epsilon = 0.01$, $\beta = 0.001$ and $|\mathcal{I}| = 2$, solve (nc-RSP) using another independent set of 3012 scenarios, and check if the cardinality of the new invariant set is smaller than 2.

In practice, we can use all available scenarios to obtain more risk-averse (i.e., smaller ϵ) solutions, without reserving a major part of datasets for out-of-sample analysis. This allows us to fully exploit the value of limited data, which is one major advantage of having theoretical guarantees.

6.3. Improve Sample Complexity

The sample complexity can be further reduced. For example, if $|\mathcal{E}| \leq k$ can be *proved* by the special structure of problem, according to *Remark 4* in [18], then equation (9) in [18] can be used to improve the sample complexity in (A.12). If we can prove $|\mathcal{I}| \leq 1$, then the sample complexity of scenarios will be reduced from 2250 to 1410; if $|\mathcal{I}| \leq 2$ is always true, then we only need 2147 scenarios instead of 3012.

7. Concluding Remarks

This paper studies the problem of energy storage planning in future power systems through a novel data-driven scenario approach. Using the two-stage robust formulation, we explicitly account for both shorter-term fluctuations (such as during hourly operation) as well as longer-term uncertainties (such as seasonable and yearly load variations) in the storage planning problem.

Methodologically, we connect two-stage RO optimization with the scenario approach theory to provide theoretical guarantees on the potential risk of planning solutions. The theoretical guarantees hold for both non-convex and convex planning models. We further show that the operation risk consists of two critical components: the risk of exceeding the estimated cost and the risk of load curtailment to avoid infeasible real-time operations.

In the computational aspect, we design numerical algorithms to tighten theoretical guarantees. We show that while solving RSP problems via the C&CG algorithm, we also obtain an invariant set, which is a key component to tighten theoretical guarantees. Numerical results on the 6-bus and 118-bus systems verify the correctness of theoretical results. Due to the structure of the two-stage robust planning problem formulation, we demonstrate that the essential set is typically small and can

be pinpointed at very low computational costs. Numerical results indicates that this observation is regardless of system size, which makes the proposed approach scalable and applicable for large-scale systems.

This paper is a first step towards utilizing theoretically rigorous and computationally scalable approaches to more integrated planning decision that will address both shorter-term and longer-term uncertainties in the future grid. Future work would include: (1) developing rigorous theories which provide upper bounds on the cardinality of essential sets of two-stage robust optimization problems; (2) applying the proposed framework on the joint planning of energy storage, renewable generation, transmission, and many other critical facilities in power systems; and (3) extending the proposed planning framework towards a more detailed data-driven modeling of long-term uncertainties.

Appendix A. Additional Background and Basic Derivations

Appendix A.1. Single-stage Decision Making

Appendix A.1.1. Epigraph Formulation

The standard form of optimization problem is (A.1).

$$\min_{\boldsymbol{x}} \ f_0(\boldsymbol{x}) \tag{A.1a}$$

$$\text{s.t.} \ f_i(\boldsymbol{x}) \leq 0, \ i = 1, 2, \cdots, m. \tag{A.1b}$$

$$h_i(\boldsymbol{x}) = 0, \ i = 1, 2, \cdots, p. \tag{A.1c}$$

The *epigraph formulation* of (A.1) is

$$\min_{\boldsymbol{x}, t} \ t \tag{A.2a}$$

$$\text{s.t.} \ f_0(\boldsymbol{x}) \leq t \tag{A.2b}$$

$$f_i(\boldsymbol{x}) \leq 0, \ i = 1, 2, \cdots, m. \tag{A.2c}$$

$$h_i(\boldsymbol{x}) = 0, \ i = 1, 2, \cdots, p. \tag{A.2d}$$

Proposition 6 (Chapter 4.1 of [46])**.** $(\boldsymbol{x}^\star, t^\star)$ *is optimal for the epigraph formulation* (A.2) *if and only if* $\boldsymbol{x}^\star$ *is optimal for the standard form* (A.1) *and* $t^\star = f_0(\boldsymbol{x}^\star)$.

We want to emphasize the fact that *Proposition 6 holds true even when the original problem* (A.1) *is non-convex.* Although the proof of Proposition 6 is elementary, we provide it below to emphasize the fact that it does not require convexity of (A.1).

Proof of Proposition 6. To prove "⇐", let $\boldsymbol{x}^*$ denote the optimal solution to (A.1) and $t^* := f_0(\boldsymbol{x}^*)$. Clearly $(\boldsymbol{x}^*, t^*)$ is the optimal solution to (A.2). Othwerwise, the optimal solution $\boldsymbol{x}^\star$ to (A.2) is a feasible solution to (A.1) with smaller objective, which causes contradiction.

The proof of "⇒" is identical with above. Clearly $\boldsymbol{x}^\star$ is feasible to (A.1), thus $f_0(\boldsymbol{x}^*) \leq t^\star$. For the purpose of contradiction, we assume that $\boldsymbol{x}^* \neq \boldsymbol{x}^\star$ and $f_0(\boldsymbol{x}^*) < t^\star$. We can define $t^* := f_0(\boldsymbol{x}^*)$, then $(\boldsymbol{x}^*, t^*)$ is a feasible solution to (A.2) with smaller objective objective $t^* < t^\star$, which is a contradiction. □

Appendix A.1.2. Chance-constrained Optimization

A typical chance-constrained program is presented below:

$$\min_{x} c^\mathsf{T} x \tag{A.3a}$$

$$\text{s.t. } f(x) \leq 0 \tag{A.3b}$$

$$\mathbb{P}_\delta(h(x, \delta) \leq 0) \geq 1 - \epsilon \tag{A.3c}$$

where $x \in \mathbb{R}^v$ is the decision variable and random variable $\delta \in \Delta$ denotes the uncertainties. All deterministic constraints are represented by $f(x) \leq 0$. The chance constraint (A.3c) ensures the inner stochastic constraint $h(x, \delta) \leq 0$ is feasible with high probability at least $1 - \epsilon$.

Definition 8 (Violation Probability [15]). The violation probability $\mathbb{V}(x)$ of a candidate solution x is defined as

$$\mathbb{V}(x) := \mathbb{P}(\delta : h(x, \delta) > 0) \tag{A.4}$$

Appendix A.1.3. The Scenario Approach for Convex Problems

The scenario approach is one of the well-known solutions to chance-constrained programs [47].

To solve (A.3), the scenario approach reformulates it to the scenario problem (A.5) with K i.i.d. scenarios $\Delta = \mathcal{K} := \{\delta^{(1)}, \delta^{(2)}, \cdots, \delta^{(K)}\}$.

$$(\text{SP})_\mathcal{K} : \min_{x} c^\mathsf{T} x \tag{A.5a}$$

$$\text{s.t. } f(x) \leq 0 \tag{A.5b}$$

$$h(x, \delta^{(1)}) \leq 0, \cdots, h(x, \delta^{(K)}) \leq 0 \tag{A.5c}$$

The scenario problem (A.5) seeks the optimal solution $x^*_\mathcal{K}$ which is feasible to all K scenarios. With a carefully chosen K, the optimal solution $x^*_\mathcal{K}$ is a feasible solution to the chance-constrained program (A.3), i.e., $\mathbb{V}(x^*_\mathcal{K}) \leq \epsilon$.

Definition 9 (Support Scenario [15]). A scenario $\delta^{(i)}$ is a *support scenario* for the scenario problem $\text{SP}_\mathcal{K}$ if its removal changes the solution of $\text{SP}_\mathcal{K}$. K is the number of all scenarios, $\mathcal{S}$ can denote the set of support scenarios, $|\mathcal{S}|$ is the number of support scenarios.

Theorem 7 (Theorem 2 of [23]). *For any convex scenario problem* (A.5), *the cardinality of its invariant set is no more than the number of decision variables.*

Definition 10 (Non-degenerate Scenario Problem[15]). Let $x^*_\mathcal{K}$ and $x^*_\mathcal{S}$ stand for the optimal solutions to the scenario problems $\text{SP}_\mathcal{K}$ and $\text{SP}_\mathcal{S}$, respectively. The scenario problem $\text{SP}_\mathcal{K}$ is *non-degenerate* if $c^\mathsf{T} x^*_\mathcal{K} = c^\mathsf{T} x^*_\mathcal{S}$.

Assumption 1 (Non-Degeneracy[15]). *For every K, the scenario problem* (A.5) *is non-degenerate with probability 1 with respect to scenarios* $\mathcal{K} := \{\delta^{(1)}, \delta^{(2)}, \cdots, \delta^{(K)}\}$.

Assumption 2 (Feasibility). *Every scenario problem* (A.5) *is feasible, and its feasibility region has a non-empty interior. The optimal solution $x^*_\mathcal{K}$ of* (A.5) *exists.* [12]

---

[12]If the problem has multiple solutions, a tie-break rule can be applied to get one unique solution.

Theorem 8 (Prior Guarantees [15]). *Under Assumptions 1 and 2, let $x^*_\mathcal{K}$ be the optimal solution to the scenario problem $SP_\mathcal{K}$, it holds that*

$$\mathbb{P}^K \{\mathbb{V}(x^*) > \epsilon\} \leq \sum_{i=0}^{v-1} \binom{K}{i} \epsilon^i (1 - \epsilon)^{K-i} \tag{A.6}$$

The probability $\mathbb{P}^K$ is taken with respect to drawing K i.i.d. scenarios, and v is the number of decision variables.

Theorem 9 (Posterior Guarantee[17]). *Given $\beta \in (0, 1)$, for any $k = 0, 1, \cdots, K$, the polynomial equation in variable τ*

$$\frac{\beta}{K+1} \sum_{i=k}^{K} \binom{i}{k} \tau^{i-k} - \binom{K}{k} \tau^{K-k} = 0 \tag{A.7}$$

has exactly one solution $\tau(k)$ in the interval $(0, 1)$. Let $\epsilon(k) := 1 - \tau(k)$. Under Assumptions 1 and 2, it holds that

$$\mathbb{P}^K \left(\mathbb{V}(x^*_\mathcal{K}) \geq \epsilon(|\mathcal{S}|) \right) \leq \beta, \tag{A.8}$$

where $|\mathcal{S}|$ is the number of support scenarios.

Appendix A.1.4. A General Scenario Theory for Non-Convex Problems

In general, if we extend the scenario approach theory to the general case, e.g. non-convex optimization problems. The above approaches usually become inefficient. In a prior approach, there maybe not exist a bound on $|\mathcal{S}|$ for non-convex $(\text{SP})_\mathcal{K}$; in a posterior approach, the *non-degeneracy assumption* is crucial for the wait-and-judge approach, however, most non-convex optimization problems are not guaranteed to be a non-degenerate scenario problem. Recently, [18] extends the scenario approach to the general case with a better bound through removing the *non-degeneracy* assumption.

For non-convex problems, it is usually computationally intractable to find global optimal solutions. There are many algorithms that are capable of finding local optimal solutions. Then, we use $\mathcal{A}_\mathcal{K}$ to represent a (maybe suboptimal) solution to $(\text{SP})_\mathcal{K}$ obtained via algorithm $\mathcal{A}$.

Definition 11 (Support Sub-sample [18]). Given a sample $\left(\delta^{(1)}, \cdots, \delta^{(K)}\right) \in \mathcal{K}$, a support sub-sample O for $\left(\delta^{(1)}, \cdots, \delta^{(K)}\right)$ is a O-tuple of elements extracted from $\left(\delta^{(1)}, \cdots, \delta^{(K)}\right)$, i.e. $O = \left(\delta^{(i_1)}, \cdots, \delta^{(i_o)}\right)$, which gives the same solution as the original sample, that is,

$$\mathcal{A}_O \left(\delta^{(i_1)}, \cdots, \delta^{(i_o)}\right) = \mathcal{A}_K \left(\delta^{(1)}, \cdots, \delta^{(K)}\right) \tag{A.9}$$

A support sub-sample $O = \left(\delta^{(i_1)}, \cdots, \delta^{(i_o)}\right)$ is said to O leaving the solution unchanged be *irreducible* if no element can be further removed from O.

Definition 12 (Irreducible Set [18]). For a support sub-sample, it is said to be *irreducible* if no elements can be removed leaving the solution unchanged.

16

It is possible that there exist many irreducible sets for a non-convex optimization problem, for the one has the minimal cardinality, it is called the essential set in [18].

Theorem 10 (Posterior Probability Guarantee on Non-convex Problem [18]). *Suppose that Assumption 2 holds true, and set a value $\beta \in (0, 1)$ (confidence parameter). Let $\epsilon : \{0, \cdots, K\} \to [0, 1]$ be a function such that*

$$\epsilon(K) = 1 \tag{A.10a}$$

$$\sum_{k=0}^{K-1} \binom{K}{k}(1 - \epsilon(k))^{K-k} = \beta \tag{A.10b}$$

Then, for any $\mathcal{A}_K$ and probability $\mathbb{P}$, it holds that

$$\mathbb{P}^K \left\{ \mathbb{V}\left(x_K^*\right) > \epsilon\left(s_K^*\right) \right\} \leq \beta \tag{A.11}$$

where s_K^* is the cardinality of one support sub-sample.

A simple choice of $\varepsilon(\cdot)$ obtained by splitting β evenly among the K terms in the sum (A.10b) is:

$$\epsilon(k) = \begin{cases} 1 & \text{if } k < K, \\ 1 - \sqrt[K-k]{\frac{\beta}{K\binom{K}{k}}} & \text{otherwise.} \end{cases} \tag{A.12}$$

To apply the probability guarantee for the non-convex problem, we first find an algorithm $\mathcal{A}$ to find the optimal solution for the non-convex problem. Then use the greedy algorithm $\mathcal{B}$ to search for a support sub-sample. The greedy algorithm in [18] can ensure to find an irreducible set, but it needs to solve K non-convex optimization problems, it is sometimes computationally intractable.

Remark 6 (Remark 4 of [18]). Choices for $\epsilon(\cdot)$ other than (A.10b) are possible and, at times, advisable. For example, if from the structure of the problem it was known that s_K^* is always less than some $\bar{s}$, then it would make sense to deliberately ignore all the situations where $s_K^* \geq \bar{s}$, thus allowing for stronger claims when $s_K^* < \bar{s}$. One possible choice is

$$\epsilon(k) = \begin{cases} 1 & \text{if } k < K, \\ 1 - \sqrt[K-k]{\frac{\beta}{\bar{s}\binom{K}{k}}} & \text{otherwise.} \end{cases} \tag{A.13}$$

Appendix B. Proofs and Algorithms

Appendix B.1. Proofs

There are two common formulations of robust optimization: using the *robust* constraint

$$\min_{x \geq 0} c^\mathsf{T} x \tag{B.1a}$$

$$\text{s.t. } Ax \leq d, \ \forall A \in \mathcal{U} \tag{B.1b}$$

or the maximin objective

$$\max_{A \in \mathcal{U}} \min_{x \geq 0} c^\mathsf{T} x \tag{B.2a}$$

$$\text{s.t. } Ax \leq d. \tag{B.2b}$$

Many researchers have studied the relationship between the two formulations above, e.g., [48, 49]. The main results is that these two formulations are equivalent if the uncertainty is constraint-wise. For robust linear programs, the constraint-wise assumption is always satisfied (see Section 2.1 of [30]), thus these two formulations are equivalent.

Theorem 11 ([48], Section 8 of [30]). *The robust linear program* (B.1) *is equivalent the maximin formulation* (B.2).

Proof of Proposition 1. Proposition 1 can be seen as an extention of Theorem 11 towards two-stage robust optimization problems. For a given x, the second-stage problem $\max_{\delta \in \Delta} \min_{y \in \mathcal{Y}(x;\delta)} d^\mathsf{T} y$ is equivalent with

$$\min_y d^\mathsf{T} y \tag{B.3a}$$

$$\text{s.t. } y \in \mathcal{Y}(x;\delta), \ \forall \delta \in \Delta \tag{B.3b}$$

Using the epigraph formulation (see Proposition 6), (B.3) can be equivalently formulated as

$$\min_{y,\gamma} \gamma \tag{B.4a}$$

$$\text{s.t. } \gamma \geq d^\mathsf{T} y \tag{B.4b}$$

$$y \in \mathcal{Y}(x;\delta), \ \forall \delta \in \Delta \tag{B.4c}$$

Putting two stages together, we get

$$\min_{x,y,\gamma} c^\mathsf{T} x + \gamma \tag{B.5a}$$

$$\text{s.t. } x \in \mathcal{X} \tag{B.5b}$$

$$\gamma \geq d^\mathsf{T} y \tag{B.5c}$$

$$y \in \mathcal{Y}(x;\delta), \ \forall \delta \in \Delta \tag{B.5d}$$

The last step is to define $\mathcal{Z}(\delta) := \{(x, \gamma) : \exists y \in \mathcal{Y}(x;\delta) \text{ and } d^\mathsf{T} y \leq \gamma\}$, which is essentially projecting the original feasible set $\{(x, y, \gamma) : \gamma \geq d^\mathsf{T} y \text{ and } y \in \mathcal{Y}(x;\delta), \ \forall \delta \in \Delta\}$ onto the x-space.

$\square$

Proof of Corollary **??**. Corollary **??** is a direct conclusion of Theorem 1 and Proposition 1. (31) is simply replacing d with $2|\mathcal{S}| + 1$ in (18).
$\square$

Remark 7. The additional variable γ might seem unnecessary, the following formulation is equivalent without γ:

$$\min_{x,y,\gamma} c^\mathsf{T} x + d^\mathsf{T} y \tag{B.6a}$$

$$\text{s.t. } x \in \mathcal{X} \tag{B.6b}$$

$$y \in \mathcal{Y}(x;\delta), \ \forall \delta \in \Delta \tag{B.6c}$$

However, the number of support scenarios of formulation (14) is much smaller than (B.6).

Proof of Proposition 3. For the convex formulation (c-RSP), there are only two first-stage decision variables E_n and P_n. Using Proposition 1, (c-RSP) can be equivalently formulated as a single-stage problem (14) with an additional variable γ. This

17

new formulation does not have any second-stage decision variables $\boldsymbol{y} = (p_{n,t}^{\text{ch}}, p_{n,t}^{\text{dis}}, p_{i,n,t}^{\mathcal{G}}, p_{n,t}^{\text{shed}}, v_{n,t})$. The last step is to apply Theorem 7 on (14), which consists of $2|\mathcal{S}| + 1$ decision variables (E_n, P_n, and γ). □

Proof of Proposition 4. When $\boldsymbol{\delta}^*$ is the unique solution to $\arg\max_{\boldsymbol{\delta} \in \mathcal{K}} f(\boldsymbol{\delta})$, simply removing $\boldsymbol{\delta}^*$ from $\mathcal{K}$ will leads to smaller optimal objective

$$\max_{\boldsymbol{\delta} \in \mathcal{K}} f(\boldsymbol{\delta}) > \max_{\boldsymbol{\delta} \in \mathcal{K} - \boldsymbol{\delta}^*} f(\boldsymbol{\delta}) \tag{B.7}$$

Thus $\boldsymbol{\delta}^*$ is clearly a support scenario. Also note that removing any other scenarios in $\mathcal{K} - \boldsymbol{\delta}^*$ will not change the optimal solution to $\max_{\boldsymbol{\delta} \in \mathcal{K}} f(\boldsymbol{\delta})$. This concludes the proof. □

Appendix B.2. Algorithms

Algorithm 2 column-and-constraint generation (C&CG) [29]

1: set LB $= -\infty$, UB $= +\infty$, $\mathcal{K} := \{\boldsymbol{\delta}^{(1)}, \boldsymbol{\delta}^{(2)}, \cdots, \boldsymbol{\delta}^{(K)}\}$, $O = \emptyset$ and $k = 0$.
2: **while** UB $-$ LB $\leq \varepsilon$ **do**
3:    solve the master problem (B.8),

$$\min_{\gamma, \boldsymbol{x} \in \mathcal{X}} \boldsymbol{c}^\mathsf{T} \boldsymbol{x} + \gamma \tag{B.8a}$$

$$\text{s.t. } \boldsymbol{d}^\mathsf{T} \boldsymbol{y}^{(l)} \leq \gamma, \ l = 1, 2, \cdots, k \tag{B.8b}$$

$$\boldsymbol{y}^{(l)} \in \mathcal{Y}(\boldsymbol{x}, \boldsymbol{\delta}^{*(l)}), \ l = 1, 2, \cdots, k \tag{B.8c}$$

obtain an optimal solution $(\boldsymbol{x}_{k+1}^*, \gamma_{k+1}^*, \boldsymbol{y}^{1*}, \cdots, \boldsymbol{y}^{k*})$.
4:    update LB $\leftarrow \boldsymbol{c}^\mathsf{T} \boldsymbol{x}_{k+1}^* + \gamma_{k+1}^*$.
5:    solve the sub-problem (B.9) with $\boldsymbol{x}_{k+1}^*$,

$$\max_{\boldsymbol{\delta} \in \mathcal{K}} \min_{\boldsymbol{y} \in \mathcal{Y}(\boldsymbol{x}_{k+1}^*, \boldsymbol{\delta})} \boldsymbol{d}^\mathsf{T} \boldsymbol{y} \tag{B.9}$$

6:    if (B.9) is feasible, obtain the optimal solution $\boldsymbol{y}_{k+1}^*$ and $\boldsymbol{\delta}^{*(k+1)}$. Create variable $\boldsymbol{y}^{(k+1)}$, add the following constraints to (B.8).

$$\boldsymbol{d}^\mathsf{T} \boldsymbol{y}^{(k+1)} \leq \gamma \tag{B.10a}$$

$$\boldsymbol{y}^{(k+1)} \in \mathcal{Y}(\boldsymbol{x}, \boldsymbol{\delta}^{*(k+1)}) \tag{B.10b}$$

7:    if (B.9) is infeasible, $\boldsymbol{c}^\mathsf{T} \boldsymbol{x}_{k+1}^* + \boldsymbol{d}^\mathsf{T} \boldsymbol{y}_{k+1}^* = +\infty$. Create variable $\boldsymbol{y}^{(k+1)}$, add the following constraints to (B.8).

$$\boldsymbol{y}^{(k+1)} \in \mathcal{Y}(\boldsymbol{x}, \boldsymbol{\delta}^{*(k+1)}) \tag{B.11}$$

where $\boldsymbol{\delta}^{*(k+1)}$ is the identified scenario for the infeasible (B.9).
8:    update UB $\leftarrow \min(\text{UB}, \boldsymbol{c}^\mathsf{T} \boldsymbol{x}_{k+1}^* + \boldsymbol{d}^\mathsf{T} \boldsymbol{y}_{k+1}^*)$.
9:    update k and O as $O \leftarrow \{O, \boldsymbol{\delta}^{*(k+1)}\}$, $k \leftarrow k + 1$;
10: **end while**
11: **return** $\boldsymbol{x}^*$, $\boldsymbol{y}^*$ and O.

Algorithm 3 Find an Essential Set for the Robust Storage Planning

---

Input: $\Delta = \mathcal{K} = \{\boldsymbol{\delta}^{(1)}, ..., \boldsymbol{\delta}^{(K)}\}$;
   Find an Invariant Set :
1: Solve (c-RSP) or (nc-RSP) with scenario-based uncertainty set $\mathcal{K}$ by C&CG algorithm, and obtain the set of scenarios in the master problem O.
2: Delete the repetitive scenarios $\boldsymbol{\delta}^*$ in O , and obtain an Invariant Set $O = \{\boldsymbol{\delta}^{(i_1)}, ..., \boldsymbol{\delta}^{(i_{|O|})}\}$
   Find an Essential Set
3: $\mathcal{I}_\mathcal{K} \leftarrow O$; solve (c-RSP) or (nc-RSP) with the uncertainty set O and obtain the solution x_{O^*};
4: **for** o=1 to $|O|$ **do**
5:    $\overline{\mathcal{I}}_\mathcal{K} = \mathcal{I}_\mathcal{K} \backslash \boldsymbol{\delta}^{(i_o)}$; solve (c-RSP) or (nc-RSP) with the uncertainty set $\overline{\mathcal{I}}_\mathcal{K}$ and obtain the solution $x_{\overline{\mathcal{I}}_\mathcal{K}}^*$.
6:    **if** $x_{\overline{\mathcal{I}}_\mathcal{K}}^* = x_O^*$ **then**
7:       $\mathcal{I}_\mathcal{K} \leftarrow \overline{\mathcal{I}}_\mathcal{K}$.
8:    **end if**
9: **end for**
10: **return** Essential Set $\mathcal{I}_\mathcal{K}$ and its cardinality $|\mathcal{I}_\mathcal{K}|$.

References

[1] Grid energy storage, Tech. rep., Department of Energy (2013).

[2] US Department of Energy, DOE OE Global Energy Storage Database. URL https://www.sandia.gov/ess-ssl/global-energy-storage-database-home/

[3] Global energy storage outlook: H2 2020, Tech. rep., Wood Mackenzie (2020).

[4] A. Conejo, L. Baringo, J. Kazempour, A. Siddiqui, Investment in Electricity Generation and Transmission, 2016. doi:10.1007/978-3-319-29501-5.

[5] V. Oree, S. Z. S. Hassen, P. J. Fleming, Generation expansion planning optimisation with renewable energy integration: A review, Renewable and Sustainable Energy Reviews 69 (2017) 790 – 803.

[6] L. Baringo, A. J. Conejo, Transmission and wind power investment, IEEE Transactions on Power Systems 27 (2) (2012) 885–893.

[7] H. Park, R. Baldick, Transmission planning under uncertainties of wind and load: Sequential approximation approach, IEEE Transactions on Power Systems 28 (3) (2013) 2395–2402.

[8] H. Park, R. Baldick, D. P. Morton, A stochastic transmission planning model with dependent load and wind forecasts, IEEE Transactions on Power Systems 30 (6) (2015) 3003–3011.

[9] R. Domínguez, A. J. Conejo, M. Carrión, Toward fully renewable electric energy systems, IEEE Transactions on Power Systems 30 (1) (2015) 316–326.

[10] P. Wu, H. Cheng, J. Xing, The interval minimum load cutting problem in the process of transmission network expansion planning considering uncertainty in demand, IEEE Transactions on Power Systems 23 (3) (2008) 1497–1506.

[11] C. Ruiz, A. Conejo, Robust transmission expansion planning, European Journal of Operational Research 242 (2) (2015) 390 – 401.

[12] L. Baringo, A. Baringo, A stochastic adaptive robust optimization approach for the generation and transmission expansion planning, IEEE Transactions on Power Systems 33 (1) (2018) 792–802.

[13] F. Verástegui, Á. Lorca, D. E. Olivares, M. Negrete-Pincetic, P. Gazmuri, An adaptive robust optimization model for power systems planning with operational uncertainty, IEEE Transactions on Power Systems 34 (6) (2019) 4606–4616.

[14] Z. Liang, H. Chen, S. Chen, Y. Wang, C. Zhang, C. Kang, Robust transmission expansion planning based on adaptive uncertainty set optimization under high-penetration wind power generation, IEEE Transactions on Power Systems (2020) 1–1.

[15] M. C. Campi, S. Garatti, The exact feasibility of randomized solutions of uncertain convex programs, SIAM Journal on Optimization (2008).

[16] G. C. Calafiore, Random convex programs, SIAM Journal on Optimization (2010).

[17] M. Campi, S. Garatti, Wait-and-judge scenario optimization, Mathematical Programming (2016) 1–35.

[18] M. C. Campi, S. Garatti, F. A. Ramponi, A general scenario theory for non-convex optimization and decision making, IEEE Transactions on Automatic Control (2018).

[19] M. Vrakopoulou, K. Margellos, J. Lygeros, G. Andersson, A probabilistic framework for reserve scheduling and $n-1$ security assessment of systems with high wind power penetration, IEEE Transactions on Power Systems 28 (4) (2013) 3885–3896.

[20] M. S. Modarresi, L. Xie, M. C. Campi, S. Garatti, A. Carè, A. A. Thatte, P. R. Kumar, Scenario-based economic dispatch with tunable risk levels in high-renewable power systems, IEEE Transactions on Power Systems 34 (6) (2019) 5103–5114.

[21] H. Ming, L. Xie, M. C. Campi, S. Garatti, P. R. Kumar, Scenario-based economic dispatch with uncertain demand response, IEEE Transactions on Smart Grid 10 (2) (2019) 1858–1868.

[22] X. Geng, L. Xie, Chance-constrained unit commitment via the scenario approach, in: 2019 North American Power Symposium (NAPS), 2019, pp. 1–6.

[23] X. Geng, L. Xie, M. S. Modarresi, Computing Essential Sets for Convex and Non-convex Scenario Problems: Theory and Application, arXiv preprint arXiv:1910.07672 (2019).

[24] D. Bertsimas, M. Sim, Tractable approximations to robust conic optimization problems, Mathematical Programming 107 (1-2) (2006) 5–36.

[25] A. Ben-Tal, L. El Ghaoui, A. Nemirovski, Robust optimization, Princeton University Press, 2009.

[26] P. Yang, A. Nehorai, Hybrid energy storage and generation planning with large renewable penetration, in: 2013 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAM-SAP), IEEE, 2013, pp. 460–463.

[27] P. Yang, A. Nehorai, Joint optimization of hybrid energy storage and generation capacity with renewable energy, IEEE Transactions on Smart Grid 5 (4) (2014) 1566–1574.

[28] C. Yan, X. Geng, L. Xie, Z. Bie, A Scenario-based Storage Planning Framework with Probabilistic Guarantees, in: 2019 IEEE Sustainable Power and Energy Conference (iSPEC), IEEE, 2019, pp. 2341–2346.

[29] B. Zeng, L. Zhao, Solving two-stage robust optimization problems using a column-and-constraint generation method, Operations Research Letters 41 (5) (2013) 457–461.

[30] B. L. Gorissen, I. Yankoglu, D. den Hertog, A practical guide to robust optimization, Omega 53 (2015) 124–137.

[31] C. Yan, X. Geng, Z. Bie, L. Xie, Two-stage Robust Energy Storage Planning with Probabilistic Guarantees: A Data-driven Approach, arXiv preprint arXiv:2103.16424 (2021).

[32] B. Xu, Y. Wang, Y. Dvorkin, R. Fernández-Blanco, C. A. Silva-Monroy, J.-P. Watson, D. S. Kirschen, Scalable planning for energy storage in energy and reserve markets, IEEE Transactions on Power Systems 32 (6) (2017) 4515–4527.

[33] E. Hsia, Energy Storage in PJM: A Perspective (2019).
URL https://insidelines.pjm.com/energy-storage-in-pjm-a-perspective/

[34] ISO New England, 2020 Regional Electricity Outlook, Tech. rep. (2020).
URL https://www.iso-ne.com/static-assets/documents/2020/02/2020_reo.pdf

[35] R. Fernández-Blanco, Y. Dvorkin, B. Xu, Y. Wang, D. S. Kirschen, Optimal energy storage siting and sizing: A wecc case study, IEEE Transactions on Sustainable Energy 8 (2) (2017) 733–743.

[36] T. Sayfutdinov, H. Patsios, P. Vorobev, E. Gryazina, D. M. Greenwood, J. W. Bialek, P. C. Taylor, Degradation and operation-aware framework for the optimal siting, sizing and technology selection of battery storage, IEEE Transactions on Sustainable Energy (2019) 1–1.

[37] T. Qiu, B. Xu, Y. Wang, Y. Dvorkin, D. S. Kirschen, Stochastic multistage coplanning of transmission expansion and energy storage, IEEE Transactions on Power Systems 32 (1) (2017) 643–651.

[38] R. A. Jabr, I. Džafić, B. C. Pal, Robust optimization of storage investment on transmission networks, IEEE Transactions on Power Systems 30 (1) (2015) 531–539.

[39] H. Akhavan-Hejazi, H. Mohsenian-Rad, Energy storage planning in active distribution grids: A chance-constrained optimization with non-parametric probability functions, IEEE Transactions on Smart Grid 9 (3) (2018) 1972–1985.

[40] M. Sun, J. Cremer, G. Strbac, A novel data-driven scenario generation framework for transmission expansion planning with high renewable energy penetration, Applied Energy 228 (2018) 546 – 555.

[41] Z. Li, Q. Guo, H. Sun, J. Wang, Sufficient conditions for exact relaxation of complementarity constraints for storage-concerned economic dispatch, IEEE Transactions on Power Systems 31 (2) (2016) 1653–1654.

[42] S. Wogrin, D. F. Gayme, Optimizing storage siting, sizing, and technology portfolios in transmission-constrained networks, IEEE Transactions on Power Systems 30 (6) (2015) 3304–3313.

[43] T. Sayfutdinov, C. Patsios, P. Vorobev, E. Gryazina, D. M. Greenwood, J. W. Bialek, P. C. Taylor, Degradation and operation-aware framework for the optimal siting, sizing, and technology selection of battery storage, IEEE Transactions on Sustainable Energy 11 (4) (2020) 2130–2140.

[44] G. Calafiore, M. C. Campi, Uncertain convex programs: randomized solutions and confidence levels, Mathematical Programming (2005).

[45] A. Bagheri, J. Wang, C. Zhao, Data-driven stochastic transmission expansion planning, IEEE Transactions on Power Systems 32 (5) (2017) 3461–3470.

[46] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge University Press, 2004.

[47] X. Geng, L. Xie, Data-driven decision making in power systems with probabilistic guarantees: Theory and applications of chance-constrained optimization, Annual Reviews in Control (2019). doi:http://doi.org/10.1016/j.arcontrol.2019.05.005.

[48] A. Beck, A. Ben-Tal, Duality in robust optimization: primal worst equals dual best, Operations Research Letters 37 (1) (2009) 1–6.

[49] A. Soyster, F. Murphy, A unifying framework for duality and modeling in robust linear programs, Omega 41 (6) (2013) 984–997, publisher: Elsevier.