Distributionally Robust Chance-Constrained Generation Expansion Planning

Farzaneh Pourahmadi, *Student Member, IEEE*, Jalal Kazempour, *Senior Member, IEEE*, Christos Ordoudis, *Member, IEEE*, Pierre Pinson, *Senior Member, IEEE*, Seyed Hamid Hosseini, *Member, IEEE*

***Abstract*—This paper addresses a centralized generation expansion planning problem, accounting for both long- and short-term uncertainties. The long-term uncertainty (demand growth) is modeled via a set of scenarios, while the short-term uncertainty (wind power production) is described by a family of probability distributions having the same first- and second-order moments obtained from historical data. In the expansion stage, the optimal units to be built are selected among discrete options. In the operational stage, a detailed representation of unit commitment constraints is considered. To make this problem tractable, we solve it with linear decision rules, and use a tight relaxation approach to convexify the unit commitment constraints. The resulting model is a distributionally robust chance-constrained optimization problem, which is eventually recast as a mixed-integer second-order cone program. We consider the IEEE 118-bus test system as a case study, and explore the performance of the proposed model using an out-of-sample analysis. It is demonstrated that handling the short-term uncertainty by the proposed distributionally robust model gives more effective out-of-sample performance in terms of system cost and reliability compared to a chance-constrained model that assumes a Gaussian distribution of uncertainty, while maintaining computational tractability. Similar out-of-sample performance is observed when comparing the proposed model against a chance-constrained program using the empirical distribution, is recast as either a robust optimization or a stochastic program.**

***Index Terms*—Distributionally robust optimization, chance constraints, conic programming, linear decision rules, generation expansion planning, unit commitment.**

I. Introduction

The increasing integration of stochastic renewable energy sources exposes power systems to a more variable and uncertain generation profile. In particular, this brings two main challenges to long-term planning decision-making problems. First, it is complicated in the long-term planning horizon to accurately model the short-term uncertainties pertaining to stochastic production. Second, the emerging variability in net-load (demand minus renewable power production) necessitates the considerations of operational limits in the planning problems. These challenges require the development of advanced decision-making tools for capacity planning studies that accurately take into account the short-term uncertainties and effectively define operational flexibility requirements.

In this paper, we focus on long-term expansion planning problems in terms of generation assets. In a market context, the generation investment falls into the internal decisions of the power producers, who are in general profit maximizers. However, a centralized generation expansion planning (GEP) model, which is the focus of this paper, can be still useful in market-related studies. The reason is that this centralized model can serve as an *ideal benchmark*, and provide the market regulators with insights into the planning decisions which are optimal for the whole system. The market regulator may then design proper policies to incentivize the producers to invest in those assets which are optimal from social welfare perspective [1].

An appropriate GEP model requires modeling both *short-term* (e.g., renewable power production) and *long-term* (e.g., demand growth) uncertainties [2]. In particular, it is complicated to model the short-term uncertainties over the whole planning horizon. In the existing literature, stochastic programming and robust optimization are commonly used to model short-term uncertainty in GEP problems, but each of those techniques has its own shortcomings [3]–[5]. The performance of stochastic GEP models highly depends on the set of scenarios representing the sources of uncertainty. It is crucial to ensure that the set of scenarios describes well the true probability distribution [6]. However, this usually requires embedding a very large number of scenarios, which will eventually end up with computational intractability. A reduction in the number of scenarios may cause failing a proper representation of the uncertainty distribution, and thereby, may yield a weak out-of-sample performance [7]. On the other hand, the GEP models based on robust optimization make the optimal planning decisions against the worst-case realization of a prescribed uncertainty set, which may achieve over conservative solutions. It is also complicated to define a proper uncertainty set, accounting for all potential distributions of the short-term uncertainties.

In addition to the aforementioned challenges, there is another one concerning two common simplifications that are made in long-term planning studies, which might be no longer valid. The first one is to ignore the chronological variations of the net-load by using a net-load duration curve over the planning horizon [14]. The second one is to constrain the GEP optimization model by economic dispatch limits only, discarding the operational unit commitment (UC) constraints, e.g., start-up, shut-down, and minimum up/down time limits. There are several recent studies showing that these simplifications are distorting the expansion planning decisions in power systems with high renewable power penetration [4], [8]–[10], [15]. In particular, these simplifications may lead to an underestimation of the need for new generation capacity, and

F. Pourahmadi and S. H. Hosseini are with the Department of Electrical Engineering, Sharif University of Technology, Tehran 11365-11155, Iran (e-mails: Pourahmadi_f@ee.sharif.edu; hosseini@sharif.edu).

J. Kazempour, C. Ordoudis, and P. Pinson are with the Department of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby 2800, Denmark (e-mails: seykaz@elektro.dtu.dk; christos.ordoudis@gmail.com; ppin@elektro.dtu.dk).

TABLE I
THE COMPARISON OF RELEVANT WORKS PROPOSED IN THE LITERATURE AND THE MODEL PROPOSED IN THIS PAPER

Ref.	Problem	Operational constraints	Uncertainty modeling technique	Correlation of uncertainties	Out-of-sample analysis	Resulting model
[8]–[10]	GEP[1]	Unit commitment	Deterministic	No	No	MILP[2]
[3]	GEP	Economic dispatch	Stochastic programming	No	No	MILP
[4], [11]	GEP	Unit commitment	Stochastic programming	No	No	MILP
[5]	GEP	Economic dispatch	Robust optimization	No	Yes	MILP
[12], [13]	TEP[3]	Economic dispatch	Distributionally robust optimization	No	Yes	MILP
This paper	GEP	Unit commitment	Distributionally robust chance-constrained optimization	Yes	Yes	MISOCP[4]

[1]GEP: Generation expansion planning; [2]MILP: Mixed-integer linear program
[3]TEP: Transmission expansion planning; [4]MISOCP: Mixed-integer second-order cone program

eventually a significant increase in the actual system cost [11].

Based on all these challenges, we address the following three research questions: How can the renewable production uncertainty be properly modeled in a GEP problem when the true probability distribution is unknown? How important is it to model the spatial and temporal correlations of renewable production uncertainty in the GEP problem? And, how to ensure that the operational limits including UC constraints are properly taken into account in the GEP problem while maintaining the GEP problem computationally tractable?

To properly address all these questions, we propose a distributionally robust chance constrained GEP model. In this model, an *ambiguity set* represents a family of probability distributions, all with identical first- and second-order moments (mean and covariance), each representing a potential distribution of short-term uncertainties[5]. In addition, the long-term uncertainties are modeled using a set of scenarios drawn from expert knowledge. The proposed model is constrained by expansion limits as well as operational limits, including a tight convex relaxation of unit commitment constraints. This allows making the expansion planning decisions, which are optimal in annualized total expected cost for both the long-term scenarios and the worst-case distribution of short-term uncertainty in the ambiguity set. By using chance constraints, we can adjust the conservativeness of planning decisions obtained by optimizing against the worst probability distribution of short-term uncertainty. Adopting linear decisions rules, the proposed model is eventually recast as a mixed-integer second-order cone program with good computational performance. Note that the resulting model is mixed-integer due to discrete representation of expansion options. Otherwise, with a simpler but continuous representation of candidate units to be built, the resulting model would be a second-order cone program without integrality constraints.

As summarized and compared in Table I, some existing GEP models focus on various types of approximations to embed the effect of UC constraints and the variability of net load in the GEP problem while maintaining its tractability [8]–[10]. In these papers, the uncertainty of stochastic power producers is neglected. Some other studies focus on characterizing uncertainties using either stochastic programming (e.g., [3], [4], [11]) or robust optimization (e.g., [5]), with or without enforcing UC constraints. To the best of our knowledge, there is no other work in the existing literature proposing a distributionally robust chance-constrained GEP model. References [12] and [13] use a distributionally robust optimization with a moment-based ambiguity set for expansion planning problems, but their purpose is to determine the optimal investment in transmission assets. Besides, it is worth mentioning that the ambiguity sets in [12] and [13] consider the first moment information of uncertainties only. Therefore, these models miss the potential temporal and spatial correlations of uncertainties, and also build very big ambiguity sets, which may end up to very conservative solutions, especially without enforcing chance constraints. Also, these two papers neglect the UC constraints.

In addition to [12] and [13], distributionally robust optimization has been recently applied to an extensive range of applications in power systems, such as UC [22]–[24], optimal power flow [25]–[29], hydroelectric reservoir optimization [30], congestion line management [31], energy management in microgrids [32], and multi-stage distribution planning [33].

Given the context above, the main contributions of this paper are as follows: we develop a moment-based distributionally robust chance constrained GEP model in which a tight relaxation of UC with chance constraints is incorporated at the operation stage of the GEP problem. This allows to incorporate the effect of both long- and short-term uncertainty as well as flexibility requirements into generation expansion planning problem, and adjust its conservatism by varying the desired confidence level of chance constraints. Moreover, by considering short-term moment-based ambiguity set including second-order moment information, the spatial and temporal correlations of short-term uncertainty on inter-temporal constraints, e.g., ramp-rate constraints, are properly modeled. In addition, our findings show that the proposed GEP model based on a distributionally robust chance-constrained optimization exhibits a better out-of-sample performance compared to a

[5]In the existing literature of distributionally robust optimization, two types of ambiguity sets are extensively used: moment-based and metric-based. In the former, the set includes all probability distributions whose moments (e.g., mean and covariance) are identical (or close enough) to those of empirical data [16]–[18]. In the latter, the ambiguity set is defined as a ball in the space of distributions, such that the empirical distribution is considered in the center, and the ball is constructed around the center using a probability distance functions, such as ϕ-divergence, Kullback-Leibler divergence, or Wasserstein metric [19], [20]. Although the Wasserstein-based ambiguity sets guarantee a comparatively stronger out-of-sample performance, the moment-based ambiguity sets provide better tractability features [20]. Besides, a distributionally robust optimization with moment-based ambiguity set is in general more tractable in comparison to its stochastic counterpart, whereas a distributionally robust model build upon Wasserstein metric is generally more computationally expensive [20]. A comprehensive review for distributionally robust optimization is available in [21]. In this paper, we limit our attention to distributionally robust optimization with a moment-based ambiguity set, where the values of the first two moments are exactly known.

similar model relying on a chance-constrained optimization following a Gaussian distribution. The same finding still holds when we consider a chance-constrained optimization using the empirical distribution, is recast as either a robust optimization or a stochastic program.

The remainder of this paper is structured as follows. Section II provides some preliminaries on expansion time horizon and uncertainty modeling, and lists our assumptions. Section III presents the proposed distributionally robust chance-constrained GEP model. Section IV explains the solution methodology. Section V provides a comprehensive case study based on the IEEE 118-bus test system. Section VI concludes the paper. Appendix A provides a nomenclature. We also include additional five appendices to provide extra technical information, which will be required throughout the paper.

II. Modeling Framework and Assumptions

A. Expansion Time Horizon

Two different approaches are commonly used in expansion planning studies to model the planning time horizon: static (one-year) and dynamic (multi-year). In the static expansion approach, a single target year only (e.g., the 20th year from now) is considered as the expansion time horizon. This model indeed finds the optimal generation assets required to be built by the target year, but does not determine the optimal years of expansion from now to the target year. On the other hand, in the dynamic expansion approach, the planning time horizon is divided into multiple years, which allows to determine the optimal time of expansion plans in addition to the new capacities to be built. More information on the differences of these two approaches are available in [1]. The proposed problem in this paper follows a static expansion model with a single target year – this means that we simply consider *annualized* expansion and operations costs. It is straightforward to extend it to a dynamic model, but the computational time may drastically increase, and a decomposition method will be eventually required.

It is also common in expansion planning studies to consider a load duration curve (LDC), or a net-load duration curve, as an approximation of hourly load curve during the expansion time horizon [34]. The aim of such LDC-based expansion models is to meet the capacity adequacy requirements only. However, its main drawback is that it ignores the chronological sequence of load profile, and thereby, the operational flexibility limitations cannot be modeled. To resolve this issue while maintaining the problem computationally tractable, the net-load during the planning horizon can be represented through a set of representative operating periods (days or weeks) [35]–[37]. In this paper, we use a set of representative days $r \in \varphi$, each including 24 hours. We assign a weight to each representative day r, indicating the number of days in the target year, whose net-load profile is represented by day r.

B. Uncertainty Modeling

In capacity expansion studies, the sources of uncertainty can be generally categorized into long- and short-term uncertainties. The former corresponds to uncertain sources to be realized in long run, e.g., load growth, future fuel cost, future share of renewables in power systems, and policy regulations. In contrast, the short-term uncertainties pertain to uncertain sources to be realized in short-run operational stage, e.g., stochastic generation of renewables and consumption of loads.

In general, there is a lack of information related to the probability distribution of long-term uncertainties, as discussed in [13]. Therefore, we model the uncertainty of demand growth and future capacity of renewables through a finite set of possible scenarios drawn by expert knowledge.

For each long-term scenario s, representative day $r \in \varphi$ and hour $t \in \{1, 2, ..., 24\}$, we model the short-term net-load uncertainty by means of an ambiguity set, i.e., a family of possible probability distributions of net-load, which is built based on the information associated with the first- and second-order moments (i.e., mean and covariance) of available historical data. We assume that the *exact* first- and second-order moments can be estimated by the available historical data, and leave modeling of the potential *inexact* moments to our future works[6]. We consider Z short-term uncertain sources. The production of uncertain sources in short run is modeled by $\mathbf{m}_{srt} + \boldsymbol{\gamma}_{srt}$, where $\mathbf{m}_{srt} \in \mathbb{R}^Z$ is the mean of forecasted future production of uncertain sources, while the forecast error $\boldsymbol{\gamma}_{srt}$ under a given probability distribution is a random variable with the mean vector $\boldsymbol{\mu}_{srt} \in \mathbb{R}^Z$ and the covariance matrix $\boldsymbol{\Sigma}_{srt} \in \mathbb{R}^{Z\times Z}$. According to these definitions, the ambiguity set $\mathcal{P}_{srt}$ associated with long-term scenario s, representative day r, and hour t can be written as

$$\mathcal{P}_{srt} = \{D \in \Psi_{srt}(\mathbb{R}^Z) : \mathbb{E}^D(\boldsymbol{\gamma}) = \boldsymbol{\mu}_{srt}, \mathbb{E}^D(\boldsymbol{\gamma}^\top \boldsymbol{\gamma}) = \boldsymbol{\Sigma}_{srt}\}, \quad (1)$$

where D is a probability distribution, belonging to family of distributions $\Psi_{srt}(\mathbb{R}^Z)$, all with the same first- and second-order moments, including the uncertainty information for all short-term uncertain sources. Note that $(.)^\top$ is the transpose operator and $\mathbb{R}$ is the set of real numbers. For notational convenience, the indices of $\boldsymbol{\gamma}_{srt}$ are dropped in (1) and in the rest of the paper. We hereafter assume that the mean vector of forecast error is zero, i.e., $\boldsymbol{\mu}_{srt} = 0, \ \forall s, r, t$.

C. Additional Assumptions

Following the majority of expansion planning studies in literature, we use a linearized loss-less DC approximation to represent the network power flow. It is assumed that the topology of network is not changed throughout the planning horizon. We also discard quick-start generating units, meaning that the on/off commitment status of all conventional generators are determined before the short-term uncertainty realization, and this status cannot be changed in real time. However, we consider that this commitment variables are dependent on long-term scenarios. Finally, we assume that all renewable sources are scheduled at their mean output power with a zero production cost.

[6]Note that inexact (estimated) moments may complicate the model, and eventually yield a semi-definite program [38]. It might be also of interest to incorporate more information into the ambiguity set, e.g., extra moments, a support set, or unimodality and log-concavity properties [39]. This allows excluding the potentially non-realistic distributions from the ambiguity set, eventually increasing the quality of the results. However, they may again complicate the model, and we leave them for our future extensions.

III. Proposed Model

In this section, we present the proposed distributionally robust chance-constrained (DRCC) model for the GEP problem. We first define all notation, though a list of notation is also provided in Appendix A.

A. Notation

First, let us define indices and parameters. We use index $i \in \{1, 2, ..., G\}$ for all conventional generating units, including both existing and candidate units. Note that G^C out of G units are the candidate generators to be built. For each unit i, parameters $\overline{p}_i$ and $\underline{p}_i$ indicate its maximum and minimum production level, respectively. Parameters $\overline{r}_i$ and $\underline{r}_i$ represent the maximum ramp-up and ramp-down capability of unit i, respectively. Parameter r_i indicates the ramp rate limit of unit i in the time of start-up or shut-down only. Parameter vectors $\mathbf{c} \in \mathbb{R}^G$ and $\mathbf{h} \in \mathbb{R}^G$ refer to the production and start-up costs of generating units, respectively. Parameter vectors $\underline{\mathbf{v}} \in \mathbb{R}^G$ and $\overline{\mathbf{v}} \in \mathbb{R}^G$ give the minimum up- and down-time of generating units. We also consider B demands. For given demand growth scenario s to be realized in long run, we assume that the short-term load $\mathbf{d}_{srt} \in \mathbb{R}^B$ is inelastic to price and exactly known. With this assumption, the net-load uncertainty in short run boils down to renewable power production uncertainty. Parameter vector $\mathbf{k} \in \mathbb{R}^{G^C}$ indicates the annualized capital cost of candidate units. Parameter $\overline{f}_l$ refers to the capacity of line l. The power network consists of U nodes and L transmission lines. The power flow throughout the network is determined using a power transfer distribution factor matrix $\mathbf{H} \in \mathbb{R}^{L \times U}$, which defines the power flow as a linear function of nodal injections. More specifically, we define three matrices $\mathbf{H}^{\mathrm{G}} \in \mathbb{R}^{L \times G}$, $\mathbf{H}^{\mathrm{W}} \in \mathbb{R}^{L \times Z}$, and $\mathbf{H}^{\mathrm{D}} \in \mathbb{R}^{L \times B}$ that incorporate the mapping of generating units, renewable sources, and demands on the network, respectively. Note that we use upper-case bold letters for matrices and lower-case bold letters for vectors. Besides, $\mathbf{1}$ and $\mathbf{0}$ are the vector of ones and zeros, respectively.

We now define the variables. The generation expansion decisions are binary variables, meaning that the optimal decisions are selected among the pre-defined expansion options. The binary variable vector $\mathbf{y} \in \{0,1\}^{G^C}$ determines which candidate units should be built. As operational stage variables, variable vector $\mathbf{x}_{srt} = [x_{1srt}, ..., x_{isrt}, ..., x_{Gsrt}] \in \{0,1\}^G$ indicates the on/off commitment status of generating units. Similarly, variable vector $\mathbf{u}_{srt} = [u_{1srt}, ..., u_{isrt}, ..., u_{Gsrt}] \in \{0,1\}^G$ gives the start-up status of units. The production level of units are given by variable vector $\mathbf{p}_{srt}(\boldsymbol{\gamma}) = [p_{1srt}(\boldsymbol{\gamma}), ..., p_{isrt}(\boldsymbol{\gamma}), ..., p_{Gsrt}(\boldsymbol{\gamma})] \in \mathbb{R}^G$. Note that the commitment and start-up variable vectors $\mathbf{x}_{srt}$ and $\mathbf{u}_{srt}$ are dependent on long-term uncertainty, but are independent of short-term uncertainty, i.e., the corresponding decisions are made before the time that the short-term uncertainty realizes. However, production variable vector $\mathbf{p}_{srt}(\boldsymbol{\gamma})$ is dependent on both long- and short-term uncertainties. Therefore, the units can alter their production level in real-time to contribute to coping with the imbalances caused by power production from renewables.

B. Formulation

The proposed DRCC model includes (2) to (5). The objective function seeks to minimize the total system cost, including expansion cost plus the worst expected operations cost, i.e.,

$$\min_{\mathbf{y}} \ \mathbf{k}^\top \mathbf{y} + \sum_{srt} \pi_s \max_{D \in \mathcal{P}_{srt}} \kappa_r \ Q_s(\mathbf{y}), \tag{2}$$

where the first term, i.e., $\mathbf{k}^\top \mathbf{y}$, is the annualized expansion cost, while the second term is the annualized expected operations cost under the worst-case probability distribution $D \in \mathcal{P}_{srt}$. Note that π_s is the probability of long-term scenario s, whereas κ_r is the weight of representative day r. Finally, $Q_s(\mathbf{y})$ represents the optimal annualized expected operations cost under probability distribution D for given investment decision $\mathbf{y}$ and long-term scenario s. This cost includes production and start-up costs of the units, and writes as

$$Q_s(\mathbf{y}) = \min_{\mathbf{p},\mathbf{x},\mathbf{u}} \mathbb{E}^D[\mathbf{c}^\top \mathbf{p}_{srt}(\boldsymbol{\gamma}) + \mathbf{h}^\top \mathbf{u}_{srt}], \tag{3}$$

where $\mathbb{E}^D[.]$ is the expectation operator. Note that the objective function is indeed a three-stage min-max-min problem, which will be recast as a single-stage problem by using linear decision rules – it will be explained later in Section IV and Appendix B. The first set of constraints is

$$x_{isrt} \leqslant \mathbf{y}_i, \quad \forall i \in G^C, s, r, t \tag{4a}$$

$$-\mathbf{x}_{sr(t-1)} + \mathbf{x}_{srt} - \mathbf{x}_{sr\tau} \leqslant \mathbf{0}, \quad \forall \tau \in \{t, ..., \overline{\mathbf{v}} + t - 1\}, \forall s, r, t \tag{4b}$$

$$\mathbf{x}_{sr(t-1)} - \mathbf{x}_{srt} + \mathbf{x}_{sr\tau} \leqslant \mathbf{1}, \quad \forall \tau \in \{t, ..., \underline{\mathbf{v}} + t - 1\}, \forall s, r, t \tag{4c}$$

$$-\mathbf{x}_{sr(t-1)} + \mathbf{x}_{srt} - \mathbf{u}_{srt} \leqslant \mathbf{0}, \quad \forall s, r, t \tag{4d}$$

$$\mathbf{1}^\top \mathbf{p}_{srt}(\boldsymbol{\gamma}) + \mathbf{1}^\top (\mathbf{m}_{srt} + \boldsymbol{\gamma}) = \mathbf{1}^\top \mathbf{d}_{srt}, \ \mathbb{P}-\text{a.s.} \ \forall s, r, t \tag{4e}$$

$$\mathbf{y} \in \{0,1\} \tag{4f}$$

$$\mathbf{x}_{srt}, \mathbf{u}_{srt} \in \{0,1\}, \quad \forall s, r, t. \tag{4g}$$

Constraint (4a) enforces the commitment status of candidate unit $i \in G^C$ to be zero if it is not selected to be built. Constraints (4b) and (4c) impose the minimum up- and down-time limits of units, whereas (4d) represents the units' state transition. Constraint (4e) enforces the power balance almost surely with probability 1 for any realization of uncertainty [40]. Although this constraint is dependent on short-term uncertainty, it is treated as the most important operational constraint of the problem, meaning that it should be respected under any potential realization. Finally, (4f) and (4g) declare the integrality conditions.

In addition to (4), the DRCC problem is constrained by a set of chance constraints. Unlike operational constraint (4e), the capacity and ramping limits of generating units as well as the capacity limits of transmission lines are imposed through individual chance constraints. This allows adjusting the conservativeness of the distributionally robust optimization. For example, the capacity constraint of unit i is enforced by

$$\min_{D \in \mathcal{P}_{srt}} \mathbb{P}[p_{isrt}(\boldsymbol{\gamma}) \leqslant \overline{p}_i x_{isrt}] \geqslant 1 - \epsilon_i, \quad \forall i, s, r, t, \tag{5a}$$

where $\mathbb{P}[.]$ is the probability operator. This specific chance constraint implies that under all distributions in ambiguity set

$\mathcal{P}_{srt}$, the probability of meeting the capacity constraint should be greater than or equal to $1-\epsilon_i$, where, parameter ϵ_i lies within zero and one. Similarly, the rest of individual chance constraints write as

$$\min_{D\in\mathcal{P}_{srt}} \mathbb{P}[p_{isrt}(\boldsymbol{\gamma}) \geqslant \underline{p}_i x_{isrt}] \geqslant 1-\epsilon_i, \quad \forall i,s,r,t \tag{5b}$$

$$\min_{D\in\mathcal{P}_{srt}} \mathbb{P}[p_{isrt}(\boldsymbol{\gamma}) - p_{isr(t-1)}(\boldsymbol{\gamma}) \leqslant \overline{r}_i\ x_{isr(t-1)} + r_i(1-x_{isr(t-1)})] \geqslant 1-\epsilon_i, \ \ \forall i,s,r,t \tag{5c}$$

$$\min_{D\in\mathcal{P}_{srt}} \mathbb{P}[p_{isr(t-1)}(\boldsymbol{\gamma}) - p_{isrt}(\boldsymbol{\gamma}) \leqslant \underline{r}_i\ x_{isrt} + r_i(1-x_{isrt})] \geqslant 1-\epsilon_i, \ \ \forall i,s,r,t \tag{5d}$$

$$\min_{D\in\mathcal{P}_{srt}} \mathbb{P}[\mathbf{H}_l^{\mathrm{G}}\mathbf{p}_{srt}(\boldsymbol{\gamma}) + \mathbf{H}_l^{\mathrm{W}}(\mathbf{m}_{srt}+\boldsymbol{\gamma}) - \mathbf{H}_l^{\mathrm{D}}\mathbf{d}_{srt} \leqslant \overline{f}_l] \geqslant 1-\epsilon_l, \ \ \forall l,s,r,t \tag{5e}$$

$$\min_{D\in\mathcal{P}_{srt}} \mathbb{P}[\mathbf{H}_l^{\mathrm{G}}\mathbf{p}_{srt}(\boldsymbol{\gamma}) + \mathbf{H}_l^{\mathrm{W}}(\mathbf{m}_{srt}+\boldsymbol{\gamma}) - \mathbf{H}_l^{\mathrm{D}}\mathbf{d}_{srt} \geqslant -\overline{f}_l] \geqslant 1-\epsilon_l, \ \ \forall l,s,r,t. \tag{5f}$$

Chance constraint (5b) imposes the minimum production level of units, while (5c) and (5d) restrict the ramping of units. Finally, (5e) and (5f) enforce the capacity of each transmission line l, where $\mathbb{E}^D[.]$ is the expectation operator. The next section explains how to solve the proposed DRCC problem (2)-(5).

IV. Solution Methodology

The proposed DRCC model (2)-(5) is computationally hard to solve (or even intractable) due to the infinite-dimensional nature of the problem and the existence of binary variables. For tractability, we solve it in linear decision rules [40], and use a tight relaxation approach to convexify the UC constraints [41]. We first introduce an approximation approach to convexify the UC constraints, and then follow an affine policy strategy to analytically reformulate the chance constraints.

A. Tight Relaxation of Unit Commitment Constraints

The expansion decisions $\mathbf{y}$ are the main decision variables of the proposed model, and therefore, their integrality conditions need to be preserved. On the contrary, it is computationally appealing to relax the integrality conditions of UC constraints in operational stage, but as tight as possible. The binary variables of the UC constraints are on/off commitment status $\mathbf{x}_{srt}$ and start-up status $\mathbf{u}_{srt}$ of generating units. To enhance the tractability, we use a tight convex relaxation approach, similar to the ones in [10] and [41], which substitutes the feasible set of each generating unit by a tractable approximation of its convex hull[7]. In this approach, we first relax each operational-stage binary variable to lie within zero and one, i.e.,

$$\mathbf{0} \leqslant \mathbf{x}_{srt} \leqslant \mathbf{1}; \quad \mathbf{0} \leqslant \mathbf{u}_{srt} \leqslant \mathbf{1}, \quad \forall s,r,t. \tag{6}$$

Then, additional inequalities are embedded to tighten the relaxation [43]. Similar to the ramping and capacity constraints of generating units in (5), we consider these additional tightening constraints in form of chance constraints:

$$\min_{D\in\mathcal{P}_{srt}} \mathbb{P}[p_{isr(t-1)}(\boldsymbol{\gamma}) \leqslant r_i\ x_{isr(t-1)} + (\overline{p}_i - r_i)(x_{isrt} - u_{isrt}))] \geqslant 1-\epsilon_i, \forall i,s,r,t \tag{7a}$$

$$\min_{D\in\mathcal{P}_{srt}} \mathbb{P}[p_{isrt}(\boldsymbol{\gamma}) \leqslant \overline{p}_i x_{isrt} - (\overline{p}_i - r_i)u_{isrt})] \geqslant 1-\epsilon_i, \quad \forall i,s,r,t \tag{7b}$$

$$\min_{D\in\mathcal{P}_{srt}} \mathbb{P}[p_{isrt}(\boldsymbol{\gamma}) - p_{isr(t-1)}(\boldsymbol{\gamma}) \leqslant (\underline{p}_i + \overline{r}_i)x_{isrt} - \underline{p}_i x_{isr(t-1)} - (\underline{p}_i + \overline{r}_i - r_i)u_{isrt}] \geqslant 1-\epsilon_i, \forall i,s,r,t \tag{7c}$$

$$\min_{D\in\mathcal{P}_{srt}} \mathbb{P}[p_{isrt}(\boldsymbol{\gamma}) - p_{isr(t-1)}(\boldsymbol{\gamma}) \geqslant (\underline{p}_i + \underline{r}_i - r_i)u_{isrt} - r_i x_{isr(t-1)} + (r_i - \underline{r}_i)x_{isrt}] \geqslant 1-\epsilon_i, \forall i,s,r,t. \tag{7d}$$

As a result, the proposed formulation for GEP can now be stated as the collection of (2)-(3), (4a)-(4f), (5)-(7).

B. Linear Decision Rules and Resulting Reformulation

To further mitigate the complexity of the proposed problem, the recourse actions are approximated with linear decision rules [40]. In this way, the production of each unit i, i.e., $p_{isrt}(\boldsymbol{\gamma})$ is described by $p_{isrt} + \alpha_{isrt}\mathbf{1}^\top\boldsymbol{\gamma}$, where p_{isrt} is the tentative day-ahead schedule of unit i, while recourse action $\alpha_{isrt}1^\top\boldsymbol{\gamma}$ represents the linear change of power produced by unit i in response to the uncertain renewable production. In particular, we refer to variable α_{isrt} as *participation factor* of unit i (in per-unit) to cover the power imbalances due to forecast errors. Using this affine policy, and assuming that $\boldsymbol{\mu}_{srt} = 0, \ \forall s,r,t$, objective function (2), power balance constraints (4e) and chance constraints (5) and (7) can be reformulated to make the original problem tractable. The details of these reformulations are given in Appendix B.

It is worth noting that ramping chance constraints (5c)-(5d) and (7c)-(7d) impose inter-temporal coupling between dispatch decisions. For these constraints, the uncertainty parameter vector contains both vectors $\boldsymbol{\gamma}_{srt}$ and $\boldsymbol{\gamma}_{sr(t-1)}$ such that

$$\widehat{\boldsymbol{\gamma}}_{srt} = \begin{bmatrix} \boldsymbol{\gamma}_{srt} \\ \boldsymbol{\gamma}_{sr(t-1)} \end{bmatrix}, \tag{8}$$

where $\widehat{\boldsymbol{\gamma}}_{srt} \in \mathbb{R}^{2Z}$. As a result, another covariance matrix needs to be defined to model the effect of temporal correlation on ramping constraints that can be written as

$$\widehat{\boldsymbol{\Sigma}}_{srt} = \begin{bmatrix} \boldsymbol{\Sigma}_{srt} & \boldsymbol{\Upsilon}_{sr(t,t-1)} \\ \boldsymbol{\Upsilon}_{sr(t,t-1)} & \boldsymbol{\Sigma}_{sr(t-1)} \end{bmatrix}, \tag{9}$$

where $\widehat{\boldsymbol{\Sigma}}_{srt} \in \mathbb{R}^{2Z\times 2Z}$ includes not only the spatial correlation of wind farms, but also their temporal correlation between hours t and $t-1$. Now, the original DRCC problem boils down to a mixed-integer second-order cone program as given in detail in Appendix C.

V. Numerical Study

We use a case study based on a modified version of the IEEE 118-bus test system [44] including 19 existing conventional units, 2 existing wind farms, 99 demands, and 186 transmission lines. By existing, we refer to those conventional or wind

[7] We do not investigate the impacts of this relaxation on pricing, but refer the interested readers to [42].

power units that either exist now and will not be phased out by the target year, or the units that do not exist now, but it is certain that they will be built by the target year. In addition to the existing conventional and wind units, we consider 22 candidate conventional units that can be built.

All simulations are run on an Intel(R) Xeon(R) E5-1650 with 12 processors clocking at 3.50 GHz and 32 GB of RAM. The source code implemented in Matlab using YALMIP and solved by Gurobi 7.5.1 is available in the online companion [45]. Depending on the values assigned for ϵ in chance constraints, the CPU time varies from 2 to 3 hours.

A. Input Data

We consider several types of conventional units, including coal, nuclear, combined cycle gas turbine (CCGT) and gas turbine (GT). Their technical parameters are given in Table II. The total capacity of existing conventional units is 6,466 MW. The full details about the characteristics of generating units, transmission lines, load profile of demands, and wind power profile of two farms are given in the online companion [45].

Using a K-means clustering technique, we provide net-load data for 10 days, each with a different profile, to represent the target year. The reason for selecting 10 representative days is that with increasing the number of these days, the investment solution does not change significantly.

Without loss of generality, two long-term scenarios with the same probabilities are considered[8]. It is assumed that the load level of the target year in the first and second scenarios is 3.4 and 4 times of the existing load level with the same pattern, respectively. The capacity of two wind farms in the target year is assumed to be 2,400 MW and 3,000 MW under the first long-term scenario, while for the second scenario, it is assumed to be 5,355 MW and 6,120 MW. Therefore, the wind penetration level, i.e., the yearly contribution of wind power to supply demand, under both long-term scenarios is 35%.

For each farm, we generate 10,000 power trajectories from the historical data, each containing the wind power profile of that farm over all hours of the 10 representative days. We then split them into two sets of 5,000 trajectories; the first set to form the training dataset (in-sample data) and the second set to be used as test data in an ex-post out-of-sample analysis.

B. Constructing the Ambiguity Sets

Using the 5,000 training trajectories, it is required to calculate the mean production vectors of wind farms $\mathbf{m}_{srt}$ and the covariance matrices of wind power forecast errors Σ_{srt} and $\hat{\Sigma}_{srt}$ for each long-term scenario, representative day and hour. For example, Fig. 1 illustrates the training data in per-unit used for constructing the covariance matrix $\hat{\Sigma}_{srt}$ in a sample representative day under both long-term scenarios. This figure illustrates a matrix of plots in which the diagonal plots display the histograms of production forecast errors pertaining to wind farms $w1$ and $w2$ at hours t and $t-1$. Besides, the scatter plots of corresponding forecast errors appear in the off-diagonal. The histogram plots show the frequency distribution of each forecast error, while the scatter plots show the correlation of two corresponding forecast errors. The scatter plots in blue and green, respectively, show temporal and spatial correlations only. However, the scatter plots in red show the combination of temporal and spatial correlations between hours t and $t-1$ and between wind farms $w1$ and $w2$. From the histogram plots, it is evident that the forecast error of wind production does not necessarily follow a specific probability distribution. Furthermore, it can be seen that the mean production of forecast errors is zero. Therefore, these observations match with the assumption that forecast errors have unknown distribution with a zero mean value.

[8]We will explore later in Section V.F the robustness of the proposed model against the long-term uncertainty. We observe that increasing the number of long-term scenarios significantly increases the CPU time, such that the same case study with 4 long-term scenarios is solved in around 6 to 7 hours, depending on the value of confidence level. To tackle this computational issue, one may consider using decomposition techniques. One potential solution is to use Benders' decomposition by treating the expansion decisions as complicating (coupling) variables, e.g., in a similar way as in [11]. However, we leave it for future extensions as it is not the main focus of the current paper.

TABLE II
DATA FOR CONVENTIONAL GENERATING UNITS

Type	Coal	Nuclear	CCGT	GT1	GT2
Oper. cost ($/MWh)	18.72	10.33	19.32	38.47	48.47
Start-up cost ($/MW)	54.11	100	16.23	25.14	28.14
Exp. cost ($10^3$$/MW-year)	174.7	224	60.1	48.5	38.5
Min. generation (p.u.)†	0.75	0.9	0.4	0.27	0.25
Ramp rate (p.u.)†	0.3	0.1	0.5	0.7	1
Min. up time (h)	24	48	6	1	1
Min. down time (h)	12	24	12	1	1
Nr. of existing units	4	2	4	2	7
Nr. of candidate units	4	2	6	4	6

†It is given as a percentage of unit's capacity.

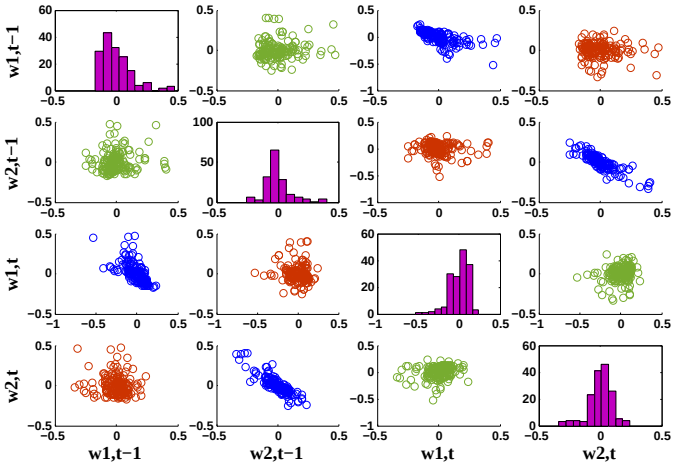

Fig. 1. Production forecast errors of wind farms $w1$ and $w2$ at hours t and $t-1$ in a sample representative day. The diagonal plots (x-axis: in per-unit; y-axis: frequency of occurrence) display the frequency of forecast error realizations. The off-diagonal plots (both axes in per-unit) display the realizations of two different corresponding forecast errors.

C. Benchmark Models

To assess the performance of the proposed DRCC model, we consider a chance-constrained (CC) model as a benchmark. In this benchmark, the short-term uncertainty follows a Gaussian probability distribution with the same two moments used in

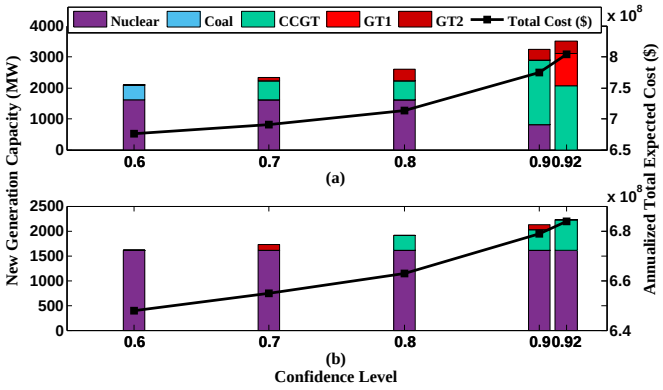

Fig. 2. In-sample results: New generation capacities to be built (left vertical axis) and annualized expected total system cost (right vertical axis) as a function of confidence level, obtained from the proposed DRCC model (upper plot) and the chance-constrained benchmark model with Gaussian probability distribution (lower plot).

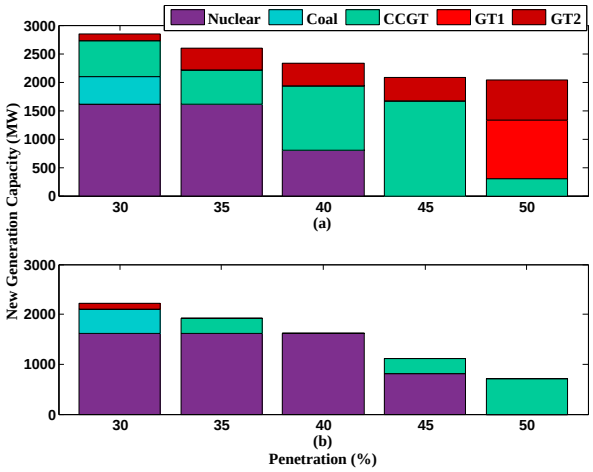

Fig. 3. In-sample results: New conventional generation capacities to be built under different values of wind power penetrations, obtained from the proposed DRCC model (upper plot) and the chance-constrained benchmark model with Gaussian probability distribution (lower plot). The confidence level is fixed to 0.8.

the DRCC model. Therefore, there is no ambiguity set in the benchmark, and it is not necessarily robust against the short-term uncertainty if it follows any other distribution. Similar to the proposed DRCC model, the benchmark model results in a mixed-integer second-order cone problem. In fact, each chance constraint with a Gaussian distribution can be analytically reformulated as a second-order cone constraint [46], as explained in Appendix D. Note that we will consider an additional benchmark in Section V.G, which is a chance-constrained optimization with no assumption on probability distribution. The chance constraints in this benchmark can no longer be analytically reformulated. Following a randomized sampling approach in [47], we will solve it as either a robust optimization or a stochastic program.

D. In-Sample Results

We solve the proposed DRCC model to determine the optimal generation expansion plans. For simplicity, we consider an identical value for ϵ_i and ϵ_l for all chance constraints, denoted as ϵ, and then refer to $1-\epsilon$ as *confidence level*.

Fig. 2 illustrates the optimal generation capacity expansion plans (left vertical axis) and the annualized expected total system cost (right vertical axis) as a function of confidence level obtained from the proposed DRCC model (upper plot) and the benchmark CC model with Gaussian distribution (lower plot). Note that the DRCC model becomes infeasible for a confidence level higher than a value between 0.95 and 0.96, meaning that all candidate units, even if are being built, cannot provide enough operational flexibility under the worst probability distribution – this implies that load curtailment will be required. In both plots, as expected, the system cost grows up with increasing confidence level, as more flexible but more expensive conventional units are required to be built. For example, the proposed DRCC model suggests investing in nuclear and coal units if confidence level is equal to 0.6, but it drastically changes in a case with confidence level of 0.92, where this model determines flexible CCGT and GT units as optimal generation mix to be built. In addition, a comparison of upper and lower plots in Fig. 2 points out that the proposed DRCC model is more conservative than the benchmark, ending up to a higher system cost and a higher investment in flexible generation capacity. However, we will demonstrate later that the proposed model has a much better out-of-sample performance with respect to the benchmark.

In order to demonstrate the effectiveness of the proposed approach in modeling the variability and uncertainty of stochastic energy sources, a sensitivity analysis with respect to the level of wind power penetration is performed. To do so, we increase the wind penetration level in both long-term scenarios from 30% to 50%. The upper plot of Fig. 3 depicts the results of the proposed generation expansion model for different penetration levels. It is observed that with increasing the penetration level, a lower conventional generation capacity is required, but as expected, this new capacity is chosen among flexible technologies. This is due to the fact that a higher wind penetration implies a lower net-load, but higher variability and uncertainty in the system. As a result, the wind power penetration level not only affects the capacity of new generating units, but also changes the type of units. The lower plot of Fig. 3 shows the results obtained by the benchmark. Compared to the solutions of the DRCC model, the benchmark model fails in capturing the need for new flexible generating units when the wind penetration grows.

E. Out-of-Sample Analysis

Using the remaining 5,000 wind trajectories described in Section V.A, we perform an out-of-sample analysis in this section to evaluate the performance of the proposed DRCC model with respect to the benchmark. The main point is that we fix the generation expansion decisions to those obtained in the in-sample study (illustrated in Fig. 2), and then solve the tightened relaxed unit commitment problem over representative days but using the out-of-sample wind data.

For the out-of-sample simulation, the 5,000 test trajectories are clustered using a k-means technique to achieve the same number of representative days to the one in the in-sample study, i.e., 10. However, the weighting factors of the 10

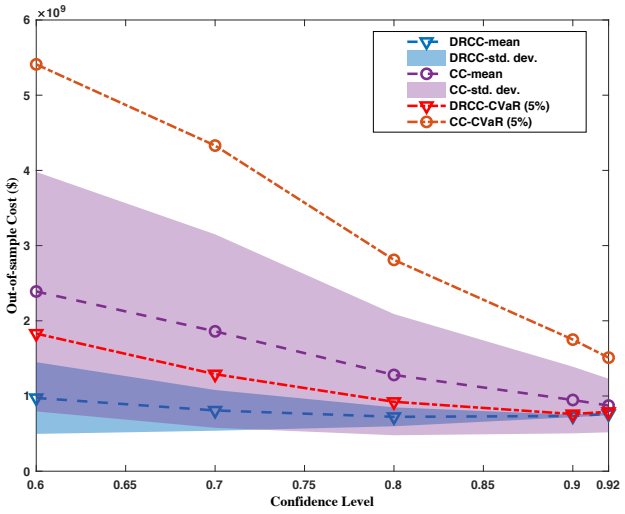

Fig. 4. Out-of-sample results: The annualized total system cost in terms of mean, standard deviation, and 5%-CVaR obtained from the proposed DRCC model and the CC benchmark model with Gaussian probability distribution.

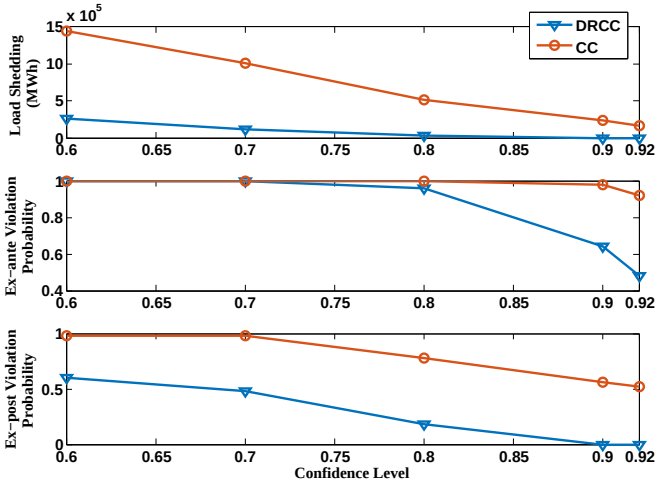

Fig. 5. Out-of-sample results: The annual amount of load curtailment (upper plot), ex-ante violation probability (middle plot), and ex-post violation probability (lower plot) obtained from the proposed DRCC model and the CC benchmark model with Gaussian probability distribution.

representative days in in-sample and out-of-sample studies might be different. After that, we randomly pick one trajectory from each cluster (i.e., representative day), such that the set of 10 selected trajectories represents the wind profile in the target year. Then, we solve the tightened relaxed unit commitment problem for the whole target year, while taking into account the extreme recourse actions, i.e., load shedding and wind spillage. We consider a value of lost load equal to \$1,000/MWh, while the wind spillage is cost free. Note that each unit commitment problem to be solved in the out-of-sample simulations is indeed a deterministic problem. We repeat this procedure until there is no trajectory left for at least one of the clusters.

For both proposed DRCC model and the CC benchmark model with Gaussian distribution, Fig. 4 shows the out-of-sample system cost (in terms of mean and standard deviation) as a function of confidence level. The range of system cost between mean plus/minus standard deviation is illustrated as a tube around the mean cost curve. In addition, this figure depicts the average system cost under the worst 5% out-of-sample simulations, measured by a conditional value-at-risk (CVaR) metric. Let us first analyze the DRCC results. With increasing values for confidence level from 0.6 to around 0.8, the mean cost of the DRCC model (blue dashed line) decreases, while it increases afterwards. The reason for this is that there is a relatively significant load curtailment when the confidence level is low, while it becomes zero with a higher confidence level, but at the cost of higher investment and operations costs. The trend of annual amount of load curtailment in DRCC model as a function of confidence level is given in the upper plot of Fig. 5. In addition, Fig. 4 shows that both standard deviation and 5%-CVaR of the system cost obtained from the proposed DRCC model decrease with increasing the confidence level. Compared to this model, the CC model (benchmark) shows a poor out-of-sample performance, such that the mean cost, standard deviation, 5%-CVaR, and the amount of load curtailment in that model are always higher than those in the proposed DRCC model. The mean cost of the two models in the highest confidence level in Fig. 4, i.e., 0.92, get closer, but note that such a difference will increase if a higher value for lost load is considered.

For further comparison of the two models, we present in the ex-post and ex-ante violation probabilities of the solutions obtained by the DRCC model and the CC benchmark model with Gaussian distribution as given in the middle and lower plots of Fig. 5, respectively. The ex-post violation probability is indeed achieved by solving the out-of-sample unit commitment problems: whenever there is a need for an extreme recourse action, i.e., load curtailment or wind spillage, we count it as a sample with an ex-post violation. In contrast, we do not solve any problem to compute the ex-ante violation probability: for given optimal in-sample values obtained for the tentative day-ahead schedule and participation factor of generating units obtained, the linear decision rule is applied to calculate the recourse power production of each unit under each test sample. This way, the satisfaction of chance constraints is investigated – see Appendix E for further details. According to Fig. 5, the ex-ante violation probability is higher than the ex-post one due to the fact that the linear decision rule restricts the power production of generating units in response to the short-term uncertainty which increases the violation probability. The key point is that with both metrices, the proposed DRCC model shows a better performance with respect to the benchmark model[9].

F. Additional Analyses

In this section, we first investigate the impacts of UC constraints on planning outcomes. To do so, we compare the results achieved by the proposed model with those of a similar

[9]It is worth mentioning that though the proposed DRCC model exhibits a good out-of-sample performance in this case study in terms of violation probability, the chance constrains in general do not necessarily restrict the *severity* of violations. If it is a concern, one may use the CVaR constraints rather than the chance constrains, but at the cost of more complicated model and potentially a more conservative solution [48].

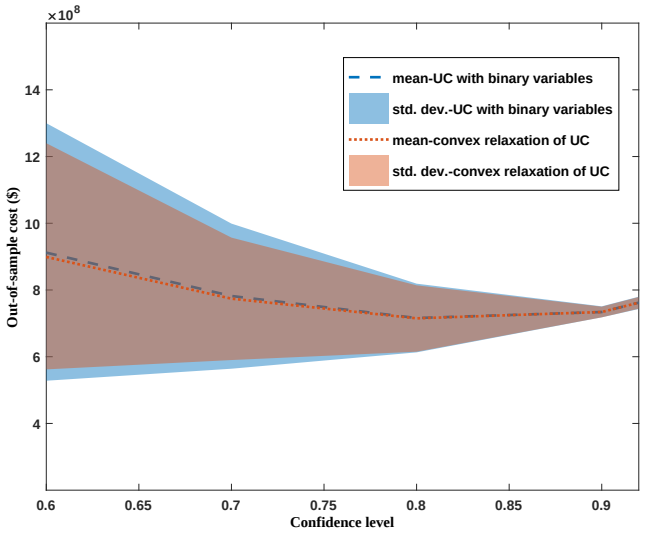

Fig. 6. Out-of-sample results: The annualized total system cost in terms of mean and standard deviation for the two cases, where the operational stage includes either exact or tightened relaxed unit commitment (UC) constraints.

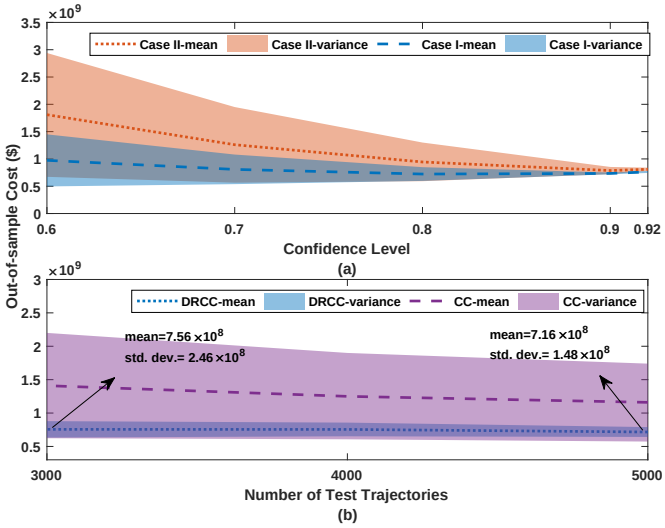

Fig. 7. Out-of-sample results: The annualized total system cost in terms of mean and standard deviation (a): obtained from the proposed DRCC model for the two cases evaluating the robustness of the proposed model against the long-term uncertainty, (b): obtained from the proposed DRCC model and the CC benchmark model with Gaussian probability distribution for different number of training trajectories showing the robustness of the proposed model against the size of training dataset.

model including economic dispatch (ED) limits only, where all UC constraints are neglected. As expected, the ED-based model invests in more inflexible generation capacity including nuclear and coal units. As a result, the out-of-sample system cost in terms of mean and standard deviation obtained from the ED-based model is comparatively higher than that of the UC-based model. For example, in a case with a confidence level of 0.8, the mean and standard deviation of the annualized system cost obtained from the out-of-sample simulation in the ED-based model are $796 and $241 millions, respectively. Those values in the UC-based model are $716 and $103 millions, respectively.

Our next analysis explores the effects of the tight relaxation of UC constraints on out-of-sample system cost. After fixing the expansion decisions to those obtained in in-sample analysis (Fig. 2), we run the deterministic out-of-sample simulations for two cases, where the operational stage includes either exact or tightened relaxed UC constraints. For both cases, Fig. 6 illustrates the out-of-sample cost in terms of mean and standard deviation for different values of confidence level. From this figure, we observe that the out-of-sample cost in case with exact UC constraints is very close to that in case with tightened relaxed UC constraints. Although this analysis does not show the impact of UC relaxation on expansion decisions, it confirms that the convex relaxation will not considerably affect the optimality of the solution obtained in the operational stage.

Another key point to highlight is that the proposed model provides a robustness against the short-term uncertainty, but not necessarily against the long-term one, as it has been modeled by a limited set of scenarios only. This is illustrated in the upper plot of Fig. 7, demonstrating the mean and standard deviation of out-of-sample system cost obtained from the proposed DRCC model for two different cases. In Case I, the two long-term scenarios for demand growth in the out-of-sample simulations are exactly the same as those in the in-sample study. This is indeed a biased out-of-sample simulation in terms of long-term uncertainty. On the contrary, in Case II, we consider a forecast error in long-term uncertainty, so that the realized load level under both scenarios are 5% higher than those in the in-sample long-term scenarios. As observed in the upper plot of Fig. 7, this long-term forecast error causes a remarkable increase in cost (both in terms of mean and standard deviation), especially for cases with comparatively lower confidence level. This assessment highlights the need for more advanced tools, which also provides robustness not only against the short-term uncertainty, but also against the long-term one. We leave this study for our future work.

In another analysis, the robustness of the proposed model is explored as a function of the size of training dataset. We run our simulation with different size of training samples, varying from 3,000 to 5,000. As shown in the lower plot of Fig. 7, the proposed DRCC model is more robust against the sample size compared to the benchmark CC model.

G. Additional Benchmark: A Chance-Constrained Optimization With no Assumption on Distribution

As mentioned earlier in Section V.C, we consider here an extra benchmark, which is again a chance-constrained optimization with a single distribution, but it does not follow any specific type of distribution. Note that it is not a DRCC model, as it considers a single distribution only. The distribution of this benchmark is identical to that of historical data, i.e., 5,000 in-sample trajectories allocated in 10 representative days. Unlike the benchmark in the previous sections with a Gaussian probability distribution, the chance constraints in this benchmark cannot be analytically reformulated. To solve this problem, we use a randomized sampling approach proposed in [47]. Based on this approach, depending on the inputs of the chance-constrained model, e.g., the confidence level,

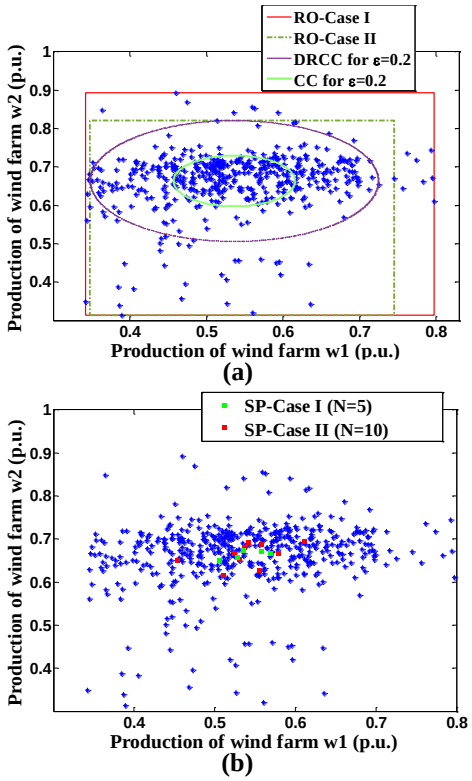

Fig. 8. The in-sample scenarios (blue dots) corresponding to the production of two wind farms ($w1$ and $w2$) in an arbitrarily selected hour, i.e., hour 1 of representative day 2. Besides, plot (a) shows two box uncertainty sets corresponding to RO-Case I and RO-Case II, as well as two ellipsoidal uncertainty sets associated with the DRCC model and the CC model with Gaussian distributions. Finally, plot (b) illustrates the representative scenarios in SP-Case I and SP-Case II with reduced number of scenarios (green and red dots).

we randomly pick a *certain* number of samples from the historical data. In particular, the minimum number of samples, M, required is

$$M \geqslant \frac{1}{1-\epsilon}\frac{e}{e-1}(1+\ln\frac{1}{\beta}), \tag{10}$$

where e is Euler number, $1-\epsilon$ is the confidence level ensuring that the resulting box uncertainty set encloses $1-\beta$ of the probability mass. It is evident that chance constraints with a higher confidence level require more samples. Similarly, the number of random samples required reduces by considering loose chance constraints. Given the samples picked, one can treat the problem as either a stochastic program, or a robust optimization by considering a box (rectangular) uncertainty set covering all samples. The formulation of resulting stochastic program and robust optimization is provided in Appendix F.

We take into account two cases, namely RO-Case I and RO-Case II, for the robust optimization benchmark. In RO-Case I, we pick *all* in-sample trajectories, without using (10), aiming to provide the most robust solution. In RO-Case II, we consider a chance-constrained program with $1-\epsilon = 0.8$ and $\beta = 0.001$ for all chance constraints, and then randomly pick a number of samples according to (10). It is obvious that the number of samples picked in RO-Case II is comparatively lower than that in RO-Case I, as it corresponds to a problem with looser chance constraints. For an arbitrarily selected hour and day, i.e., hour 1 in representative day 2, the samples picked for the two cases, and their corresponding box uncertainty sets are depicted in Fig. 8(a). It is also worth mentioning that there exists a close relation and a complete equivalence between chance-constrained program and robust optimization with an ellipsoidal uncertainty set, as discussed in [49] and [50]. Fig. 8(a) shows the two ellipsoidal uncertainty sets for robust optimizations, which are equivalent to the chance-constrained program with Gaussian distribution (smaller ellipsoid) and the distributionally robust chance-constrained program (larger ellipsoid). One interesting observation is that the robust optimization equivalent to DRCC (i.e., the one with larger ellipsoid) covers less samples than the robust optimization with a smaller box uncertainty set (i.e., RO-Case II)[10]. Therefore, we would expect that the robust optimization corresponding to the chance-constrained benchmark model with no assumption on probability distribution provides a more conservative solution than the proposed DRCC model.

We compare the out-of-sample performance of the two robust cases RO-Case I and RO-Case II with the proposed DRCC model. The upper plot of Fig. 9 illustrates the system cost and the amount of load shedding in these three cases. It is observed that both robust cases even the one corresponding to the benchmark with a lower confidence level (i.e., RO-Case II) provide more conservative solutions than the proposed DRCC model – the system cost is higher, while the amount of load shedding is lower. This is consistent with our earlier observation in Fig. 8(a) that both box uncertainty sets cover more samples than the equivalent robust optimization to DRCC.

We now take the other potential approach with the randomly picked samples obtained from (10), and treat them as scenarios within a two-stage stochastic program, whose first stage determines the expansion decisions, and whose second stage makes the operational decisions under each scenario. Similar to robust optimization approach, we consider two cases, namely SP-Case I and SP-Case II. By allocating the samples within 10 representative days and using a K-means clustering technique, we end up to 10 scenarios in SP-Case I, and 5 scenarios in SP-Case II. Note that each scenario embodies the production profile of two wind farms over all hours of 10 representative days. To make a fair comparison with other approaches, we apply linear decision rules to the two-stage stochastic program too, which recasts it as a single-stage model – see the corresponding formulation in Appendix F. Fig. 8(b) depicts the production of two wind farms under selected scenarios at hour 1 of representative day 2. The lower plot of Fig. 9 illustrates the better out-of-sample performance of the proposed DRCC model compared to both stochastic cases in terms of system cost as well as amount of load shedding.

One important observation for stochastic cases is their computational requirement. The CPU time for the stochastic model with 50 and 100 scenarios is 10 and 17 hours, respectively.

[10]Note that this might not be the case if the randomly selected samples are enveloped by a polyhedron with more than four vertices, taking into account the potential correlation of random variables. The interested readers are referred to [51] for constructing uncertainty sets for robust optimization based on the available samples.

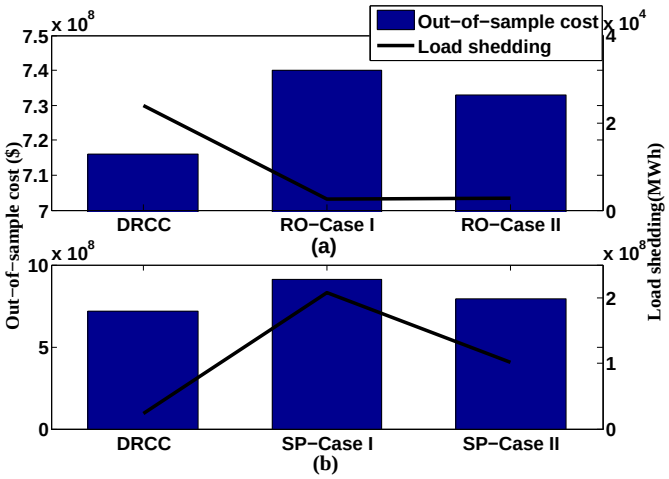

Fig. 9. Out-of-sample results: The annualized expected total system cost and the amount of load shedding obtained from the proposed DRCC model, RO-Case I and RO-Case II (the upper plot), as well as SP-Case I and SP-Case II (the lower plot).

TABLE III
THE COMPARISON OF DIFFERENT MODELS IN TERMS OF COMPUTATIONAL BURDEN

Model	CPU time (h)	Nr. of variables	Nr. of constraints
DRCC	3	78,744	518,080
CC	3	78,744	518,080
RO-Case I	8	1,737,624	2,471,760
RO-Case II	8	1,737,624	2,471,760
SP-Case I	10	78,744	1,845,680
SP-Case II	17	78,744	3,544,540

This is significantly higher than the solution time in the proposed DRCC model (between 2 to 3 hours depending on the value of confidence level). Our numerical analysis implies that the proposed DRCC model presents a better out-of-sample solution with comparatively lower computational burden. It is obvious that by increasing the number of scenarios, the stochastic optimization will exhibit a better out-of-sample performance, but at the cost of increased computational burden. Table III summarizes the computational complexity of the proposed model and the benchmarks which depicts the promising performance of the DRCC model.

VI. CONCLUSION

This paper proposes a distributionally robust chance-constrained generation expansion model, which considers the short-term uncertainty using a family of probability distributions, and characterizes the long-term uncertainty by a set of scenarios. A tight relaxed version of operational unit commitment constraints is incorporated into the model. The resulting model is a mixed-integer second-order cone problem, which is computationally tractable, and demonstrates a better out-of-sample performance compared to a benchmark which considers a specific uncertainty distribution only. The future work needs to robustify the model against the long-term uncertainty too. An additional uncertainty on moments can also be considered, resulting in a distributionally robust model with inexact moments. It is also of interest to explore other possibilities to build the ambiguity sets, e.g., using a Wasserstein metric.

APPENDIX A: NOMENCLATURE

A. Indices

i Index for conventional generating units.
l Index for transmission lines.
r Index for representative days.
s Index for long-term scenarios.
t Index for operating hours.

C. Parameters

κ_r Weight of representative day r [day].
$\mathbf{c}$ Production cost of generating units [$/MWh].
$\mathbf{d}_{srt}$ Load level in short run in representative day r, hour t under long-term scenario s [MW].
$\mathbf{h}$ Start-up cost of generating units [$].
$\mathbf{k}$ Annualized capital cost of candidate units [$].
$\mathbf{m}_{srt}$ Mean vector of production of uncertain sources in representative day r, hour t under long-term scenario s [MW].
$\overline{\mathbf{v}}$,$\underline{\mathbf{v}}$ Minimum up- and down-time of generating units [hour].
$\overline{f}_l$ Capacity of transmission line l [MW].
$\overline{p}_i$,$\underline{p}_i$ Maximum and minimum production level of generating unit i [MW].
$\overline{r}_i$,$\underline{r}_i$ Maximum ramp-up and ramp-down capability of generating unit i [MW/h].
π_s Probability of long-term scenarios s.
D Worst probability distribution.
r_i Ramp rate limit of generating unit i in the time of start-up and shut-down [MW/h].

B. Variables

α_{isrt} Participation factor of unit i in response to short-term uncertainty in representative day r, hour t under long-term scenario s [per-unit].
$\mathbf{H}^{[.]}$ Power transfer distribution factor mapping generating units (G), uncertain sources (W), and demand (D) on network.
$\mathbf{y}$ Investment decision [0,1] of candidate generating units.
p_{isrt} Tentative schedule of unit i in representative day r, hour t under long-term scenario s [MW].
u_{isrt} Start-up status [0,1] of unit i in representative day r, hour t under long-term scenario s.
x_{isrt} On/off commitment status [0,1] of generating unit i in representative day r, hour t under long-term scenario s.

D. Uncertainty modeling

$\mathcal{P}_{srt}$ Ambiguity set associated with long-term scenario s, representative day r, and hour t.
$\boldsymbol{\gamma}_{srt}$ Random variable of forecast error at hour t of representative day r under long-term scenario s.
$\boldsymbol{\mu}_{srt}$ Mean vector of forecast error at hour t of representative day r under long-term scenario s.
$\boldsymbol{\Sigma}_{srt}$ Covariance matrix of forecast error at hour t of representative day r under long-term scenario s.
Ψ_{srt} Family of potential distributions associated with long-term scenario s, representative day r, and hour t.

APPENDIX B: REFORMULATION

In the following three sections of this appendix, we elaborate on the reformulations of objective function (2), power balance constraints (4e), and chance constraints (5) and (7), respectively. Recall that we use linear decision rules in this paper, and thereby, we substitute power production of each unit i, i.e., $p_{isrt}(\boldsymbol{\gamma})$ with $p_{isrt} + \alpha_{isrt}\mathbf{1}^\top\boldsymbol{\gamma},\ \forall i,s,r,t$.

A. Objective Function

By implementing linear decisions rules, three-stage objective function (2) is written as

$$\min_{\mathbf{y}} \mathbf{k}^\top\mathbf{y} + \sum_{srt}\pi_s \max_{D\in\mathcal{P}_{srt}} \kappa_r \min_{\mathbf{p},\alpha,\mathbf{x},\mathbf{u}} \mathbb{E}^D[\mathbf{c}^\top(\mathbf{p}_{srt} + \alpha_{srt}\mathbf{1}^\top\boldsymbol{\gamma}) + \mathbf{h}^\top\mathbf{u}_{srt}]. \tag{11a}$$

Recall that the mean of the forecast error uncertainty vector $\boldsymbol{\gamma}$ is assumed to be zero, i.e., $\mathbb{E}^D(\boldsymbol{\gamma}) = \boldsymbol{\mu} = 0$. Therefore, the second row of (11a) reduces to $\mathbf{c}^\top\mathbf{p}_{srt} + \mathbf{h}^\top\mathbf{u}_{srt}$. Now, the forecast error $\boldsymbol{\gamma}$ does no longer exist in the objective function. Similarly, constraints (4e), (5) and (7) will be reformulated in the next two sections of this appendix in a way that they will not include $\boldsymbol{\gamma}$. As a result, maximization operator $\max_{D\in\mathcal{P}_{srt}}$ in min-max-min objective function (11a) can be removed [26], yielding a min-min objective function, which can be combined to a single minimization function as

$$\min_{\mathbf{y},\mathbf{p},\alpha,\mathbf{x},\mathbf{u}} \mathbf{k}^\top\mathbf{y} + \sum_{s,r,t}\pi_s\kappa_r(\mathbf{c}^\top\mathbf{p}_{srt} + \mathbf{h}^\top\mathbf{u}_{srt}). \tag{11b}$$

B. Power Balance Constraints

The linear decisions rules reformulate the power balance constraints (4e) to

$$\mathbf{1}^\top(\mathbf{p} + \boldsymbol{\alpha}_{srt}\mathbf{1}^\top\boldsymbol{\gamma}) + \mathbf{1}^\top(\mathbf{m}_{srt} + \boldsymbol{\gamma}) = \mathbf{1}^\top\mathbf{d}_{srt},\ \ \forall s,r,t. \tag{12a}$$

In order to satisfy (12a) for all realizations of the forecast error $\boldsymbol{\gamma}$, we set the first-order coefficient of $\boldsymbol{\gamma}$ equal to zero [40]. Thus, by matching the zero- and first-order coefficients of $\boldsymbol{\gamma}$ on both sides of (12a), we obtain

$$\mathbf{1}^\top\boldsymbol{\alpha}_{srt} = -\mathbf{1}, \quad \forall s,r,t \tag{12b}$$

$$\mathbf{1}^\top\mathbf{p}_{srt} + \mathbf{1}^\top\mathbf{m}_{srt} = \mathbf{1}^\top\mathbf{d}_{srt}, \quad \forall s,r,t. \tag{12c}$$

C. Chance Constraints

By applying Chebyshev inequality to a distributionally robust individual chance constraint as explained in [52] and [38], we can analytically reformulate chance constraints (5) and (7). However, one may use this reformulation with caution, as the accuracy of Chebyshev approximation might be reduced when the value of confidence level $1-\epsilon$ is very close to one. It may end up to a unnecessarily conservative or even infeasible solution[11] [26], [53].

[11]To tackle this issue, a more complicated but exact reformulation is proposed in [26], which is not implemented in this paper. Reference [53] provides a comparison of these two approximate and exact reformulations by applying them to a market-clearing problem.

For example, we provide here the reformulation of chance constraint (5a). Other chance constraints can be analytically reformulated in the same manner. By implementing the linear decision rules, (5a) is written as

$$\min_{D\in\mathcal{P}_{srt}} \mathbb{P}\left(p_{isrt} + \alpha_{isrt}\mathbf{1}^\top\boldsymbol{\gamma} \leqslant \overline{p}_i x_{isrt}\right) \geqslant 1-\epsilon_i, \quad \forall i,s,r,t. \tag{13a}$$

Assuming that $\boldsymbol{\mu}_{srt} = 0,\ \forall s,r,t$, and following the approach in [52] and [38] to implement Chebyshev inequality, the above chance constraint is analytically reformulated as

$$p_{isrt} \leqslant \overline{p}_i x_{isrt} - \sqrt{\frac{1-\epsilon_i}{\epsilon_i}}\sqrt{\alpha_{isrt}\mathbf{1}^\top\boldsymbol{\Sigma}_{srt}\mathbf{1}\alpha_{isrt}}, \quad \forall i,s,r,t, \tag{13b}$$

which can be eventually rewritten as

$$p_{isrt} \leqslant \overline{p}_i x_{isrt} - \sqrt{\frac{1-\epsilon_i}{\epsilon_i}}\|\alpha_{isrt}\mathbf{1}^\top\boldsymbol{\Sigma}_{srt}^{1/2}\|_2, \ \ \forall i,s,r,t. \tag{13c}$$

Note that the resulting reformulation (13c) is a second-order cone constraint. It can be reduced to a linear constraint by assuming $\alpha_{isrt} \geqslant 0$. However, it is a strong assumption, especially in problems with network and unit commitment constraints, and we avoid taking such an assumption. Therefore, each chance constraint (5) and (7) is reformulated to a second-order cone constraint.

APPENDIX C: RESULTING MODEL

According to all reformulations explained in Appendix B, the resulting mixed-integer second-order cone problem for the proposed DRCC model (2)-(5) is

$$\min_{\mathbf{y},\mathbf{p},\alpha,\mathbf{x},\mathbf{u}} \mathbf{k}^\top\mathbf{y} + \sum_{s,r,t}\pi_s\kappa_r(\mathbf{c}^\top\mathbf{p}_{srt} + \mathbf{h}^\top\mathbf{u}_{srt}) \tag{14a}$$

subject to:

(4a)-(4d),(4f),(6),(12b)-(12c),(13c)

$$p_{isrt} \geqslant \underline{p}_i x_{isrt} + \sqrt{\frac{1-\epsilon_i}{\epsilon_i}}\|\alpha_{isrt}\mathbf{1}^\top\boldsymbol{\Sigma}_{srt}^{1/2}\|_2, \ \ \forall i,s,r,t \tag{14b}$$

$$p_{isrt} - p_{isr(t-1)} \leqslant \overline{r}_i x_{isr(t-1)} + r_i(1 - x_{isr(t-1)}) - \sqrt{\frac{1-\epsilon_i}{\epsilon_i}}\|\mathbf{v}_{isrt}^\top\widehat{\boldsymbol{\Sigma}}_{srt}^{1/2}\|_2, \ \ \forall i,s,r,t \tag{14c}$$

$$p_{isrt} - p_{isr(t-1)} \geqslant -\underline{r}_i x_{isrt} - r_i(1 - x_{isrt}) + \sqrt{\frac{1-\epsilon_i}{\epsilon_i}}\|\mathbf{v}_{isrt}^\top\widehat{\boldsymbol{\Sigma}}_{srt}^{1/2}\|_2, \ \ \forall i,s,r,t \tag{14d}$$

$$\mathbf{H}_l^{\mathrm{G}}\mathbf{p}_{srt} + \mathbf{H}_l^{\mathrm{W}}\mathbf{m}_{srt} - \mathbf{H}_l^{\mathrm{D}}\mathbf{d}_{srt} \leqslant \overline{f}_l - \sqrt{\frac{1-\epsilon_l}{\epsilon_l}}\|(\mathbf{H}_l^{\mathrm{G}}\alpha_{srt}\mathbf{1}^\top + \mathbf{H}_l^{\mathrm{W}})\boldsymbol{\Sigma}_{srt}^{1/2}\|_2, \ \ \forall l,s,r,t \tag{14e}$$

$$\mathbf{H}_l^{\mathrm{G}}\mathbf{p}_{srt} + \mathbf{H}_l^{\mathrm{W}}\mathbf{m}_{srt} - \mathbf{H}_l^{\mathrm{D}}\mathbf{d}_{srt} \geqslant -\overline{f}_l + \sqrt{\frac{1-\epsilon_l}{\epsilon_l}}\|(\mathbf{H}_l^{\mathrm{G}}\alpha_{srt}\mathbf{1}^\top + \mathbf{H}_l^{\mathrm{W}})\boldsymbol{\Sigma}_{srt}^{1/2})\|_2, \ \ \forall l,s,r,t \tag{14f}$$

$$p_{isr(t-1)} \leqslant r_i x_{isr(t-1)} + (\overline{p}_i - r_i)(x_{isrt} - u_{isrt}) - \sqrt{\frac{1-\epsilon_i}{\epsilon_i}}\|\alpha_{isr(t-1)}\mathbf{1}^\top\boldsymbol{\Sigma}_{sr(t-1)}^{1/2}\|_2, \ \ \forall i,s,r,t \tag{14g}$$

$$p_{isrt} \leqslant \overline{p}_i x_{isrt} - (\overline{p}_i - r_i)u_{isrt}$$

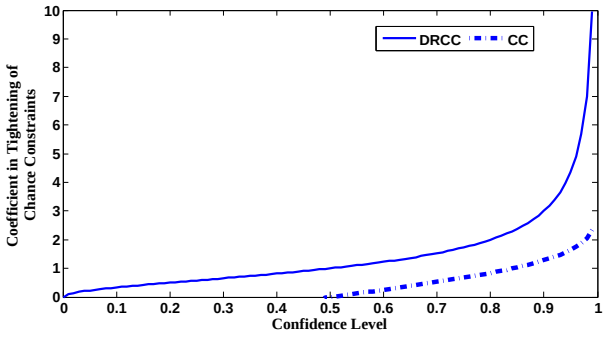

Fig. 10. The value of tightening coefficient of chance constraints as a function of confidence level $1-\epsilon$. This value in the DRCC model is $\sqrt{\frac{1-\epsilon}{\epsilon}}$, while it is equal to $\Phi^{-1}(1-\epsilon)$ in the CC model assuming a Gaussian probability distribution.

$$-\sqrt{\frac{1-\epsilon_i}{\epsilon_i}}\|\alpha_{isrt}\mathbf{1}^\top\mathbf{\Sigma}_{srt}^{1/2}\|_2,\quad \forall i,s,r,t \tag{14h}$$

$$p_{isrt}-p_{isr(t-1)}\leqslant(\underline{p}_i+\overline{r}_i)x_{isrt}-(\underline{p}_i+\overline{r}_i-r_i)u_{isrt}$$
$$-\underline{p}_i x_{isr(t-1)}-\sqrt{\frac{1-\epsilon_i}{\epsilon_i}}\|\mathbf{v}_{isrt}^\top\widehat{\mathbf{\Sigma}}_{srt}^{1/2}\|_2,\quad \forall i,s,r,t \tag{14i}$$

$$p_{isrt}-p_{isr(t-1)}\geqslant -r_i x_{isr(t-1)}-(\underline{p}_i+\underline{r}_i-r_i)u_{isrt}$$
$$-(r_i-\underline{r}_i)x_{isrt}+\sqrt{\frac{1-\epsilon_i}{\epsilon_i}}\|\mathbf{v}_{isrt}^\top\widehat{\mathbf{\Sigma}}_{srt}^{1/2}\|_2,\quad \forall i,s,r,t, \tag{14j}$$

where vector $\mathbf{v}_{isrt}$ is defined as

$$\mathbf{v}_{isrt}=\begin{bmatrix}\alpha_{isrt}\\-\alpha_{isr(t-1)}\end{bmatrix},\quad \forall i,s,r,t. \tag{14k}$$

Appendix D: Chance Constraints with Gaussian Distribution

Assuming a Gaussian probability distribution while still assuming $\boldsymbol{\mu}_{srt}=0,\ \forall s,r,t$, the arbitrarily selected chance constraint (5a) is reformulated in an exact way to a second-order cone constraint [46] as

$$p_{isrt}\leqslant\overline{p}_i x_{isrt}-\Phi^{-1}(1-\epsilon_i)\sqrt{\alpha_{isrt}\mathbf{1}^\top\mathbf{\Sigma}_{srt}\mathbf{1}\alpha_{isrt}},\ \forall i,s,r,t, \tag{15}$$

where $\Phi^{-1}(1-\epsilon)$ is the inverse cumulative distribution function. The key difference between (15) and (13b) is the coefficient in their last term. Fig. 10 illustrates the values of coefficients $\sqrt{\frac{1-\epsilon}{\epsilon}}$ and $\Phi^{-1}(1-\epsilon)$ as a function of confidence level, and thereby, their effects on tightening the chance constraints. For the same value of confidence level, it can be concluded that the DRCC model provides comparatively tighter chance constraints, which results in a more robust solution. This tightening effect is more remarkable for higher values of confidence level.

Appendix E: Violation Probabilities

To assess the *ex-post* and *ex-ante* violation probabilities, we use the test dataset containing 5,000 trajectories. The ex-post violation probability is calculated through the out-of-sample analysis based on the simulation of the unit commitment problem with considering the recourse actions of load shedding and wind spillage. To do so, for each sample trajectory j, the tight relaxed unit commitment problem is carried out to calculate the violation indicator as follows:

$$\widehat{I}_j=\begin{cases}1 & \text{if } ls_j>0 \text{ or } ws_j>0,\\ 0 & \text{otherwise,}\end{cases} \tag{16}$$

where ls_j and ws_j indicate the value of load shedding and wind spillage for sample trajectory j. We use (16) to assess the number of violations, and subsequently, the ex-post violation probability is calculated by

$$\widehat{V}=\frac{1}{N}\sum_{j=1}^{N}\widehat{I}_j, \tag{17}$$

where N is the number of sample trajectories. In contrast, for calculating the ex-ante violation probability, we do not run the tight relaxed unit commitment problem. According to the optimal in-sample values, we calculate the recourse power production of generating units for each sample trajectory by using linear decision rules. For a set of distributionally robust individual chance constraints in form of $\min_{D\in\mathcal{P}}\mathbb{P}\left(v^\top\boldsymbol{\gamma}\leqslant b\right)\geqslant 1-\epsilon$, we obtain the violation indicator for each trajectory j by

$$I_j=\begin{cases}1 & \text{if } v_k^\top\boldsymbol{\gamma}_j>b_k,\\ 0 & \text{otherwise,}\end{cases} \tag{18}$$

where k indicates each chance constraint from (5) and (7). We investigate the satisfaction of all chance constraints by I_j. Then, we assess the ex-ante violation probability with

$$V=\frac{1}{N}\sum_{j=1}^{N}I_j. \tag{19}$$

Appendix F: Resulting Robust Optimization and Stochastic Program

A. Robust Optimization

In order to recast the benchmark model in Section V.G as a robust optimization, we first calculate the mean value of selected samples, i.e., $\mathbf{m}_{srt}$, for each long-term scenario s, representative day r, and hour t. We then construct the uncertainty set of forecast errors as $\boldsymbol{\gamma}_{srt}\in\mathcal{U}_{srt}=[\boldsymbol{\gamma}_{srt}^{\min},\boldsymbol{\gamma}_{srt}^{\max}]$. For simplicity, we neglect the potential correlation between random variables. Note that the mean of forecast error is zero. To formulate the robust optimization problem, the chance constraints in the proposed DRCC model should be replaced by regular constraints, meaning that for any realization within the uncertainty set $\mathcal{U}_{srt}$, the constraints are always satisfied. For example, we provide below the formulation for the robust counterpart of (5a), imposing the capacity limit of conventional generator i. By implementing the linear decision rules, the corresponding robust constraint can be written as

$$p_{isrt}+\alpha_{isrt}\mathbf{1}^\top\boldsymbol{\gamma}_{srt}\leqslant\overline{p}_i x_{isrt},\ \forall i,s,r,t, \tag{20}$$

where $\boldsymbol{\gamma}_{srt}\in\mathcal{U}_{srt}$. By using duality theory, the robust counterpart of (20) can be formulated as

$$p_{isrt}+\boldsymbol{\eta}_{isrt}^{\mathrm{up}\top}\boldsymbol{\gamma}_{srt}^{\max}-\boldsymbol{\eta}_{isrt}^{\mathrm{dn}\top}\boldsymbol{\gamma}_{srt}^{\min}\leqslant\overline{p}_i x_{isrt},\forall i,s,r,t \tag{21a}$$

$$\eta_{zisrt}^{\mathrm{up}}-\eta_{zisrt}^{\mathrm{dn}}=\alpha_{isrt},\quad \forall z,i,s,r,t \tag{21b}$$

$$\eta^{\mathrm{up}}_{zisrt} \geqslant 0,\ \eta^{\mathrm{dn}}_{zisrt} \geqslant 0, \quad \forall z, i, s, r, t. \tag{21c}$$

where $\boldsymbol{\eta}^{\mathrm{up}}_{isrt} \in \mathbb{R}^Z$ and $\boldsymbol{\eta}^{\mathrm{dn}}_{isrt} \in \mathbb{R}^Z$ are dual variables, and z is an index for wind farms. A further elaboration on the derivation of linear robust counterpart is given in [54]–[56]. Given that fact that the mean of forecast error is zero, the objective function of the resulting robust optimization falls into the one of the proposed DRCC model, i.e., (11b).

B. Stochastic Programming

By clustering the selected samples in Section V.G, we obtain a set of short-term scenarios $\omega \in \Omega_{srt}$ associated with long-term scenario s, representative day r, and hour t. According to the number of samples in each cluster, a probability $\upsilon_{srt\omega}$ is assigned to each short-term scenario ω such that $\sum_\omega \upsilon_{srt\omega} = 1,\ \forall s, r, t$. To have a fair comparison among models, we implement linear decisions rules in this stochastic benchmark. Under each short-term scenario, all operational constraints need to be satisfied. For example, the capacity constraint of generator i imposed by (5a) in the proposed DRCC model can be now rewritten as

$$p_{isrt} + \alpha_{isrt}\mathbf{1}^\top \boldsymbol{\gamma}_{srt\omega} \leqslant \overline{p}_i x_{isrt}, \quad \forall i, s, r, t, \forall \omega \in \Omega_{srt} \tag{22}$$

where $\boldsymbol{\gamma}_{srt\omega}$ is the vector of forecast errors under short-term scenario ω. Finally, the objective function of the resulting stochastic program is

$$\min_{\mathbf{y},\mathbf{p},\boldsymbol{\alpha},\mathbf{x},\mathbf{u}} \mathbf{k}^\top \mathbf{y} + \sum_{srt\omega} \pi_s \kappa_r [\mathbf{c}^\top (\mathbf{p}_{srt} + \boldsymbol{\alpha}_{srt} \upsilon_{srtw} \mathbf{1}^\top \boldsymbol{\gamma}_{srtw}) + \mathbf{h}^\top \mathbf{u}_{srt}]. \tag{23}$$

ACKNOWLEDGMENT

The authors would like to thank the three anonymous reviewers. We also thank Stein-Erik Fleten (Norwegian University of Science and Technology) and Yury Dvorkin (New York University) for their valuable suggestions and comments, and participants at the Mathematics of Energy Systems Workshop at Isaac Newton Institute, University of Cambridge, March 2019, for helpful discussions.

REFERENCES

[1] A. J. Conejo, L. Baringo, S. J. Kazempour, and A. S. Siddiqui, *Investment in Electricity Generation and Transmission: Decision Making Under Uncertainty*. Springer, Switzerland, 2016.

[2] X. Zhang and A. J. Conejo, "Robust transmission expansion planning representing long- and short-term uncertainty," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1329–1338, Mar. 2018.

[3] S. Jin, S. M. Ryan, J. Watson, and D. L. Woodruff, "Modeling and solving a large-scale generation expansion planning problem under uncertainty," *Energy Systems*, vol. 2, no. 3-4, pp. 209–242, Nov. 2011.

[4] S. Jin, A. Botterud, and S. M. Ryan, "Temporal versus stochastic granularity in thermal generation capacity planning with wind power," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2033–2041, Sep. 2014.

[5] S. Dehghan, N. Amjady, and A. Kazemi, "Two-stage robust generation expansion planning: A mixed integer linear programming model," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 584–597, Mar. 2014.

[6] F. Maggioni, M. Cagnolari, and L. Bertazzi, "The value of the right distribution in stochastic programming with application to a newsvendor problem," *Computational Management Science*, pp. 1–20, 2019, to be published.

[7] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Oper. Res.*, vol. 62, no. 6, pp. 1358–1376, 2014.

[8] B. S. Palmintier and M. D. Webster, "Impact of operational flexibility on electricity generation planning with renewable and carbon targets," *IEEE Trans. Sustain. Energy*, vol. 7, no. 2, pp. 672–684, Apr. 2016.

[9] J. Ma, V. Silva, R. Belhomme, D. S. Kirschen, and L. F. Ochoa, "Evaluating and planning flexibility in sustainable power systems," *IEEE Trans. Sustain. Energy*, vol. 4, no. 1, pp. 200–209, Jan. 2013.

[10] B. Hua, R. Baldick, and J. Wang, "Representing operational flexibility in generation expansion planning through convex relaxation of unit commitment," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 2272–2281, Mar. 2018.

[11] A. Schwele, J. Kazempour, and P. Pinson, "Do unit commitment constraints affect generation expansion planning? A scalable stochastic model," *Energy Systems*, Jan. 2019, to be published.

[12] D. Alvarado, A. Moreira, R. Moreno, and G. Strbac, "Transmission network investment with distributed energy resources and distributionally robust security," *IEEE Trans. Power Syst.*, 2018, to be published.

[13] D. Pozo, A. Street, and A. Velloso, "An ambiguity-averse model for planning the transmission grid under uncertainty on renewable distributed generation," in *Power Systems Computation Conf.*, Jun. 2018, pp. 1–7.

[14] T. Zhang, R. Baldick, and T. Deetjen, "Optimized generation capacity expansion using a further improved screening curve method," *Elect. Power Syst. Res.*, vol. 124, pp. 47–54, Jul. 2015.

[15] A. Shortt, J. Kiviluoma, and M. O'Malley, "Accommodating variability in generation planning," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 158–169, Feb. 2013.

[16] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Oper. Res.*, vol. 58, no. 3, pp. 595–612, 2010.

[17] G. A. Hanasusanto, V. Roitch, D. Kuhn, and W. Wiesemann, "Ambiguous joint chance constraints under mean and dispersion information," *Oper. Res.*, vol. 65, no. 3, pp. 751–767, 2017.

[18] S. Zymler, D. Kuhn, and B. Rustem, "Distributionally robust joint chance constraints with second-order moment information," *Math. Program.*, vol. 137, no. 1-2, pp. 167–198, Nov. 2011.

[19] G. A. Hanasusanto and D. Kuhn, "Conic programming reformulations of two-stage distributionally robust linear programs over Wasserstein balls," *Oper. Res.*, vol. 66, no. 3, pp. 849–869, 2018.

[20] P. Mohajerin Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations," *Math. Program.*, Jul. 2017.

[21] H. Rahimian and S. Mehrotra, "Distributionally robust optimization: A review," 2019. [Online]. Available: https://arxiv.org/abs/1908.05659

[22] C. Zhao and R. Jiang, "Distributionally robust contingency-constrained unit commitment," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 94–102, Jan. 2018.

[23] P. Xiong, P. Jirutitijaroen, and C. Singh, "A distributionally robust optimization model for unit commitment considering uncertain wind power generation," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 39–49, Jan. 2017.

[24] Y. Chen *et al.*, "A distributionally robust optimization model for unit commitment based on Kullback-Leibler divergence," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5147–5160, Sep. 2017.

[25] W. Wei, F. Liu, and S. Mei, "Distributionally robust co-optimization of energy and reserve dispatch," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 289–300, Jan. 2016.

[26] W. Xie and S. Ahmed, "Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1860–1867, Mar. 2018.

[27] R. Mieth and Y. Dvorkin, "Data-driven distributionally robust optimal power flow for distribution systems," *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 363–368, 2018.

[28] Y. Guo, K. Baker, E. Dall'Anese, Z. Hu, and T. H. Summers, "Data-based distributionally robust stochastic optimal power flow–Part I: Methodologies," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1483–1492, Mar. 2019.

[29] C. Ordoudis, V. A. Nguyen, D. Kuhn, and P. Pinson, "Energy and reserves dispatch with distributionally robust joint chance constraints," 2018. [Online]. Available: http://www.optimization-online.org/DB_HTML/2018/12/6962.html

[30] A. B. Philpott, V. L. de Matos, and L. Kapelevich, "Distributionally robust SDDP," *Computational Management Science*, vol. 15, no. 3–4, pp. 431–454, Oct. 2018.

[31] F. Qiu and J. Wang, "Distributionally robust congestion management with dynamic line ratings," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2198–2199, Jul. 2015.

[32] Z. Shi, H. Liang, S. Huang, and V. Dinavahi, "Distributionally robust chance-constrained energy management for islanded microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2234–2244, Mar. 2019.

[33] A. Zare, C. Y. Chung, J. Zhan, and S. O. Faried, "A distributionally robust chance-constrained MILP model for multistage distribution system planning with uncertain renewables and loads," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5248–5262, Sep. 2018.
[34] P. Maghouli, S. H. Hosseini, M. Oloomi Buygi, and M. Shahidehpour, "A scenario-based multi-objective model for multi-stage transmission expansion planning," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 470–478, Feb. 2011.
[35] K. Poncelet, H. Höschle, E. Delarue, A. Virag, and W. D'haeseleer, "Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 1936–1948, May 2017.
[36] Y. Liu, R. Sioshansi, and A. J. Conejo, "Hierarchical clustering to find representative operating periods for capacity-expansion modeling," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3029–3039, May 2018.
[37] D. A. Tejada-Arango, M. Domeshek, S. Wogrin, and E. Centeno, "Enhanced representative days and system states modeling for energy storage investment analysis," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6534–6544, Nov. 2018.
[38] Y. Zhang, S. Shen, and J. L. Mathieu, "Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1378–1388, Mar. 2017.
[39] B. Li, R. Jiang, and J. L. Mathieu, "Distributionally robust chance constrained optimal power flow assuming unimodal distributions with misspecified modes," *IEEE Transactions on Control of Network Systems*, 2019, to be published.
[40] D. Kuhn, W. Wiesemann, and A. Georghiou, "Primal and dual linear decision rules in stochastic and robust optimization," *Math. Program.*, vol. 130, no. 1, pp. 177–209, 2011.
[41] B. Hua and R. Baldick, "A convex primal formulation for convex hull pricing," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3814–3823, Sep. 2017.
[42] H.-P. Chao, "Incentives for efficient pricing mechanism in markets with non-convexities," *Journal of Regulatory Economics*, pp. 1–29, 2019, to be published.
[43] K. Pan and Y. Guan, "A polyhedral study of the integrated minimum-up/-down time and ramping polytope," *arXiv preprint arXiv:1604.02184*, 2016.
[44] I. Peña, C. B. Martinez-Anido, and B. Hodge, "An extended IEEE 118-bus test system with high renewable penetration," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 281–289, Jan. 2018.
[45] [Online]. Available: https://github.com/farzanehpourahmadi/DRCC.git
[46] D. Bienstock, M. Chertkov, and S. Harnett, "Chance-constrained optimal power flow: Risk-aware network control under uncertainty," *SIAM Review*, vol. 56, no. 3, pp. 461–495, 2014.
[47] K. Margellos, P. Goulart, and J. Lygeros, "On the road between robust optimization and the scenario approach for chance constrained optimization problems," *IEEE Trans. Autom. Control*, vol. 59, no. 8, pp. 2258–2263, Aug. 2014.
[48] B. P. G. Van Parys, D. Kuhn, P. J. Goulart, and M. Morari, "Distributionally robust control of constrained stochastic systems," *IEEE Trans. Autom. Control*, vol. 61, no. 2, pp. 430–442, Feb. 2016.
[49] G. C. Calafiore and L. El Ghaoui, "On distributionally robust chance-constrained linear programs," *Journal of Optimization Theory and Applications*, vol. 130, no. 1, pp. 1–22, 2006.
[50] B. L. Gorissen, İ. Yanıkoğlu, and D. den Hertog, "A practical guide to robust optimization," *Omega*, vol. 53, pp. 124–137, Jun. 2015.
[51] D. Bertsimas, V. Gupta, and N. Kallus, "Data-driven robust optimization," *Math. Program.*, vol. 167, no. 2, pp. 235–292, 2018.
[52] M. Wagner, "Stochastic 0-1 linear programming under limited distributional information," *Oper. Res. Lett.*, vol. 36, no. 2, pp. 150–156, 2008.
[53] Y. Dvorkin, "A chance-constrained stochastic electricity market," 2019. [Online]. Available: https://arxiv.org/abs/1906.06963
[54] D. Bertsimas and M. Sim, "Tractable approximations to robust conic optimization problems," *Math. Program.*, vol. 107, no. 1-2, pp. 5–36, 2006.
[55] F. Pourahmadi, H. Heidarabadi, S. H. Hosseini, and P. Dehghanian, "Dynamic uncertainty set characterization for bulk power grid flexibility assessment," *IEEE Systems Journal*, pp. 1–11, 2019.
[56] R. A. Jabr, "Adjustable robust OPF with renewable energy sources," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4742–4751, 2013.