European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

Asymptotically tight conic approximations for chance-constrained AC optimal power flow

Abolhassan Mohammadi Fathabad[a], Jianqiang Cheng[a,*], Kai Pan[b], Boshi Yang[c]

[a] Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721, USA
[b] Department of Logistics and Maritime Studies, Faculty of Business, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[c] School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC, USA

ARTICLE INFO

Article history:
Received 22 May 2021
Accepted 8 June 2022
Available online xxx

Keywords:
Stochastic programming
Two-sided chance constraint
AC optimal power flow
Second-order cone programming
Piecewise linear approximation

ABSTRACT

The increasing penetration of renewable energy in power systems calls for secure and reliable system operations under significant uncertainty. To that end, the chance-constrained AC optimal power flow (CC-ACOPF) problem has been proposed. Most research in the literature of CC-ACOPF focuses on one-sided chance constraints; however, two-sided chance constraints (TCCs), albeit more complex, provide more accurate formulations as both upper and lower bounds of the chance constraints are enforced simultaneously. In this paper, we introduce a fully two-sided CC-ACOPF problem (TCC-ACOPF), in which the active/reactive generation, voltage, and power flow all remain within their upper/lower bounds simultaneously with a predefined probability. Instead of applying Bonferroni approximation or scenario-based approaches, we present an efficient second-order cone programming (SOCP) approximation of the TCCs under Gaussian Mixture (GM) distribution via a piecewise linear (PWL) approximation. Compared to the conventional normality assumption for forecast errors, the GM distribution adds an extra level of accuracy representing the uncertainties. Moreover, we show that our SOCP formulation has adjustable rates of accuracy and its optimal value enjoys asymptotic convergence properties. Furthermore, an algorithm is proposed to speed up the solution procedure by optimally selecting the PWL segments. Finally, we demonstrate the effectiveness of our proposed approaches with both real historical data and synthetic data on the IEEE 30-bus and 118-bus systems. We show that our formulations provide significantly more robust solutions (about 60% reduction in constraint violation) compared to other state-of-art ACOPF formulations.

1. Introduction

The penetration of renewable energy such as wind and solar increases rapidly in the power system. At the same time, the uncertain output of these non-dispatchable renewable sources prompts various issues in power system operation. In particular, with many renewable energy sources, it is well known that supply-and-demand balance control becomes very difficult to manage. Indeed, the integration of the uncertain power sources increases the risk of supply-and-demand mismatch, leading to significant challenges from voltage fluctuations to overloaded branches and can shut down the entire or parts of the power system (De Rubira & Hug, 2016; Filabadi & Azad, 2020; Kundur et al., 2004). Such complexities raise the need for fast and reliable optimization methods that can securely and economically schedule the power system operation. As a solution, the optimal power flow (OPF) problem provides real-time control measures to support the system's stability. It guarantees the system's secure operation by enforcing constraints such as voltage limits, generation limits, and line capacity.

First introduced by Carpentier (1962) over half a century ago, the OPF problem has gained great attention due to its importance in power system operations. Due to the physical complexity in power systems, the majority of the OPF research in the literature oversimplifies the real OPF problem using direct current (DC) and deterministic formulations (Lin, Magnago, & Alemany, 2018; Skolfield & Escobedo, 2022). The DC formulation is an approximation of the actual nonlinear alternating current (AC) power flow formulas and the deterministic formulation is a rough estimation of the stochastic nature of the power system. In traditional centralized power systems, such simplifications, albeit imperfect, are acceptable to run the system without major catastrophic problems.

* Corresponding author.
E-mail addresses: amohammadi@email.arizona.edu (A. Mohammadi Fathabad), jqcheng@arizona.edu, jqcheng@email.arizona.edu (J. Cheng), kai.pan@polyu.edu.hk (K. Pan), boshiy@clemson.edu (B. Yang).

https://doi.org/10.1016/j.ejor.2022.06.020
0377-2217/

Indeed, the deterministic DC models are still being used in many commercial and industrial applications (Stott & Alsaç, 2012).

However, the large-scale integration of renewable generations in recent years is exposing the deficiencies of the traditional OPF formulations. In particular, the DC approximation neglects both network losses and reactive power, which are natural components of a power system network. Omitting these components would lead to decisions that are not feasible to the realistic AC optimal power flow (ACOPF) problem. On the other hand, as penetration levels of volatile and intermittent sources, such as wind and solar power, reach massive fractions of the total supplied power, the risk of extreme catastrophic outcomes (e.g., power supply interruptions, power mismatch, increased power losses, and network instability) escalates rapidly. The traditional deterministic models, which rely on day-ahead or hour-ahead OPF analyses, do not work as well as before because it is more difficult to predict the renewable power outputs as residential houses are becoming both the load and generation points. These uncertain forecast errors can lead to insufficient or excess electricity generation, and if unaccounted for, can result in major reliability issues and cause significant damage to the system via costly blackouts or other catastrophic events. In fact, major power outages and the risk for power interruptions increased rapidly in the past few decades (Bloomenergy, 2021; NREL, 2021; WirfsBrock, 2014) as more renewables are integrated in the system. To account for the increasing risks caused by uncertainty, stochastic ACOPF problems have been addressed in recent years (Capitanescu, 2016).

An accurate representation of the stochastic ACOPF is achieved using chance-constrained optimization (Bienstock, Chertkov, & Harnett, 2014). In particular, the ACOPF problem enforces limits on voltage, active power, reactive power, and power flow, and a chance-constrained ACOPF (CC-ACOPF) enforces those limit constraints with a certain probability. In other words, chance constraints restrict the feasible region so that the confidence level of the solution is high, and thus reflect the stochastic nature of the problem. Moreover, in a recent survey of the actual power system operators, it was determined that using the chance-constrained formulation to choose a predetermined level of violation for constraints is both an intuitive and transparent way of representing an ACOPF problem under uncertainty (Roald & Andersson, 2017).

While the CC-ACOPF presents an intuitive way to model forecast uncertainty, the nonlinearity of the AC power flow equations and the probabilistic constraints render the problem computationally intractable (Nemirovski & Shapiro, 2007; Paudyal, Canizares, & Bhattacharya, 2011). To develop a practical method, it is necessary to consider both an approximation of the AC power flow equations and a pertinent convex approximation of the chance constraints.

First, as discussed above, the underlying physical characteristics of a power system include nonconvex and nonlinear AC power flow equations; thus, most works in literature approximate these nonlinear equations with an imprecise DC model to reduce the complexity (Aigner, Clarner, Liers, & Martin, 2022; Bienstock et al., 2014; Overbye, Cheng, & Sun, 2004; Vrakopoulou, Katsampani, Margellos, Lygeros, & Andersson, 2013; Xie & Ahmed, 2017; Zhang, Shen, & Mathieu, 2016). Modeling the actual AC power flow, on the other hand, allows us to accurately consider new constraints and chance constraints on reactive power, power angles, and power transmission capacity (Zohrizadeh et al., 2020). More importantly, the ACOPF problem needs to be solved at different levels of the power system (e.g., transmission and distribution grids) and under different stages (e.g., long-term planning and short-term operations). Thus, a tractable and accurate formulation of the ACOPF problem is highly demanded. Several different methods for linear approximation of full CC-ACOPF are proposed in the literature that attempt to represent the output variables as linear combinations of input variables. For example, Dall'Anese, Baker, & Summers (2017) distinctly models linearized AC power flow equations around a given voltage profile. In another effort, Vrakopoulou et al. (2013) uses SDP relaxations for the AC power equations; however, the resulting formulation cannot provide robust guarantees and it is computationally expensive. Others (Hojjat & Javidi, 2015; Zhang & Li, 2011) have considered full linearization of the responses around expected values of the random variables using linear decision rules. In this study, similar to Fu & McCalley (2001); Roald & Andersson (2017) and Lubin, Dvorkin, & Roald (2019), we use Taylor expansion to linearize the AC power flow equations around a predicted operating point. This predicted point is identified by solving the deterministic ACOPF problem, and our model approximates the impact of uncertainty as a linear function of the uncertain power injections only around that particular operating point. Hence, it is more accurate than the other techniques as it linearizes the system around one operating point instead of fully linearizing the system (Roald & Andersson, 2017). This also further allows for the development of analytical chance-constrained reformulations.

Second, to facilitate the development of analytical chance-constrained reformulations, a one-sided chance constraint (OCC) relaxation is commonly used. In particular, most works in the literature of CC-ACOPF treat the physical bounds separately (Dall'Anese et al., 2017; Lubin et al., 2019; Roald & Andersson, 2017), i.e., a single chance constraint is imposed on the upper bound and another chance constraint is imposed on the lower bound. While using OCCs is convenient as they can be reformulated and implemented more easily (Baker, Dall'Anese, & Summers, 2016; Bienstock et al., 2014; Zhang & Li, 2011), the OCC relaxation provides an inexact approximation of the OPF problem. More specifically, it is known that active power, reactive power, and voltage at each bus, as well as the power flow at each branch in general, have both lower and upper bound limits. Hence, it is more accurate to represent each by a two-sided chance constraint (TCC). To the best of our knowledge, Lubin, Bienstock, & Vielma (2015), Pena-Ordieres, Molzahn, Roald, & Waechter (2019), and Xie & Ahmed (2017) are the only known studies treating the lower and upper bounds simultaneously. Among them, the distributionally robust results in Pena-Ordieres et al. (2019) and Xie & Ahmed (2017) depend on inaccurate DC approximations, and the analytical results in Lubin et al. (2015) only consider a subset of constraints with a TCC formulation that is limited to a Gaussian assumption.

In fact, many papers in the literature that study closed-form analytical reformulations of CC-ACOPF simply model the forecast errors through Gaussian distribution (Bienstock et al., 2014; Li, Vrakopoulou, & Mathieu, 2017; Lubin et al., 2015; Lubin et al., 2019; Roald & Andersson, 2017). This Gaussian assumption is often criticized in the literature as it may lead to further inaccuracies that can cause cascading shutdowns and power interruptions in the grid. To be more specific, recent statistical analyses of renewable forecast errors have shown that the forecast error distribution differs greatly from the commonly assumed normal distribution (Hodge & Milligan, 2011; Lange, 2005). In particular, the large-scale availability of historical data on renewable generation and forecast can be analyzed to obtain data on renewable forecast errors. This forecast data then can be analyzed to obtain an estimation of the forecast error distribution. For example, Hodge et al. (2012) study historical wind generation data from multiple countries to analyze the underlying probability distributions of forecast errors, and the results show that forecast error distributions are skewed in many cases. They further show that the simple Gaussian distribution performs poorly when skewness is present. In view of this, in our research, we have modeled the forecast error distribution through Gaussian mixture (GM) modeling. With the GM modeling approach, the forecast error distribution can be modeled as a

convex combination of multiple normal distributions with respective means and variances. Consequently, it not only encompasses the normal distribution but also can be used to model a continuous distortion of the latter, such as skewness (Bertholon, Monfort, & Pegoraro, 2007). More importantly, any distribution can be approximated by a GM distribution through nonparametric estimation method. For instance, if we use the kernel density estimation, a commonly used nonparametric method proposed by Rosenblatt (1956) and Parzen (1962), to estimate the distribution, then the well-known normal kernel density estimator (Dattatreya & Kanal, 1990; Reynolds, 2009; Zhuang, Huang, Palaniappan, & Zhao, 1996) becomes a GM distribution. Hence, we can approximate the distribution of any uncertainty using a GM distribution. Furthermore, GM is stable by convolution and easy to simulate. Therefore, employing the GM distribution helps us represent the forecast error uncertainties more accurately and develop OPF solutions that are significantly robust against such uncertainties, thereby adding great accuracy and flexibility to our proposed chance-constrained model.

The resulting two-sided chance-constrained ACOPF (TCC-ACOPF) problem under the GM distribution, however, relies on nonconvex and nonlinear TCCs that need to be convexly approximated. In this paper, we present a tractable approximation methodology for the TCC-ACOPF to provide an accurate solution for the problem in a timely manner. We first provide a convex approximation of the TCC under a GM distribution. Furthermore, we show that our approximation is exact for the case of simple Gaussian and some special GM distributions, so it generalizes the convexity results in Lubin et al. (2015). In spite of being convex, this approximation is nonlinear, which makes it less efficient for large-scale applications. The source of nonlinearity, i.e., the Gaussian cumulative distribution function (CDF), is then approximated by a piecewise linear (PWL) function, which is often used in the existing studies (see, e.g., Ardestani-Jaafari & Delage, 2016 and Kuryatnikova, Ghaddar, & Molzahn, 2021) for the tractability purpose. The resulting final formulation is a series of linear and second-order conic (SOC) constraints, which can be solved efficiently by many commercial solvers.

The quality of our final SOC approximation of the TCC is adjustable depending on the quality of the PWL approximation of the Gaussian CDF. Intuitively, the more linear pieces we have, the more accurate the approximation is. However, a PWL approximation with more pieces also leads to more constraints in our final SOC formulation, which means a higher computational cost. Hence, it is of our interest to choose an optimal PWL approximation that guarantees a given accuracy threshold with the minimum number of segments.

Several studies have investigated the idea of finding the optimal PWL fit to a uni-variate function. For example, Hamann & Chen (1994) finds the optimal locations of segments based on a root-mean-square error tolerance. Tomek (1974) develops two heuristics to minimize the number of approximating segments subject to an error limit. More recently, Rebennack & Krasko (2020) and Kong & Maravelias (2020) developed algorithms that use mixed-integer techniques for finding segments that incite the exact amount of error required; however, their methodology is only applicable to bounded functions. In this paper, to choose an optimal PWL function, we introduce an algorithm based on the linear interpolation error. Our methodology relies on the monotonicity and concavity of the target function, and it avoids introducing mixed-integer variables. Moreover, it can be generalized to approximate any strictly monotone concave or convex function that is bounded from at least one side. In our theoretical results, we prove that our algorithm provides the best fit, and in our numerical results, we compare this algorithm with the uni-distance algorithm that is commonly used to obtain a PWL approximation. We show that our method can achieve similar accuracy with significantly fewer linear segments (e.g., 40% reduction in number of piece when the error tolerance is 0.05%). Hence, employing our algorithm speeds up the computational time for our SOC approximation of the TCCs significantly.

In summary, the main contributions of this paper are the following:

- We present a TCC-ACOPF to model the stochastic ACOPF problem, which models the forecast error uncertainty more accurately using the GM distribution. In particular, the GM distribution addresses non-normalities such as skewness in power forecast data, which are not captured by the common normal distribution.
- We present a convex inner approximation of a TCC on a GM distribution with K components, and we show that our convex formulation is exact under an easily verifiable condition. As this intermediate convex formulation may be expensive to solve on large-scale applications, we introduce a tractable SOC approximation of it based on a PWL approximation of the standard normal CDF.
- Moreover, the quality of our SOC approximation depends on the quality of the PWL approximation, which can be improved by increasing the number of well-positioned PWL segments. We prove that our SOC approximation enjoys asymptotic convergence properties. As higher number of PWL segments lead to a higher computational cost, we propose an algorithm that obtains the minimum number of segments (and their optimal positioning) required for a PWL function to speed up the computation.
- We report computational results with both synthetic and real-world datasets, which show that TCC-ACOPF can be solved efficiently (in a similar timeframe) as compared to the state-of-art OCC-ACOPF and deterministic ACOPF. Our results show that the TCC-ACOPF approach significantly improves the robustness and feasibility of solutions.

The remainder of the paper is organized as follows. Section 2 introduces the ACOPF problem and the TCC-ACOPF problem. Section 3 shows how to reformulate the TCC-ACOPF into a convex problem and then approximate it by a second-order cone program. An algorithm to speed up the computation is also proposed in this Section. Section 4 numerically illustrates the strengths of the proposed model. Section 5 concludes the paper.

2. Mathematical model

In this section, we introduce mathematical formulations for deterministic ACOPF and TCC-ACOPF, in the following Sections 2.1 and 2.2, respectively. We use boldface and normal symbols to represent vectors and scalars, respectively, throughout the paper.

2.1. Deterministic ACOPF

We use $\mathcal{B}$, $\mathcal{G}$, and $\mathcal{R}$ to denote the set of all buses, thermal generators, and renewable generators respectively. For each bus $i \in \mathcal{B}$, we use $\mathcal{G}_i$ (resp. $\mathcal{R}_i$) to denote the set of thermal generators (resp. renewable generators) at this bus, and $\bar{v}_i$ (resp. $\underline{v}_i$) to denote the upper (resp. lower) bound on nodal voltage magnitude at this bus. For each bus $i \in \mathcal{B}$ and bus $j \in \mathcal{B}$, we use $\mathcal{L}$ to denote the set of tuples (i, j) such that there is a branch between bus i and bus j, and $\bar{I}_{ij}$ to denote this branch's apparent power flow limit. For each thermal generator $g \in \mathcal{G}$, we use $\bar{P}_g$ (resp. $\underline{P}_g$) to denote its maximum (resp. minimum) active power generation amount, and $\bar{Q}_g$ (resp. $\underline{Q}_g$) to denote its maximum (resp. minimum) reactive power generation amount. To define decision variables, we let p_g (resp. q_g)

represent the active (resp. reactive) power output of each thermal generator $g \in \mathcal{G}$, e^{p}_{ij} (resp. e^{q}_{ij}) represent the active (resp. reactive) power flow of each tuple $(i, j) \in \mathcal{L}$, and v_i (resp. θ_i) represent the nodal voltage magnitude (resp. angle) at each bus $i \in \mathcal{B}$. Therefore, the ACOPF model can be described as follows:

$$\min_{\boldsymbol{p},\boldsymbol{q},\boldsymbol{v},\boldsymbol{\theta},\, e^{\mathrm{p}},e^{\mathrm{q}}} \sum_{g\in\mathcal{G}} c_g(p_g) \tag{1a}$$

$$\text{s.t. } \underline{P}_g \le p_g \le \bar{P}_g, \qquad \forall\, g \in \mathcal{G}, \tag{1b}$$

$$\underline{Q}_g \le q_g \le \bar{Q}_g, \qquad \forall\, g \in \mathcal{G}, \tag{1c}$$

$$\underline{v}_i \le v_i \le \bar{v}_i, \qquad \forall\, i \in \mathcal{B}, \tag{1d}$$

$$(e^{\mathrm{p}}_{ij})^2 + (e^{\mathrm{q}}_{ij})^2 \le (\bar{I}_{ij})^2, \qquad \forall\, (i, j) \in \mathcal{L}, \tag{1e}$$

$$\varphi(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{v}, \boldsymbol{\theta}) = 0, \tag{1f}$$

$$\theta_{ref} = 0. \tag{1g}$$

In the objective function (1a), $c_g(p_g)$ represents the power production cost of generator g and it is typically a convex quadratic function (Wood, Wollenberg, & Sheblé, 2013). Constraints (1b) and (1c) restrict the active and reactive generation amount of each unit, respectively. Constraints (1d) set the upper/lower bounds of the nodal voltage magnitude at each bus $i \in \mathcal{B}$. The capacity of each branch is bounded by (1e) and the voltage angle at the reference bus is set to zero by (1g). The set of active and reactive power balance equations from the Kirchhoff's current and voltage laws are enforced by (1f). Specifically, the active and reactive power flows are functions of the voltage magnitudes $\boldsymbol{v}$ and voltage angles $\boldsymbol{\theta}$, as illustrated in the following line power flow equations:

$$e^{\mathrm{p}}_{ij} = v_i v_j \big(G_{ij}\cos(\theta_i - \theta_j) + B_{ij}\sin(\theta_i - \theta_j)\big), \ \ \forall (i, j) \in \mathcal{L}, \tag{2a}$$

$$e^{\mathrm{q}}_{ij} = v_i v_j \big(G_{ij}\sin(\theta_i - \theta_j) - B_{ij}\cos(\theta_i - \theta_j)\big), \ \ \forall (i, j) \in \mathcal{L}, \tag{2b}$$

where parameters G_{ij} and B_{ij} represent the real and imaginary parts of network admittance for each tuple $(i, j) \in \mathcal{L}$, respectively. It follows that the nodal power flow equations can be represented by as follows:

$$\sum_{g\in\mathcal{G}_i} p_g - D^{\mathrm{p}}_i + \sum_{g\in\mathcal{R}_i} r^{\mathrm{p}}_g = v_i^2 G_{ii} + \sum_{j:(i,j)\in\mathcal{L}} e^{\mathrm{p}}_{ij}, \ \ \forall i \in \mathcal{B}, \tag{3a}$$

$$\sum_{g\in\mathcal{G}_i} q_g - D^{\mathrm{q}}_i + \sum_{g\in\mathcal{R}_i} r^{\mathrm{q}}_g = -v_i^2 B_{ii} + \sum_{j:(i,j)\in\mathcal{L}} e^{\mathrm{q}}_{ij}, \ \ \forall i \in \mathcal{B}, \tag{3b}$$

which shows that the net active (resp. reactive) power injection at each bus $i \in \mathcal{B}$ is equal to the active (resp. reactive) power flow leaving this bus. In (3), D^{p}_i (resp. D^{q}_i) denotes the active (resp. reactive) power demand at each bus $i \in \mathcal{B}$, $\mathcal{R}_i$ denotes the set of renewable generators at each bus $i \in \mathcal{B}$, and r^{p}_g (resp. r^{q}_g) denotes the forecast active (resp. reactive) power injections from a given renewable generator $g \in \mathcal{R}$. Note that since for any $(i, j) \in \mathcal{L}$, if v_i, v_j, θ_i, and θ_j are given, then e^{p}_{ij} and e^{q}_{ij} are uniquely determined. Thus, e^{p} and e^{q} are not involved in the function $\varphi(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{v}, \boldsymbol{\theta}) = 0$ as arguments in (1f).

From (3), we can see that there are more variables than equations, which implies that some variables can be chosen independently and the others will be implicitly determined. In fact, certain physical structure settings of the power systems enable us to choose such variables independently, thereby controlling the balance in (3). Specifically, to that end, three bus types are considered in power system operations, i.e., pv buses, pq buses, and a $v\theta$ bus. The pv buses (referred to as "generation buses" and collected in set $\mathcal{B}_{pv}$) maintain constant values of active power generation and voltage magnitude; the pq buses (referred to as "load buses" or "buses without generation or load" and collected in set $\mathcal{B}_{pq}$) maintain constant active and reactive power outputs; the $v\theta$ bus (referred to as "reference bus" and collected in singleton $\mathcal{B}_{v\theta}$) is unique and maintains constant values of the voltage magnitude and angle (Kothari & Nagrath, 2003; Roald, 2016).

The deterministic ACOPF (1) assumes that the renewable generation outputs (i.e., r^{p}_g and r^{q}_g for each $i \in \mathcal{B}$) are known exactly; that is, the forecast values are exact. However, due to the intermittent nature (e.g., wind fluctuates and solar relies on sunny weather), renewable energy actually creates huge uncertainties to power system operations, thereby requesting advanced decision making approaches under uncertainty for the ACOPF problem.

Remark 1. The deterministic ACOPF (1) is solved to support the power system operations at different levels and stages. For instance, at the transmission level, the transmission grid operator may solve the day-ahead unit commitment and economic dispatch problems with the ACOPF formulation incorporated to perform market clearing. The operator may also solve the ACOPF (1) in the real-time market to evaluate the imbalances of power supply and demand and thus activate the ancillary services, thereby ensuring the system stability. Note that the ACOPF formulation becomes more crucial than before at the transmission level because the high penetration of renewable energy leads to significant uncertainties in the system and a more accurate OPF formulation can help the grid operator make a decision that can better hedge against the uncertainties. In addition, at the distribution level, the distribution grid operator may solve the ACOPF formulation in various stages, such as generation expansion planning and short-term operations. Specifically, as a distribution grid is usually represented as a radial network, the generic ACOPF formulation (1) can be equivalently reformulated as an SOC formulation (Farivar & Low, 2013) by using the special radial network structure. Note that our proposed approach in the remainder of this paper also works for this equivalent ACOPF formulation.

Remark 2. Note that the non-dispatchable renewable generation may have a large impact on the overall cost of the power system. Besides adjusting the generation outputs of the existing dispatchable thermal generators in the system, the system imbalances due to the renewable generation uncertainty may also be compensated by other market tools, such as ancillary services. These services could range from energy storage devices and immediate power purchase strategies, to reactive power generation devices. These services can be implemented at different levels and stages of the power system (as mentioned in Remark 1), including transmission and distribution levels, to balance the network. Thus, a direct cost component for each service can be considered in the objective function of model (1), and we leave this for the future research.

2.2. TCC-ACOPF

To account for the uncertainty due to renewable generation, we introduce a TCC-ACOPF that ensures a secure system operation under uncertainty. In particular, for a given renewable generator $g \in \mathcal{R}$, we denote this generator's uncertain active power generation amount by $r^{\mathrm{p}}_g(\boldsymbol{\xi}) = r^{\mathrm{p}}_g + \xi_g$, where r^{p}_g denotes the forecast

active power generation amount of generator g, ξ_g denotes a random variable representing the real-time deviation of this generator from the forecast value r_g^{p}, and $\boldsymbol{\xi}$ denotes the vector of ξ_g's with any $g \in \mathcal{R}$. Facing the active power generation variation of each renewable generator $g \in \mathcal{R}$, the power system has to respond to the variation by adjusting the values of other decision variables, i.e., $r_g^{\mathrm{q}}(\boldsymbol{\xi})$ for any $g \in \mathcal{R}$, $p_g(\boldsymbol{\xi})$ for any $g \in \mathcal{G}$, $q_g(\boldsymbol{\xi})$ for any $g \in \mathcal{G}$, $v_i(\boldsymbol{\xi})$ for any $i \in \mathcal{B}$, and $\theta_i(\boldsymbol{\xi})$ for any $i \in \mathcal{B}$, thereby controlling system balance and stability in the real-time operations. Technically, such responses maintain the ACOPF model (1) to be feasible. To track such adjustments and avoid high complexities, system operators often adopt a family of affine response control policies for practical purpose (Jaleeli, VanSlyck, Ewart, Fink, & Hoffmann, 1992). In this paper the response policies are selected as follows. First, the reactive power output of a renewable generator g changes following the active power output variations of this renewable generator according to

$$r_g^{\mathrm{q}}(\boldsymbol{\xi}) = \gamma_g r_g^{\mathrm{p}}(\boldsymbol{\xi}), \ \ \forall g \in \mathcal{R}, \tag{4}$$

where γ_g is a decision variable that depends on operational requirements (Cabrera-Tobar, Bullich-Massagué, Aragüés-Peñalba, & Gomis-Bellmunt, 2019; Roald, 2016). Second, according to the automatic generation control (AGC) policy (Borkowska, 1974; Venzke, Halilbasic, Markovic, Hug, & Chatzivasileiadis, 2017), the active power output of a thermal (dispatchable) generator under the renewable generation uncertainty is adjusted by the following equation:

$$p_g(\boldsymbol{\xi}) = p_g - \alpha_g \Xi, \ \ \forall g \in \mathcal{G}, \tag{5}$$

where α_g is the participation factor of each thermal generator $g \in \mathcal{G}$ and $\Xi = \sum_{g \in \mathcal{R}} \xi_g$. In our model, α_g will be optimized as a decision variable for any $g \in \mathcal{G}$, and it indicates the fraction of the total forecast error that is compensated by thermal generator g. Third, a distinction between pv, pq, and $v\theta$ buses becomes important for the reactive power balancing and voltage control. Considering common practice, we assume that the voltage is adjusted at pq buses to keep the reactive power constant with uncertainty, however, pv and $v\theta$ buses can adjust their reactive power to keep the voltage magnitude constant with uncertainty.

With the uncertainty representation and response policies described above, we then can enforce the probability that each set of constraints of (1b)–(1e) hold with respect to the distribution of $\boldsymbol{\xi}$, leading to the following four sets of chance constraints, respectively:

$$\mathbb{P}\big(\underline{P}_g \le p_g(\boldsymbol{\xi}) \le \bar{P}_g\big) \ge 1 - \epsilon_{\mathrm{p}}, \qquad \forall\, g \in \mathcal{G}, \tag{6a}$$

$$\mathbb{P}\Big(\underline{Q}_g \le q_g(\boldsymbol{\xi}) \le \bar{Q}_g\Big) \ge 1 - \epsilon_{\mathrm{q}}, \qquad \forall\, g \in \mathcal{G}, \tag{6b}$$

$$\mathbb{P}\big(\underline{v}_i \le v_i(\boldsymbol{\xi}) \le \bar{v}_i\big) \ge 1 - \epsilon_{\mathrm{v}}, \qquad \forall\, i \in \mathcal{B}, \tag{6c}$$

$$\mathbb{P}\left(\left(e_{ij}^{\mathrm{p}}(\boldsymbol{\xi})\right)^2 + \left(e_{ij}^{\mathrm{q}}(\boldsymbol{\xi})\right)^2 \le \left(\bar{I}_{ij}\right)^2\right) \ge 1 - \epsilon_{\mathrm{L}}, \quad \forall\, (i,j) \in \mathcal{L}. \tag{6d}$$

Here, $\mathbb{P}$ denotes the probability of an event with respect to the distribution of $\boldsymbol{\xi}$. Constraints (6a) enforce both the upper and lower bounds on active power generation amount of a given generator $g \in \mathcal{G}$ to be held simultaneously with a probability that is no less than $1 - \epsilon_{\mathrm{p}}$. Similarly, the probabilistic bounds on reactive power flow, voltage magnitude, and apparent power flow are enforced through (6b), (6c), and (6d), with violation probability less than ϵ_{q}, ϵ_{v}, and ϵ_{L}, respectively.

Chance constraints (6a)–(6c) are TCCs where the stochastic constraints are linear in the variables $p_g(\boldsymbol{\xi})$, $q_g(\boldsymbol{\xi})$, and $v_i(\boldsymbol{\xi})$, respectively. As the stochastic constraints $\left(e_{ij}^{\mathrm{p}}(\boldsymbol{\xi})\right)^2 + \left(e_{ij}^{\mathrm{q}}(\boldsymbol{\xi})\right)^2 \le \left(\bar{I}_{ij}\right)^2$ are nonlinear in the variables $e_{ij}^{\mathrm{p}}(\boldsymbol{\xi})$ and $e_{ij}^{\mathrm{q}}(\boldsymbol{\xi})$, constraints (6d) are more challenging to deal with. Here, we follow the approach in Lubin et al. (2015) to inner approximate (6d) with the following constraints

$$\mathbb{P}\Big(|e_{ij}^{\mathrm{p}}(\boldsymbol{\xi})| \le \zeta_{ij}^{\mathrm{p}}\Big) \ge 1 - \beta\,\epsilon_{\mathrm{L}}, \qquad \forall\, (i,j) \in \mathcal{L}, \tag{7a}$$

$$\mathbb{P}\Big(|e_{ij}^{\mathrm{q}}(\boldsymbol{\xi})| \le \zeta_{ij}^{\mathrm{q}}\Big) \ge 1 - (1-\beta)\epsilon_{\mathrm{L}}, \qquad \forall\, (i,j) \in \mathcal{L}, \tag{7b}$$

$$\left(\zeta_{ij}^{\mathrm{p}}\right)^2 + \left(\zeta_{ij}^{\mathrm{q}}\right)^2 \le \left(\bar{I}_{ij}\right)^2, \qquad \forall\, (i,j) \in \mathcal{L}, \tag{7c}$$

where ζ_{ij}^{p} and ζ_{ij}^{q} are auxiliary decision variables and $\beta \in (0,1)$ is a given parameter that is typically set as 0.5. The resulting constraints (7a) and (7b) are in the same form as (6a)–(6c), and (7c) is a deterministic convex quadratic constraint.

Therefore, the mathematical formulation of TCC-ACOPF can be described as follows:

$$\min_{\substack{\boldsymbol{p},\boldsymbol{q},\boldsymbol{v},\boldsymbol{\theta} \\ \zeta^{\mathrm{p}},\zeta^{\mathrm{q}},\boldsymbol{\alpha},\boldsymbol{\gamma}}} \ \sum_{g \in \mathcal{G}} \mathbb{E}\big[c_g(p_g(\boldsymbol{\xi}))\big] \tag{8a}$$

s.t. (6a), (6b), (6c), (7a), (7b), (7c), (1g),

$$\hat{\varphi}\big(\boldsymbol{p}(\boldsymbol{\xi}), \boldsymbol{q}(\boldsymbol{\xi}), \boldsymbol{v}(\boldsymbol{\xi}), \boldsymbol{\theta}(\boldsymbol{\xi}) \,|\, \hat{\boldsymbol{p}}, \hat{\boldsymbol{q}}, \hat{\boldsymbol{v}}, \hat{\boldsymbol{\theta}}\big) = 0, \ \ \forall\, \boldsymbol{\xi}, \tag{8b}$$

where $\mathbb{E}$ denotes the expectation with respect to the distribution of $\boldsymbol{\xi}$, objective function (8a) minimizes the expected total operating cost, and constraints (8b) describe the linearized version of constraints (2) and (3). This linearization follows the existing studies (Fu & McCalley, 2001; Lubin et al., 2019; Roald & Andersson, 2017). Specifically, constraints (8b) are obtained from a Taylor expansion of (2) and (3) around a feasible solution (denoted by $(\hat{\boldsymbol{p}}, \hat{\boldsymbol{q}}, \hat{\boldsymbol{v}}, \hat{\boldsymbol{\theta}})$) to the deterministic ACOPF model (1). As constraints (8b) hold for all the possible realizations of $\boldsymbol{\xi}$, these constraints ensure that the solution of (8) satisfies the linearized power balance equations under all possible realizations of uncertainty. These existing studies show that this approximation is very tight (see Lemma 1 in Lubin et al. (2019) and the numerical results therein). Note that, (6a)–(6c) and (7a)–(7b) represent all the TCCs in our model, and they are nonlinear and nonconvex in general. However, the inner stochastic constraints of the TCCs are all defined by functions which are affine in decision variables and affine in random variables. In the following section, a tractable approximation for such two-sided chance-constraints is developed.

3. Reformulation techniques for TCC

In this section, we develop a tractable convex formulation of (8) by approximating the two-sided chance constraints (6a)–(6c) and (7a)–(7b) by a series of linear and SOC constraints. To that end, we consider a general form of TCC as follows:

$$\mathbb{P}\big(l_b \le \boldsymbol{h}_1(\boldsymbol{x})^{\top}\boldsymbol{\xi} + h_0(\boldsymbol{x}) \le u_b\big) \ge 1 - \epsilon, \tag{9}$$

where $\boldsymbol{h}_1(\boldsymbol{x})$ and $h_0(\boldsymbol{x})$ are affine functions of a vector of decision variables $\boldsymbol{x}$, l_b and u_b are lower and upper bounds of the TCC, respectively, and $\boldsymbol{\xi}$ represents a vector of random variables. For ease of exposition, we define $h(\boldsymbol{x}, \boldsymbol{\xi}) := \boldsymbol{h}_1(\boldsymbol{x})^{\top}\boldsymbol{\xi} + h_0(\boldsymbol{x})$, which is affine in $\boldsymbol{x}$.

In general, TCCs are very difficult to solve. As a result, many existing studies, including Bienstock et al. (2014), Roald & Andersson

(2017), and Dall'Anese et al. (2017), approximate the TCC (9) by two one-sided chance constraints as follows:

$$\mathbb{P}\big(h(\boldsymbol{x},\boldsymbol{\xi}) \le u_b\big) \ge 1-\hat{\epsilon}, \tag{10a}$$

$$\mathbb{P}\big(h(\boldsymbol{x},\boldsymbol{\xi}) \ge l_b\big) \ge 1-\hat{\epsilon}. \tag{10b}$$

When $\hat{\epsilon}=\epsilon$, it is clear that (10) provides an outer approximation of (9). When $\hat{\epsilon}=\epsilon/2$, (10) provides an inner approximation for (9), which is known as the Bonferroni approximation of TCCs (Hanasusanto, Roitch, Kuhn, & Wiesemann, 2017; Nemirovski & Shapiro, 2007).

The approximation (10) is inexact and often very weak (Xie & Ahmed, 2017), and it is accordingly followed by further studies seeking better approximations. For instance, Lubin et al. (2015) developed an SOC approximation of a TCC, in which $\boldsymbol{\xi}$ follows Gaussian distribution with known mean and covariance. In many practical settings, however, the Gaussian distribution is known to lack accuracy in modeling uncertainty, especially when it comes to skewness in the distribution, which though is commonly present in power system applications (Hodge et al., 2012). In the following Section 3.1, we provide a tight approximation for a general TCC under a GM distribution, which allows for significantly more accurate modeling of uncertainties. Our approximation has a controllable degree of accuracy, which is of high value as it provides an additional degree of flexibility to the model.

3.1. Tight approximation

In this section, we show that TCC (9) under GM distributions can be inner approximated by a set of convex constraints. A sufficient condition is also provided under which the approximation becomes exact. Moreover, the resulting convex constraints can be efficiently approximated by a set of linear and SOC constraints, which converges to the feasibility set of the original TCC (9). For notational brevity, we define $[n]=\{1,2,\ldots,n\}$ for any positive integer number n. We make the following assumption throughout this paper.

Assumption 1. The random variable $\boldsymbol{\xi}$ follows a GM distribution with K components as follows:

$$\boldsymbol{\xi} \sim \sum_{k=1}^{K} w_k \mathcal{N}(\boldsymbol{\mu}_k, \eta_k \Sigma), \tag{11}$$

where component $k\in[K]$ is a Gaussian distribution with mean $\boldsymbol{\mu}_k$, positive definite covariance $\eta_k\Sigma$, and weight $w_k>0$ such that $\sum_{k=1}^{K} w_k = 1$.

We use F to denote the CDF of $\boldsymbol{\xi}$ and accordingly have $F(\boldsymbol{\xi}) = \sum_{k=1}^{K} w_k F_k(\boldsymbol{\xi})$, where F_k denotes the corresponding CDF of each component $k\in[K]$. Note that the covariance of each component is a positive multiple of $\Sigma\succ 0$, which is the covariance matrix base for all components. Moreover, if the distribution of $\boldsymbol{\xi}$ only consists of one Gaussian distribution, i.e., $K=1$, the uncertainty distribution simplifies to the Gaussian distribution.

Using the GM distribution to represent the uncertainty, the following proposition, in which we develop a convex reformulation of the TCC, is key to our main results. Our results are related to the perspective function of the standard normal CDF $\Phi(\cdot)$. In particular, the perspective function of Φ, see Combettes (2016), is defined as

$$\tilde{\Phi}:\mathbb{R}\times[0,\infty)\to\mathbb{R}, \quad \tilde{\Phi}(z,\lambda) := \begin{cases} \lambda\,\Phi(z/\lambda), & \text{if } \lambda>0 \\ 0, & \text{if } \lambda=0. \end{cases}$$

Since Φ is a bounded continuous function, it is clear that $\tilde{\Phi}$ is also continuous. Also, since Φ is a continuous concave function on $[0,\infty)$, $\tilde{\Phi}$ is concave on $[0,\infty)\times[0,\infty)$ (Combettes, 2016). For ease of exposition, we assume $0\cdot\Phi(z/0)=0$ for all $z\in\mathbb{R}$ without mentioning the perspective function. We also define $\mu'_k(\boldsymbol{x}) := \boldsymbol{h}_1(\boldsymbol{x})^\top\boldsymbol{\mu}_k + h_0(\boldsymbol{x})$ and $\Sigma'(\boldsymbol{x}) := \sqrt{\boldsymbol{h}_1(\boldsymbol{x})^\top\Sigma\,\boldsymbol{h}_1(\boldsymbol{x})}$ for the rest of the paper.

Proposition 1. *Under Assumption 1, the two-sided chance constraint (9) can be inner approximated by the following convex constraints in $(\boldsymbol{x},\lambda)$:*

$$\lambda\left(\sum_{k=1}^{K} w_k\left(\Phi\left(\frac{u_b-\mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\lambda}\right)+\Phi\left(\frac{\mu'_k(\boldsymbol{x})-l_b}{\sqrt{\eta_k}\lambda}\right)\right)\right) \ge \lambda(2-\epsilon), \tag{12a}$$

$$l_b \le \mu'_k(\boldsymbol{x}) \le u_b, \ \ \forall\, k\in[K], \tag{12b}$$

$$\Sigma'(\boldsymbol{x}) \le \lambda, \tag{12c}$$

where $\Phi(\cdot)$ is the standard normal CDF and $\lambda\in\mathbb{R}$ is a nonnegative auxiliary variable. Moreover, if $\epsilon\le(1/2)\min\{w_1,\ldots,w_K\}$, then the approximation is exact; that is, $\boldsymbol{x}$ satisfies (9) if and only if there exists $\lambda\in\mathbb{R}$ such that $(\boldsymbol{x},\lambda)$ satisfies (12).

Proof. First, we show the convexity of constraints (12). Note that $\Phi(\cdot)$ is concave in the restricted domain $[0,\infty)$, which is enforced by (12b). As discussed above, its perspective function $\tilde{\Phi}(z,\lambda)$ is concave over $[0,\infty)\times[0,\infty)$. Since affine substitution and nonnegative weighted summation maintain concavity, (12a) is a convex constraint in $(\boldsymbol{x},\lambda)$. It is easy to see that (12b) and (12c) are affine and SOC constraints, respectively, and thus are convex.

Next, we show that (12) is an inner approximation of the TCC (9). We consider two possible cases: $\Sigma'(\boldsymbol{x})=0$ and $\Sigma'(\boldsymbol{x})>0$. (i) When $\Sigma'(\boldsymbol{x})=0$, we have $\boldsymbol{h}_1(\boldsymbol{x})=\boldsymbol{0}$ because $\Sigma\succ 0$ and accordingly (9) is reduced to (12b). (ii) When $\Sigma'(\boldsymbol{x})>0$, under Assumption 1, we have

$$\begin{aligned}
&\mathbb{P}_F\big(l_b\le h(\boldsymbol{x},\boldsymbol{\xi})\le u_b\big)\\
&=\sum_{k=1}^{K} w_k\mathbb{P}_{F_k}\big(l_b\le h(\boldsymbol{x},\boldsymbol{\xi})\le u_b\big)\\
&=\sum_{k=1}^{K} w_k\mathbb{P}_{F_k}\left(\frac{l_b-\mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\le\frac{h(\boldsymbol{x},\boldsymbol{\xi})-\mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\le\frac{u_b-\mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right)\\
&=\sum_{k=1}^{K} w_k\left(\Phi\left(\frac{u_b-\mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right)-\Phi\left(\frac{l_b-\mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right)\right)\\
&=\sum_{k=1}^{K} w_k\left(\Phi\left(\frac{u_b-\mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right)+\Phi\left(\frac{\mu'_k(\boldsymbol{x})-l_b}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right)\right)-1,
\end{aligned}$$

where the last equation holds because $\Phi(x)=1-\Phi(-x)$. As a result, (9) can be recast as

$$\sum_{k=1}^{K} w_k\left(\Phi\left(\frac{u_b-\mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right)+\Phi\left(\frac{\mu'_k(\boldsymbol{x})-l_b}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right)\right)\ge 2-\epsilon. \tag{13}$$

Consider $(\boldsymbol{x},\lambda)$ that satisfies constraints (12). Because $\Phi(\cdot)$ is an increasing function, $l_b\le\mu'_k(\boldsymbol{x})\le u_b$, and $\lambda\ge\Sigma'(\boldsymbol{x})>0$, we conclude that

$$\Phi\left(\frac{u_b-\mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right)\ge\Phi\left(\frac{u_b-\mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\lambda}\right) \quad\text{and}\quad \Phi\left(\frac{\mu'_k(\boldsymbol{x})-l_b}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right)\ge\Phi\left(\frac{\mu'_k(\boldsymbol{x})-l_b}{\sqrt{\eta_k}\lambda}\right),$$

which, together with (12a), implies (13). Thus, for any $(\boldsymbol{x},\lambda)$ that satisfies constraints (12) under both the above two cases, the $\boldsymbol{x}$ satisfies (9).

Finally, we show the equivalence between (12) and (9) when $\epsilon \le (1/2)\min\{w_1,\dots,w_K\}$. (i) Suppose that $\boldsymbol{x}$ satisfies constraint (9) with $\boldsymbol{h}_1(\boldsymbol{x}) = \boldsymbol{0}$. Then, $\Sigma'(\boldsymbol{x}) = 0$ and $h(\boldsymbol{x},\boldsymbol{\xi}) = h_0(\boldsymbol{x}) = \mu'_k(\boldsymbol{x})$ for all $k \in [K]$. By setting $\lambda = 0$, we see that constraints (12) are satisfied by $(\boldsymbol{x},\lambda)$. (ii) Suppose that $\boldsymbol{x}$ satisfies constraint (9) with $\boldsymbol{h}_1(\boldsymbol{x}) \neq \boldsymbol{0}$. Then (13) is satisfied by $\boldsymbol{x}$. We claim that such an $\boldsymbol{x}$ satisfies (12b), that is, for all $k \in [K]$,

$$\Phi\left(\frac{u_b - \mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right) \ge 0.5 \quad \text{and} \quad \Phi\left(\frac{\mu'_k(\boldsymbol{x}) - l_b}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right) \ge 0.5.$$

If not, then there exists some $\kappa \in [K]$ such that

$$\Phi\left(\frac{u_b - \mu'_\kappa(\boldsymbol{x})}{\sqrt{\eta_\kappa}\Sigma'(\boldsymbol{x})}\right) + \Phi\left(\frac{\mu'_\kappa(\boldsymbol{x}) - l_b}{\sqrt{\eta_\kappa}\Sigma'(\boldsymbol{x})}\right) < 1.5$$

because both $\Phi\left(\frac{u_b-\mu'_\kappa(\boldsymbol{x})}{\sqrt{\eta_\kappa}\Sigma'(\boldsymbol{x})}\right)$ and $\Phi\left(\frac{\mu'_\kappa(\boldsymbol{x})-l_b}{\sqrt{\eta_\kappa}\Sigma'(\boldsymbol{x})}\right)$ are no larger than 1 and one of them is less than 0.5 by contradiction. It further follows that,

$$\sum_{k=1}^{K} w_k\left(\Phi\left(\frac{u_b - \mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right) + \Phi\left(\frac{\mu'_k(\boldsymbol{x}) - l_b}{\sqrt{\eta_k}\Sigma'(\boldsymbol{x})}\right)\right)$$
$$< \sum_{k\neq\kappa}(2w_k) + 1.5w_\kappa = 2\sum_{k=1}^{K} w_k - 0.5w_\kappa = 2 - 0.5w_\kappa \le 2 - \epsilon,$$

which contradicts (13). Therefore, for any $\boldsymbol{x}$ satisfying (13), by setting $\lambda = \Sigma'(\boldsymbol{x})$, we see that constraints (12) are satisfied by $(\boldsymbol{x},\lambda)$. □

Remark 3. The exactness condition in Proposition 1, i.e., $\epsilon \le (1/2)\min\{w_1,\dots,w_K\}$, is always satisfied in the case of simple Gaussian ($K = 1$) if $\epsilon \le 0.5$. Hence, our proof recovers and generalizes the convexity results in Lubin et al. (2015).

Remark 4. The exactness condition in Proposition 1 is usually satisfied in practice where smaller values of ϵ are of interest. For example, in the case where $w_k = 1/K$ for each $k \in [K]$, our reformulation is exact for $\epsilon \le 1/(2K)$. That means a mixture of 4 Gaussians with $\epsilon \le 0.125$ is convex and can be exactly reformulated by Proposition 1.

Although (12) consists of convex constraints, it can be computationally expensive to solve for large-scale applications. In the rest of this section, we discuss how to approximate (12) using linear and SOC constraints. By Proposition 1, the domain of $\Phi(\cdot)$ can be restricted to $[0,\infty)$ in our model (12) without loss of generality, on which $\Phi(\cdot)$ is concave. As a result, a PWL inner approximation of $\Phi(\cdot)$ on the interval $[0,\infty)$ can be found by a modified linear interpolation. Specifically, we call a vector $\boldsymbol{t} = (t_0, t_1, \cdots, t_M)^\top \in \mathbb{R}^{M+1}$ a *valid interpolation vector* if $0 = t_0 < t_1 < \cdots < t_M$. For a valid interpolation vector $\boldsymbol{t} \in \mathbb{R}^{M+1}$, we define $\hat{\Phi}^{\boldsymbol{t}}_M(z) := \min_{m\in[M+1]}\left\{a^{\boldsymbol{t}}_m z + b^{\boldsymbol{t}}_m\right\}$, the PWL function from $[0,\infty)$ to $\mathbb{R}$ with $M+1$ segments such that

$$\begin{cases} a^{\boldsymbol{t}}_m t_{m-1} + b^{\boldsymbol{t}}_m = \Phi(t_{m-1}), \\ a^{\boldsymbol{t}}_m t_m + b^{\boldsymbol{t}}_m = \Phi(t_m), \end{cases} \qquad \forall\, m \in [M],$$

$a^{\boldsymbol{t}}_{M+1} = 0$, and $b^{\boldsymbol{t}}_{M+1} = \Phi(t_M)$, as illustrated in Fig. 1. The following lemma about $\hat{\Phi}^{\boldsymbol{t}}_M(\cdot)$ is straightforward due to the concavity of $\Phi(\cdot)$, and the proof is omitted.

Lemma 1. *For any valid interpolation vector $\boldsymbol{t} \in \mathbb{R}^{M+1}$, $\hat{\Phi}^{\boldsymbol{t}}_M(\cdot)$ is a concave increasing PWL function over $[0,\infty)$, and $\hat{\Phi}^{\boldsymbol{t}}_M(z) \le \Phi(z)$ for all $z \in [0,\infty)$.*

Replacing the standard normal CDF $\Phi(\cdot)$ with a PWL function $\hat{\Phi}^{\boldsymbol{t}}_M(\cdot)$, we have an inner approximation of (12a):

$$\lambda\left(\sum_k w_k\left(\hat{\Phi}^{\boldsymbol{t}}_M\left(\frac{u_b - \mu'_k(\boldsymbol{x})}{\sqrt{\eta_k}\lambda}\right) + \hat{\Phi}^{\boldsymbol{t}}_M\left(\frac{\mu'_k(\boldsymbol{x}) - l_b}{\sqrt{\eta_k}\lambda}\right)\right)\right) \ge \lambda(2-\epsilon). \tag{14}$$

Let $\Delta^{\boldsymbol{t}}_m := t_m - t_{m-1}$ for each $m \in [M]$, $\Delta^{\boldsymbol{t}} := \max_{m\in[M]}\Delta^{\boldsymbol{t}}_m$, and $E(\boldsymbol{t}) := \sup_{z\ge 0}\{\Phi(z) - \hat{\Phi}^{\boldsymbol{t}}_M(z)\}$. We are interested in the asymptotic behavior of the approximation as $E(\boldsymbol{t}) \to 0$. To that end, we introduce the following definition.

Definition 1. Let $\{\boldsymbol{t}^r\}_r := \{\boldsymbol{t}^1, \boldsymbol{t}^2, \dots,\}$ be a sequence of valid interpolation vectors where $\boldsymbol{t}^r = (t^r_0, t^r_1, \dots, t^r_{M_r})^\top$. We call $\{\boldsymbol{t}^r\}_r$ a "fine" sequence if $\hat{\Phi}^{\boldsymbol{t}^r}_{M_r}(z) \to \Phi(z)$ as $r \to \infty$ for all $z \in [0,\infty)$.

In other words, the choice of valid interpolation vectors in a "fine" sequence guarantees pointwise convergence of $\{\hat{\Phi}^{\boldsymbol{t}^r}_{M_r}\}_r$ to Φ on $[0,\infty)$. The following lemma shows that a "fine" sequence converges not only pointwisely but also uniformly.

Lemma 2. *Let $f : [a,\infty) \to [b,c)$ be an increasing surjective function. For each $r \ge 0$, let $f_r : [a,\infty) \to [b,c)$ be an increasing function such that $f_r(z) \le f(z)$ for all $z \in [a,\infty)$. If $\lim_{r\to\infty} f_r(z) = f(z)$ for every $z \in [a,\infty)$, then f_r converges to f uniformly. That is, for any $\epsilon > 0$, there exists $R \ge 0$ such that $|f_r(z) - f(z)| < \epsilon$ for all $r \ge R$ and $z \in [a,\infty)$.*

Proof. For any $\epsilon > 0$, let $M := \lceil 2(c-b)/\epsilon\rceil - 1$. Since f is an increasing surjective function, $f(a) = b$ and f is invertible. Let $s_i := f^{-1}(b + i\epsilon/2)$ for $i = 0,\dots,M$. Then, $f(s_{i+1}) - f(s_i) = \epsilon/2$ for $i = 0,\dots,M-1$, and $f(s_M) \ge c - \epsilon/2$. For each $i = 0,\dots,M$, since $\lim_{r\to\infty} f_r(s_i) = f(s_i)$, there exists $R_i \ge 0$ such that $f(s_i) - f_r(s_i) < \epsilon/2$ for all $r \ge R_i$. Due to the monotonicity of f and f_r, for any $z \ge a = s_0$, if $z \in [s_i, s_{i+1}]$ for some $i = 0,\dots,M-1$, then for all $r \ge R_i$,

$$f(z) - f_r(z) \le f(s_{i+1}) - f_r(s_i) = (f(s_{i+1}) - f(s_i)) + (f(s_i) - f_r(s_i)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon;$$

if $z \in [s_M,\infty)$, then for all $r \ge R_M$,

$$f(z) - f_r(z) \le c - f_r(s_M) = (c - f(s_M)) + (f(s_M) - f_r(s_M)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Let $R := \max_{i\in\{0,\dots,M\}}\{R_i\}$. We see that $0 \le f(z) - f_r(z) < \epsilon$ for all $r \ge R$ and $z \in [a,\infty)$. □

With Lemma 1 and Lemma 2, it is clear that a sequence of valid interpolation vectors $\{\boldsymbol{t}^r\}_r$ is "fine" if and only if $E(\boldsymbol{t}^r) \to 0$ as $r \to \infty$.

Note that it may be inconvenient to check the convergence of a sequence to determine whether it is "fine." The following lemma provides an easy-to-check sufficient condition.

Lemma 3. *A sequence of valid interpolation points $\{\boldsymbol{t}^r\}_r$ is "fine" if $\Delta^{\boldsymbol{t}^r} \to 0$ and $t^r_{M_r} \to \infty$ as $r \to \infty$.*

Proof. For any valid interpolation vector $\boldsymbol{t} \in \mathbb{R}^{M+1}$, note that

$$E(\boldsymbol{t}) \le \max\left\{\max_{m\in[M]}\{\Phi(t_m) - \Phi(t_{m-1})\},\ 1 - \Phi(t_M)\right\}.$$

For any small $0 < \epsilon < \frac{1}{2}$, since $t^r_{M_r} \to \infty$ as $r \to \infty$, there is $\bar{r} > 0$ such that $t^r_{M_r} > \Phi^{-1}(1-\epsilon)$ for all $r \ge \bar{r}$. That is,

$$1 - \Phi(t^r_{M_r}) < \epsilon, \qquad \forall\, r \ge \bar{r}.$$

Also, since $\Delta^{\boldsymbol{t}^r} \to 0$ as $r \to \infty$, there is $\hat{r} > 0$ such that $\Delta^{\boldsymbol{t}^r} < \epsilon/\Phi'(0) = \epsilon\sqrt{2\pi}$ for all $r \ge \hat{r}$, where $\Phi'(\cdot)$ is the derivative function of $\Phi(\cdot)$. Due to the concavity and monotonicity of Φ over

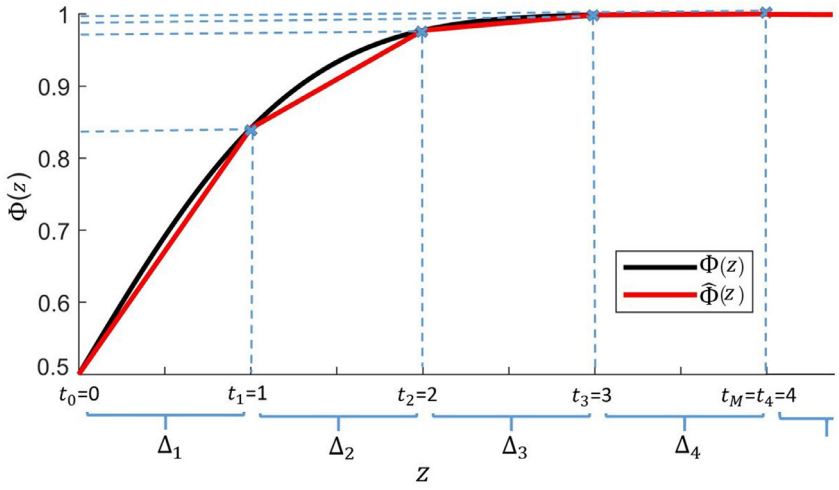

Fig. 1. An example of PWL fitting of the standard normal CDF ($M = 4$).

$[0, \infty)$, for all $r \geq \hat{r}$,

$$\Phi(t_m^r) - \Phi(t_{m-1}^r) \leq \Phi'(t_{m-1}^r)(t_m^r - t_{m-1}^r) \leq \Phi'(0)\Delta^{t^r} < \epsilon \qquad \forall\, m \in [M_r].$$

Consequently, $E(\boldsymbol{t}^r) < \epsilon$ for all $r \geq \max\{\bar{r}, \hat{r}\}$, which completes the proof. □

As an example, the sequence $\{\boldsymbol{t}^r\}_r$, where $\boldsymbol{t}^r = (0, \frac{1}{r}, \frac{2}{r}, \cdots, \frac{r^2}{r})^\top \in \mathbb{R}^{r^2+1}$, is a "fine" sequence. Another example of a "fine" sequence is $\{\boldsymbol{t}^r\}_r$, where $\boldsymbol{t}^r = (0, \Phi^{-1}(\frac{r+1}{2r}), \Phi^{-1}(\frac{r+2}{2r}), \cdots, \Phi^{-1}(\frac{2r-1}{2r}))^\top \in \mathbb{R}^r$. Both of the examples can be checked easily using Lemma 3. We remark here that Lemma 3 is not a necessary condition. For example, the sequence $\{\boldsymbol{t}^r\}_r$, where $\boldsymbol{t}^r = (0, \frac{1}{r}, \frac{1}{r} + \frac{1}{r-1}, \cdots, \frac{1}{r} + \frac{1}{r-1} + \cdots + \frac{1}{2} + 1)^\top \in \mathbb{R}^{r+1}$, is a "fine" sequence by the definition — we leave the details to the readers — however, $\Delta^{\boldsymbol{t}^r} = 1$ for all $r > 0$, which does not satisfy the condition in Lemma 3.

When $\{\boldsymbol{t}^r\}_r$ is a "fine" sequence, constraint (14) converges to (12a) in the following sense.

Proposition 2. *Suppose that $f(\cdot,\cdot)$ is a continuous function of $(\boldsymbol{x}, \lambda)$. Let $(\boldsymbol{x}^*, \lambda^*)$ be an optimal solution of*

$$\inf\{f(\boldsymbol{x}, \lambda) | (\boldsymbol{x}, \lambda) \text{ satisfies (12)}\}, \tag{15}$$

and let $(\boldsymbol{x}_{\boldsymbol{t}}^*, \lambda_{\boldsymbol{t}}^*)$ be an optimal solution of

$$\inf\{f(\boldsymbol{x}, \lambda) | (\boldsymbol{x}, \lambda) \text{ satisfies (14), (12b), (12c)}\}. \tag{16}$$

Then, we have $f(\boldsymbol{x}_{\boldsymbol{t}}^*, \lambda_{\boldsymbol{t}}^*) \geq f(\boldsymbol{x}^*, \lambda^*)$ for any valid interpolation vector $\boldsymbol{t}$. Moreover, if there exists $(\bar{\boldsymbol{x}}, \bar{\lambda})$ such that (12) are satisfied and (12a) is strictly satisfied, then

$$\lim_{r\to\infty} f(\boldsymbol{x}_{\boldsymbol{t}^r}^*, \lambda_{\boldsymbol{t}^r}^*) = f(\boldsymbol{x}^*, \lambda^*)$$

for any "fine" sequence $\{\boldsymbol{t}^r\}_r$.

The proof of Proposition 2 relies on the following technical lemma.

Lemma 4. *Let $X \subseteq \mathbb{R}^n$ be a nonempty convex set, $\Psi : \mathbb{R}^n \to \mathbb{R}$ be a concave function over X, and $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. Suppose that $\{\Psi_r\}_r$ is a sequence of functions from $\mathbb{R}^n$ to $\mathbb{R}$ such that for any $\boldsymbol{y} \in X$,*

$$\Psi_r(\boldsymbol{y}) \leq \Psi(\boldsymbol{y}) \quad \text{and} \quad \lim_{r\to\infty} \Psi_r(\boldsymbol{y}) = \Psi(\boldsymbol{y}).$$

Let $\boldsymbol{y}^*$ be an optimal solution of $\inf\{f(\boldsymbol{y}) | \Psi(\boldsymbol{y}) \geq 0, \boldsymbol{y} \in X\}$, and let $\boldsymbol{y}_r^*$ be an optimal solution of $\inf\{f(\boldsymbol{y}) | \Psi_r(\boldsymbol{y}) \geq 0, \boldsymbol{y} \in X\}$. If there exists $\bar{\boldsymbol{y}} \in X$ such that $\Psi(\bar{\boldsymbol{y}}) > 0$, then $f(\boldsymbol{y}_r^*) \geq f(\boldsymbol{y}^*)$ for all $r \geq 0$, and $\lim_{r\to\infty} f(\boldsymbol{y}_r^*) = f(\boldsymbol{y}^*)$.

Proof. For any $\epsilon > 0$, since $f(\cdot)$ is continuous, there exists $\sigma_1 \in (0, \|\bar{\boldsymbol{y}} - \boldsymbol{y}^*\|)$ such that $|f(\boldsymbol{y}) - f(\boldsymbol{y}^*)| < \epsilon$ whenever $\|\boldsymbol{y} - \boldsymbol{y}^*\| \leq \sigma_1$. Let

$$\hat{\boldsymbol{y}} := \sigma_1 \frac{\bar{\boldsymbol{y}} - \boldsymbol{y}^*}{\|\bar{\boldsymbol{y}} - \boldsymbol{y}^*\|} + \boldsymbol{y}^* = \frac{\sigma_1}{\|\bar{\boldsymbol{y}} - \boldsymbol{y}^*\|}\bar{\boldsymbol{y}} + \left(1 - \frac{\sigma_1}{\|\bar{\boldsymbol{y}} - \boldsymbol{y}^*\|}\right)\boldsymbol{y}^*,$$

which represents a convex combination of $\bar{\boldsymbol{y}}$ and $\boldsymbol{y}^*$. It follows that $\|\hat{\boldsymbol{y}} - \boldsymbol{y}^*\| \leq \sigma_1$, implying $f(\hat{\boldsymbol{y}}) - f(\boldsymbol{y}^*) < \epsilon$ due to the continuity of $f(\cdot)$, and that $\hat{\boldsymbol{y}} \in X$ due to the convexity of X. Moreover, by the concavity of $\Psi(\cdot)$ over X, we have

$$\Psi(\hat{\boldsymbol{y}}) \geq \frac{\sigma_1}{\|\bar{\boldsymbol{y}} - \boldsymbol{y}^*\|}\Psi(\bar{\boldsymbol{y}}) + \left(1 - \frac{\sigma_1}{\|\bar{\boldsymbol{y}} - \boldsymbol{y}^*\|}\right)\Psi(\boldsymbol{y}^*) > 0.$$

Since $\lim_{r\to\infty} \Psi_r(\hat{\boldsymbol{y}}) = \Psi(\hat{\boldsymbol{y}}) > 0$, there exists $R \geq 0$ such that $\Psi_r(\hat{\boldsymbol{y}}) > 0$ for any $r \geq R$. It follows that $\hat{\boldsymbol{y}}$ is a feasible solution to the minimization problem $\inf\{f(\boldsymbol{y}) | \Psi_r(\boldsymbol{y}) \geq 0, \boldsymbol{y} \in X\}$ for any $r \geq R$, and accordingly $f(\boldsymbol{y}_r^*) \leq f(\hat{\boldsymbol{y}})$. Therefore, we can conclude that $f(\boldsymbol{y}_r^*) - f(\boldsymbol{y}^*) \leq f(\hat{\boldsymbol{y}}) - f(\boldsymbol{y}^*) < \epsilon$ for any $r \geq R$. That is, we have

$$\limsup_{r\to\infty} f(\boldsymbol{y}_r^*) \leq f(\boldsymbol{y}^*). \tag{17}$$

In addition, for any $r \geq 0$, since $\Psi_r(\boldsymbol{y}) \leq \Psi(\boldsymbol{y})$ for any $\boldsymbol{y} \in X$, we have $\inf\{f(\boldsymbol{y}) | \Psi(\boldsymbol{y}) \geq 0, \boldsymbol{y} \in X\}$ is a relaxation of $\inf\{f(\boldsymbol{y}) | \Psi_r(\boldsymbol{y}) \geq 0, \boldsymbol{y} \in X\}$. Therefore, we conclude that for any $r > 0$,

$$f(\boldsymbol{y}_r^*) - f(\boldsymbol{y}^*) \geq 0. \tag{18}$$

By combining (17) and (18), we complete the proof. □

Lemma 4 shows that for an optimization problem under the setting mentioned therein, its inner approximations converge to one formulation that has the same optimal value of the original problem. We prove this conclusion by creating a sequence of feasible regions (represented based on $\{\Psi_r\}_r$) contained in the original feasible region (represented based on Ψ), where this sequence converges to one that produces the optimal value of the original problem. With this lemma, we are now ready to prove Proposition 2.

Proof of Proposition. 2.. For any valid interpolation vector $\boldsymbol{t}$, Lemma 1 implies that (16) is an inner approximation of (15). Therefore, $f(\boldsymbol{x}_{\boldsymbol{t}}^*, \lambda_{\boldsymbol{t}}^*) \geq f(\boldsymbol{x}^*, \lambda^*)$ for any valid interpolation vector $\boldsymbol{t}$. Now consider a "fine" sequence $\{\boldsymbol{t}^r\}_r$. Let

$$\Psi(\boldsymbol{x}, \lambda) := \lambda\left(\sum_k w_k\left(\Phi\left(\frac{u_b - \mu_k'(\boldsymbol{x})}{\sqrt{\eta_k}\lambda}\right) + \Phi\left(\frac{\mu_k'(\boldsymbol{x}) - l_b}{\sqrt{\eta_k}\lambda}\right)\right) + \epsilon - 2\right),$$

$$\Psi_r(\boldsymbol{x}, \lambda) := \lambda\left(\sum_k w_k\left(\hat{\Phi}_{M_r}^r\left(\frac{u_b - \mu_k'(\boldsymbol{x})}{\sqrt{\eta_k}\lambda}\right) + \hat{\Phi}_{M_r}^r\left(\frac{\mu_k'(\boldsymbol{x}) - l_b}{\sqrt{\eta_k}\lambda}\right)\right) + \epsilon - 2\right),$$

and $X := \{(\boldsymbol{x}, \lambda) | (\boldsymbol{x}, \lambda)$ satisfies (12b), (12c)$\}$. By Proposition 1 and the existence of $(\bar{\boldsymbol{x}}, \bar{\lambda})$, we have that X is a nonempty convex set, $\Psi(\cdot, \cdot)$ is a concave function over X, and there exists $(\bar{\boldsymbol{x}}, \bar{\lambda}) \in X$ such that $\Psi(\bar{\boldsymbol{x}}, \bar{\lambda}) > 0$. Moreover, by Lemma 1 and Definition 1, it is easy to check that

$$\Psi_r(\boldsymbol{x}, \lambda) \le \Psi(\boldsymbol{x}, \lambda) \quad \text{and} \quad \lim_{r\to\infty} \Psi_r(\boldsymbol{x}, \lambda) = \Psi(\boldsymbol{x}, \lambda)$$

for all $(\boldsymbol{x}, \lambda) \in X$. Therefore, the proof is completed by a direct application of Lemma 4. □

Remark 5. The assumption about the existence of $(\bar{\boldsymbol{x}}, \bar{\lambda})$ in Proposition 2 is easy to satisfy in practice. In particular, if there exists $\bar{\boldsymbol{x}}$ such that $\mathbb{P}(l_b \le \boldsymbol{h}(\bar{\boldsymbol{x}}, \boldsymbol{\xi}) \le u_b) > 1 - \epsilon$ and $\Sigma'(\bar{\boldsymbol{x}}) > 0$, then under Assumption 1, (13) is satisfied strictly at $\bar{\boldsymbol{x}}$, and thus $(\bar{\boldsymbol{x}}, \sqrt{\boldsymbol{h}_1(\bar{\boldsymbol{x}})^\top \Sigma \boldsymbol{h}_1(\bar{\boldsymbol{x}})})$ satisfies (12) and satisfies (12a) strictly.

In summary, constraints (14), (12b), and (12c) provide an inner approximation of constraints (12), and more importantly, they guarantee asymptotic convergence in terms of optimal values. In the following proposition, we show that the set defined by (14) is in fact polyhedral.

Proposition 3. *Consider $(\boldsymbol{x}, \lambda)$ that satisfies (12b) and (12c). Then, $(\boldsymbol{x}, \lambda)$ satisfies (14) if and only if there exists $\boldsymbol{\pi} \in \mathbb{R}^{2K}$ such that $(\boldsymbol{x}, \lambda, \boldsymbol{\pi})$ satisfies the following linear constraints:*

$$\begin{cases} a_m^{\boldsymbol{t}}(u_b - \mu_k'(\boldsymbol{x})) + b_m^{\boldsymbol{t}}\sqrt{\eta_k}\lambda \ge \sqrt{\eta_k}\pi_{k,1}, \\ a_m^{\boldsymbol{t}}(\mu_k'(\boldsymbol{x}) - l_b) + b_m^{\boldsymbol{t}}\sqrt{\eta_k}\lambda \ge \sqrt{\eta_k}\pi_{k,2}, \end{cases} \quad \forall\, m \in [M+1],\ \forall\, k \in [K], \tag{19a}$$

$$\sum_{k=1}^{K} w_k(\pi_{k,1} + \pi_{k,2}) \ge \lambda(2 - \epsilon). \tag{19b}$$

Proof. First, if $(\boldsymbol{x}, \lambda, \boldsymbol{\pi})$ satisfies (19) for some $\boldsymbol{\pi} \in \mathbb{R}^{2K}$, then

$$\begin{aligned} \lambda \hat{\Phi}_M^{\boldsymbol{t}}\left(\frac{u_b - \mu_k'(\boldsymbol{x})}{\sqrt{\eta_k}\lambda}\right) &= \lambda \min_{m\in[M+1]} \left\{ a_m^{\boldsymbol{t}}\left(\frac{u_b - \mu_k'(\boldsymbol{x})}{\sqrt{\eta_k}\lambda}\right) + b_m^{\boldsymbol{t}} \right\} \\ &= \frac{1}{\sqrt{\eta_k}} \min_{m\in[M+1]} \left\{ a_m^{\boldsymbol{t}}(u_b - \mu_k'(\boldsymbol{x})) + b_m^{\boldsymbol{t}}\sqrt{\eta_k}\lambda \right\} \\ &\ge \pi_{k,1}, \end{aligned} \tag{20}$$

where the inequality holds due to (19a). With similar arguments, we have

$$\lambda \hat{\Phi}_M^{\boldsymbol{t}}\left(\frac{\mu_k'(\boldsymbol{x}) - l_b}{\sqrt{\eta_k}\lambda}\right) \ge \pi_{k,2}. \tag{21}$$

Combining (19b) and (20)-(21), we see that (14) is satisfied by $(\boldsymbol{x}, \lambda)$.

Next, if $(\boldsymbol{x}, \lambda)$ satisfies (14), then we can set

$$\pi_{k,1} := \hat{\Phi}_M^{\boldsymbol{t}}\left(\frac{u_b - \mu_k'(\boldsymbol{x})}{\sqrt{\eta_k}\lambda}\right) \quad \text{and} \quad \pi_{k,2} := \hat{\Phi}_M^{\boldsymbol{t}}\left(\frac{\mu_k'(\boldsymbol{x}) - l_b}{\sqrt{\eta_k}\lambda}\right)$$

for $k \in [K]$ if $\lambda > 0$. If $\lambda = 0$, we can set $\pi_{k,j} := 0$ for $k \in [K]$, $j \in \{1, 2\}$. It follows in both cases that $(\boldsymbol{x}, \lambda, \boldsymbol{\pi})$ satisfies (19). □

For any continuous function $f(\cdot)$ and any valid interpolation vector $\boldsymbol{t}$, let $f^* := \inf\{f(\boldsymbol{x}) | \boldsymbol{x}$ satisfies (9)$\}$ and $f_{\boldsymbol{t}}^* := \inf\{f(\boldsymbol{x}) | \boldsymbol{x}$ satisfies (19), (12b), (12c) with some $\lambda, \boldsymbol{\pi}\}$.
Note that (12c) is an SOC constraint as it is equivalent to $\|\Sigma^{1/2}\boldsymbol{h}_1(\boldsymbol{x})\| \le \lambda$, where $\Sigma^{1/2}$ is a symmetric positive definite matrix such that $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$. We summarize our results in this section in the following theorem.

Theorem 1. *Under Assumption 1, TCC (9) can be inner approximated by the set of linear constraints (19), (12b) and SOC constraint (12c). In addition, if $\epsilon \le (1/2)\min\{w_1, \ldots, w_K\}$ and (9) is strictly satisfied by some $\bar{\boldsymbol{x}}$, the inner approximation converges to (9) in the sense that $f_{\boldsymbol{t}^r}^* \to f^*$ as $r \to \infty$ for any "fine" sequence $\{\boldsymbol{t}^r\}_r$.*

Using the results from Theorem 1, one can provide an SOCP approximation of the TCC-ACOPF (8), where each of the TCCs (6a) – (6c) and (7a) – (7b) can be replaced with its counterpart as in (19), (12b), and (12c).

3.2. Selection of interpolation points

The quality of our SOCP approximation in Section 3.1 is related to the choice of the valid interpolation vector $\boldsymbol{t} = (t_0, \ldots, t_M)^\top$. Loosely speaking, a smaller maximum error $E(\boldsymbol{t}) = \max_{z\in[0,\infty)}\{\Phi(z) - \hat{\Phi}_M^{\boldsymbol{t}}(z)\}$ is likely to lead to a better SOCP approximation. The maximum error $E(\boldsymbol{t})$ depends on two factors: the number of interpolation points and where the interpolation points are positioned. Intuitively, more interpolation points (i.e., a larger M) often lead to more accurate approximation. Nevertheless, a larger M results in more constraints in (19) and thus a higher computational cost. For a given δ, we are interested in how to choose the minimum number of interpolation points such that $E(\boldsymbol{t}) \le \delta$. First, we provide the following lemma to help characterize the maximum error of a linear interpolation.

Lemma 5. *Given two interpolation points $b > a \ge 0$, let $\hat{\Phi}_{a,b}(\cdot)$ be the linear interpolation of $\Phi(\cdot)$ such that $\hat{\Phi}_{a,b}(a) = \Phi(a)$ and $\hat{\Phi}_{a,b}(b) = \Phi(b)$. The maximum error of the linear interpolation on $[a, b]$ is*

$$\begin{aligned} \max_{z\in[a,b]} \left\{\Phi(z) - \hat{\Phi}_{a,b}(z)\right\} = \Phi\left(\sqrt{-\ln(2\pi s^2)}\right) \\ - s\left(\sqrt{-\ln(2\pi s^2)} - a\right) - \Phi(a), \end{aligned} \tag{22}$$

where $s = (\Phi(b) - \Phi(a))/(b - a)$. Moreover, define two functions

$$E_a(b) = \bar{E}_b(a) := \max_{z\in[a,b]} \left\{\Phi(z) - \hat{\Phi}_{a,b}(z)\right\}.$$

Then, $E_a(\cdot)$ is an increasing continuous function on (a, ∞) and $\bar{E}_b(\cdot)$ is a decreasing continuous function on $(-\infty, b)$.

Proof. The linear interpolation of $\Phi(\cdot)$ on $[a, b]$ can be expressed as $\hat{\Phi}_{a,b}(z) = \Phi(a) + s(z - a)$. Since $\Phi(\cdot)$ is strictly concave on $[0, \infty)$, the maximum error on $[a, b]$, i.e.,

$$\max_{z\in[a,b]} \{\Phi(z) - s(z - a) - \Phi(a)\},$$

occurs at $z^* \in (a, b)$ such that $\Phi'(z^*) = s$, i.e., $(1/\sqrt{2\pi})e^{-\frac{(z^*)^2}{2}} = s$. Therefore, $z^* = \sqrt{-\ln(2\pi s^2)}$, and the maximum error is $\Phi(z^*) - s(z^* - a) - \Phi(a)$, which proves (22).

Now consider any $0 \le a < b < c$. Since $\Phi(\cdot)$ is strictly concave on $[0, \infty)$,

$$\frac{\Phi(b) - \Phi(a)}{b - a} > \frac{\Phi(c) - \Phi(a)}{c - a}.$$

Therefore, for any $z \in (a, b]$,

$$\begin{aligned} \hat{\Phi}_{a,b}(z) &= \Phi(a) + \frac{\Phi(b) - \Phi(a)}{b - a}(z - a) > \Phi(a) \\ &\quad + \frac{\Phi(c) - \Phi(a)}{c - a}(z - a) = \hat{\Phi}_{a,c}(z). \end{aligned}$$

As a result,

$$\begin{aligned} E_a(b) &= \max_{z\in(a,b)} \left\{\Phi(z) - \hat{\Phi}_{a,b}(z)\right\} < \max_{z\in(a,b)} \left\{\Phi(z) - \hat{\Phi}_{a,c}(z)\right\} \\ &\le \max_{z\in[a,c]} \left\{\Phi(z) - \hat{\Phi}_{a,c}(z)\right\} = E_a(c). \end{aligned}$$

That is, $E_a(\cdot)$ is an increasing function. Similarly, we can show that $\bar{E}_b(\cdot)$ is decreasing. By (22), it is clear that both $E_a(\cdot)$ and $\bar{E}_b(\cdot)$ are continuous. □

Based on Lemma 5, we propose Algorithm 1 to calculate the positions of interpolation points. Given a tolerance δ of the maximum error $E(\boldsymbol{t})$, Algorithm 1 finds a PWL approximation of $\Phi(\cdot)$ with the least interpolation points. The minimality of the number of interpolation points is guaranteed by the following theorem.

Algorithm 1: Interpolation points positioning.

Data: approximation tolerance $\delta > 0$.
Result: a valid interpolation vector $\boldsymbol{t} = (t_0, \ldots, t_M)^\top$ such that $E(\boldsymbol{t}) \le \delta$ and M is minimal.
$m \leftarrow 0$, $t_0 \leftarrow 0$;
while $\Phi(t_m) < 1 - \delta$ **do**
  use line search to find $t_{m+1} \in [t_m, \infty)$ such that $E_{t_m}(t_{m+1}) = \delta$;
  $m \leftarrow m + 1$;
$M \leftarrow m$;

Theorem 2. *For $\delta > 0$, let $\boldsymbol{t} = (t_0, \ldots, t_M)^\top$ be the output of Algorithm 1. Let $\hat{\Phi}^{\boldsymbol{t}'}_{M'}(\cdot)$ be a PWL approximation of $\Phi(\cdot)$ defined by a valid interpolation vector $\boldsymbol{t}' = (t'_0, \ldots, t'_{M'})^\top$. If $E(\boldsymbol{t}') \le \delta$, then $M' \ge M$.*

Proof. We prove by contradiction. Suppose that $M' < M$. We first show that $t'_m \le t_m$ for all $m \in [M']$. Note that for any $m \in [M']$, $E_{t'_{m-1}}(t'_m) = \max_{z \in [t'_{m-1}, t'_m]}\{\Phi(z) - \hat{\Phi}^{\boldsymbol{t}'}_{M'}(z)\} \le E(\boldsymbol{t}') \le \delta$ by Lemma 5. As $\boldsymbol{t}'$ is a valid interpolation vector, $t'_0 = 0$. Since

$$E_0(t_1) = E_{t_0}(t_1) = \delta \ge E(\boldsymbol{t}') \ge E_{t'_0}(t'_1) = E_0(t'_1)$$

and $E_0(\cdot)$ is increasing on $(0, \infty)$, we have $t'_1 \le t_1$. Now suppose that $t'_m \le t_m$ for some $m \in [M' - 1]$. Then,

$$E_{t'_m}(t_{m+1}) = \bar{E}_{t_{m+1}}(t'_m) \ge \bar{E}_{t_{m+1}}(t_m) = E_{t_m}(t_{m+1}) = \delta \ge E_{t'_m}(t'_{m+1}),$$

where the first inequality holds because $\bar{E}_{t_{m+1}}(\cdot)$ is decreasing on $(-\infty, t_{m+1})$. Since $E_{t'_m}(\cdot)$ is increasing on (t'_m, ∞), we have $t'_{m+1} \le t_{m+1}$. By induction, $t'_m \le t_m$ for all $m \in [M']$.

Since $M' < M$, $t'_{M'} \le t_{M'} \le t_{M-1}$. Therefore, $\Phi(t_{M-1}) \ge \Phi(t'_{M'})$. On the other hand, since $\hat{\Phi}^{\boldsymbol{t}'}_{M'}(z) = \Phi(t'_{M'})$ for all $z \in [t'_{M'}, \infty)$,

$$\delta \ge \Phi(z) - \hat{\Phi}^{\boldsymbol{t}'}_{M'}(z) = \Phi(z) - \Phi(t'_{M'}) \qquad \forall\, z \in [t'_{M'}, \infty).$$

Taking $z \to \infty$, we have $\Phi(t'_{M'}) \ge 1 - \delta$. Therefore, $\Phi(t_{M-1}) \ge \Phi(t'_{M'}) \ge 1 - \delta$, which contradicts the assumption that Algorithm 1 does not terminate at $M - 1$. □

It is worth noting that although our proposed PWL algorithm is tailored for the standard normal CDF in this paper, it can be generalized to approximate any other strictly monotone convex or concave functions that are bounded from at least one side.

4. Computational experiments

In this section, we implement the proposed TCC-ACOPF model and approximation from Theorem 1 on modified IEEE 30-bus and 118-bus test systems. First, in Section 4.1, we introduce the test systems, the real historical data that we collect, and the synthetic data that we create. Next, in Section 4.2, we present computational results to demonstrate the effectiveness of our proposed TCC-ACOPF model in comparison to the state-of-art methods. All computational experiments are performed on a PC with an Intel Core i7-7700 CPU and 16 GB RAM. We use JuMP in Julia (Dunning, Huchette, & Lubin, 2017) to implement all of the models. Ipopt solver (Wächter & Biegler, 2006) is used to solve the nonlinear deterministic ACOPF model (1), and Gurobi 9.0 solver is used for solving the SOCP formulations.

4.1. Data setting

Our modified IEEE 118-bus system is based on the original IEEE 118-bus system available online at MATPOWER (Zimmerman, Murillo-Sánchez, & Thomas, 2010), which includes 118 buses, 54 thermal generators, and 186 transmission lines. The following modifications are made: the value of $\bar{P}_g$ for each $g \in \mathcal{G}$ is reduced by 30% and the values of D_i^{p} and D_i^{q} for each $i \in \mathcal{B}$ are increased by 10%. We also include 11 wind farms, which provide about 33% of the total demand, and their forecast power outputs are listed in Table 1. Moreover, the standard IEEE 30-bus system has 30 buses, 6 thermal generators, and 41 transmission lines, and it is also available online at Zimmerman et al. (2010). We will consider different number of wind farms (leading to different penetration levels of renewable energy) in our modified IEEE 30-bus system in Section 4.2.3. In our computational experiments, we set the risk control parameters $\epsilon_{\mathrm{p}} = \epsilon_{\mathrm{q}} = \epsilon_{\mathrm{v}} = \epsilon_{\mathrm{L}}$ in (8) and use a single risk parameter ϵ to denote all of them thereafter.

We consider the renewable power generation forecast errors (i.e., $\boldsymbol{\xi}$) of the wind farms in the test systems to be uncertain. The real wind power outputs and the hour-ahead wind power forecasts from Wind Integration National Dataset Toolkit of National Renewable Energy Laboratory (NREL) are analyzed to obtain historical data on forecast errors (Draxl, Hodge, Clifton, & McCaa, 2015). Note that, the wind power forecast errors are then scaled based on the wind power capacity. In addition to the historical data, we also generate synthetic data that can reflect skewness in the forecast error distribution. Specifically, three synthetic datasets are generated and they are referred to as "Left-skewed," "Normally distributed," and "Right-skewed" datasets, respectively, with each of them having a size of $N = 20,000$ data samples and $|\mathcal{R}|$ variates, where $|\mathcal{R}|$ represents the cardinality of set $\mathcal{R}$, i.e., the total number of renewable wind generators in our studied test systems. For each dataset, we create three groups of data samples (referred to as "Group 1," "Group 2," and "Group 3" data samples) and merge them to create the entire dataset, as specified in the following:

- First, we randomly create an $|\mathcal{R}| \times |\mathcal{R}|$ correlation matrix ρ.
- Second, we create Group 1 data samples with size of $N/2$. Each data sample follows a multivariate normal distribution $\mathcal{N}(\breve{\boldsymbol{\mu}}_1, 0.1\rho^\top\rho)$, where each element of $\breve{\boldsymbol{\mu}}_1$ (i.e., $\breve{\mu}_{1r}$, $\forall r \in \mathcal{R}$) is randomly generated within the interval $[-0.05, 0.05]$. That is, most data samples are around zero.
- Third, we introduce a parameter $\varpi \in (0, 1)$, which is used to control the skewness of the dataset, and create Group 2 data samples with size of $(N/2)\varpi$. Each data sample follows a multivariate normal distribution $\mathcal{N}(\breve{\boldsymbol{\mu}}_2/(1-\varpi), \rho^\top\rho)$, where each element of $\breve{\boldsymbol{\mu}}_2$ is randomly generated within the interval $[-0.15, -0.05]$. That is, most data samples are below zero.
- Fourth, we create Group 3 data samples with size of $(N/2)(1-\varpi)$. Each data sample follows a multivariate normal distribution $\mathcal{N}(\breve{\boldsymbol{\mu}}_3/\varpi, \rho^\top\rho)$, where each element of $\breve{\boldsymbol{\mu}}_3$ is randomly generated within the interval [0.05,0.15]. That is, most data samples are above zero.

For the modified IEEE 118-bus system with $|\mathcal{R}| = 11$ wind farms, we set ϖ to be 0.7, 0.5, and 0.3 for creating the "Left-skewed," "Normally distributed," and "Right-skewed" datasets, respectively.

For each of the real and synthetic datasets described above, we can fit it with a Gaussian distribution ($K = 1$) and a GM distribution with two components ($K = 2$) separately, both of which

Table 1
Hourly wind power forecast (MW).

Bus i	3	8	11	20	24	26	31	38	43	49	53
$\sum_{g\in\mathcal{R}_i} r_g^{\mathrm{p}}(0)$	70	147	102	105	113	84	59	250	118	76	72

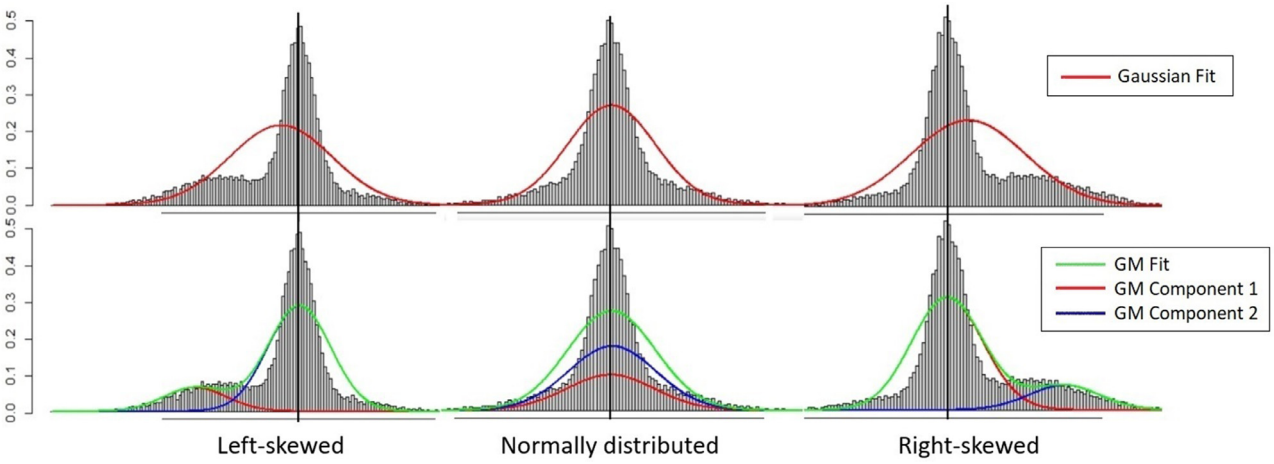

Fig. 2. Synthetic data alongside Gaussian and GM fits for forecast error of a wind farm.

are used in our computational experiments, thereby obtaining necessary parameters therein. It leads to "Gaussian Fit" and "GM Fit" correspondingly for each dataset. For the Gaussian Fit, the mean vectors and covariance matrices are simply calculated from the data. For the GM Fit, the weighted means and covariance are obtained using the *mvnormalmixEM* package in R, which fits a multivariate GM distribution to the data via an expectation-maximization algorithm (Benaglia, Chauveau, Hunter, & Young, 2010). We take a renewable generator as an example, and show the fitting results to the synthetic forecast errors of this generator based on the histogram of the data visually in Fig. 2. We can observe that both GM and Gaussian distributions fit similarly to the "Normally distributed" dataset, but the GM distribution fits much better for the "Left-skewed" and "Right-skewed" datasets.

4.2. Results and discussions

We implement our proposed TCC-ACOPF model and approximation approaches by considering (i) a Gaussian distribution ($K = 1$) and (ii) a GM distribution with two components ($K = 2$), leading to two specific models denoted by TCC-ACOPF-K1 and TCC-ACOPF-K2, respectively. Our approaches are compared with the one-sided chance-constrained model (denoted by "OCC-ACOPF"), where Lubin et al. (2019) assume that the active power, reactive power, and voltage limits are specified in one-sided chance constraints and the apparent flow constraints (7a)–(7b) are approximated by a TCC SOC formulation. The computational results are presented in the follow sequence: (i) comparisons are presented regarding the optimality and violation probability of each approach, as shown in Section 4.2.1; (ii) sensitivity analyses are performed to verify the accuracy and computational efficiency of our proposed approach, as shown in Section 4.2.2; (iii) the scalability of our proposed approach is discussed under various levels of renewable energy penetration, as shown in Section 4.2.3; and (iv) real case studies are provided to demonstrate the out-of-sample performance of the solutions generated by our proposed model and approaches, leading to real-world nonlinear ACOPF feasibility analyses in practice, as shown in Section 4.2.4.

4.2.1. Optimality and violation probability

First, we compare the computational performance of OCC-ACOPF, TCC-ACOPF-K1, and TCC-ACOPF-K2 over the synthetic data. The in-sample optimal values and out-of-sample violation probabilities are evaluated on the modified 118-bus system with $|\mathcal{R}| = 11$ wind farms. We set $\epsilon = 0.2$ and obtain 10 optimal segments via Algorithm 1 by setting $\delta = 0.002$ when performing the PWL approximation. In addition, for each of the three synthetic datasets, we randomly select 5,000 data samples therein to fit a Gaussian distribution ($K = 1$) and a GM distribution with two components ($K = 2$). We correspondingly solve all the three aforementioned models, recording the solutions and reporting the optimal values (represented by "Opt. Val. ($)" in Table 2).

Based on the solution induced by a model, we further evaluate its quality over the remaining 15,000 data samples by calculating the maximum violation probability (represented by "Vio. Prob." in Table 2) across all the nominal constraints in the model. Specifically, given an uncertainty realization and the obtained solution, we check whether each set of the nominal two-sided constraints on active power, reactive power, voltage, and power flow is violated or not. For example, given an uncertainty realization $\hat{\boldsymbol{\xi}}$, an optimal solution (p_g^*, α_g^*), and a generator $g \in \mathcal{G}$, we check whether $\underline{P}_g \le p_g(\hat{\boldsymbol{\xi}}) - \alpha_g^* \sum_{r\in\mathcal{R}} \hat{\xi}_r \le \bar{P}_g$ in (1b) is violated. By running evaluation tests over 15,000 data samples, we can obtain the corresponding violation probability of this constraint (i.e., the number of times that this constraint is violated divided by 15,000). We then report the maximum violation probability across all of such two-sided nominal constraints enforced by chance constraints in each model, as shown in Table 2.

In terms of the optimal values, we can observe that there is no significant difference (within 5%) among the three models, irrespective of which dataset is applied. This is mainly because there is no direct renewable generation cost in objective function (8a), and accordingly the effect of renewable generation uncertainty appears to be small on the optimal cost.

In terms of the maximum violation probability, we can observe that as compared to the other two models, the OCC-ACOPF is subject to a relatively high violation probability. It indicates that a nominal constraint in ACOPF is likely to fail with a very high probability (e.g., 43.1%), which is much higher than the pre-set risk

Table 2
Performance on Optimality and Violation Probability with $\epsilon = 0.2$.

Datasets		OCC-ACOPF	TCC-ACOPF-K1	TCC-ACOPF-K2
Right-skewed	Opt. Val. ($)	91289.2	92291.0	93003.1
	Vio. Prob.	0.431	0.286	0.178
Normally distributed	Opt. Val. ($)	89699.4	90081.2	90105.7
	Vio. Prob.	0.314	0.184	0.182
Left-skewed	Opt. Val. ($)	89568.1	90225.1	90282.4
	Vio. Prob.	0.378	0.233	0.182

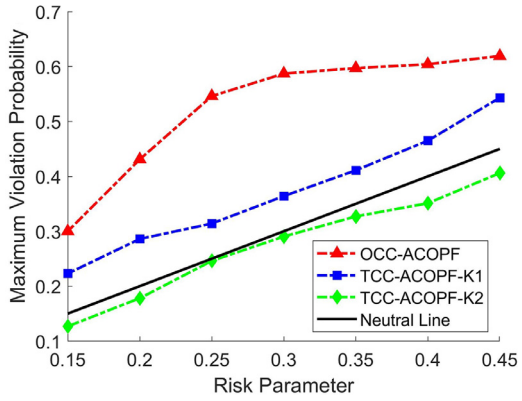

Fig. 3. Sensitivity analyses on the maximum violation probability with respect to ϵ.

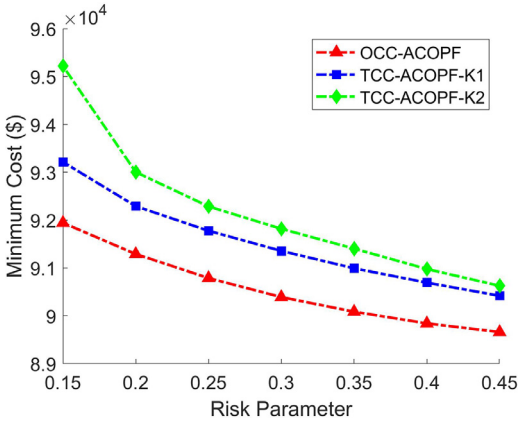

Fig. 4. Sensitivity analyses on the total cost with respect to ϵ.

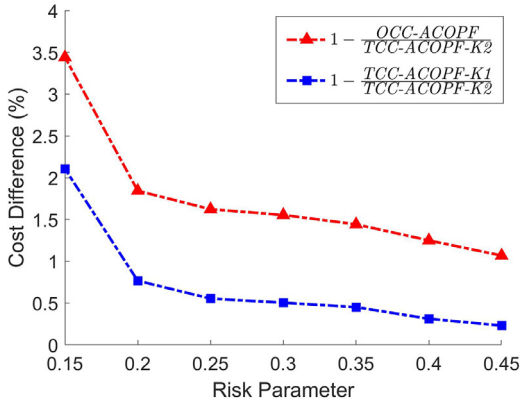

Fig. 5. Sensitivity analyses on cost difference with respect to ϵ.

control parameter $\epsilon = 0.2$. In addition, while TCC-ACOPF-K1 clearly performs better than the OCC-ACOPF by ensuring less violation probability, it does not well account for the data skewness. In fact, the violation probabilities induced by TCC-ACOPF-K1 over the Left-skewed and Right-skewed datasets can be up to 23.3% and 28.6%, respectively, both of which are higher than the pre-set risk control parameter $\epsilon = 0.2$. Finally, we can observe that TCC-ACOPF-K2, which considers a GM distribution with two components, significantly outperforms the above two models by providing the most robust solutions. The corresponding maximum violation probabilities under various datasets are all less than the pre-set risk control parameter $\epsilon = 0.2$.

Next, we use the Right-skewed dataset to perform sensitivity analyses with respect to the value of ϵ, with results illustrated in Figs. 3–5. Figure 3 shows that the maximum violation probability guaranteed by the OCC-ACOPF model, which always exceeds the pre-set risk control parameter ϵ, is also very sensitive to ϵ. It indicates that the solution provided by OCC-ACOPF is not robust and may lead to unstable operations in practice. The blue curve representing TCC-ACOPF-K1 stays lower than the red curve representing OCC-ACOPF and accordingly ensure less violation probability, but it is always above the neutral line representing the value of ϵ. It indicates that the solution provided by TCC-ACOPF-K1, albeit further incorporating the TCCs, is not robust enough. In contrast, the green curve representing TCC-ACOPF-K2 turns out to be very robust with respect to the risk parameter ϵ because the maximum violation probability always stays below the neutral line, regardless of the specific value that ϵ takes. In short, TCC-ACOPF-K1 confirms the quality of our proposed TCC approach over the OCC approach by enforcing joint probabilistic bound on two-sided constraints, and TCC-ACOPF-K2 further confirms the advantages of adopting GM distributions over simple Gaussian distributions by considering skewness in the data, as illustrated in Fig. 2.

Furthermore, Fig. 4 shows the optimal total cost of the models, while the cost difference between TCC-ACOPF-K2 and OCC-ACOPF (in square-dotted blue) and that between TCC-ACOPF-K2 and TCC-ACOPF-K1 (in triangle-dotted red) are shown in Fig. 5. We observe that as the risk parameter ϵ increases, (i) the total costs of all models decrease because each model prepares less electrical generation to hedge against the decreasing risk; (ii) both cost differences decrease as well due to the same reason. Both cost differences are relatively small (within 5%) because the renewable generation uncertainty does not significantly affect the total generation cost, as there is no need to include renewable generation cost in the objective function of each model. Nevertheless, as Remark 2 notes, more cost components can be considered in the objective function of model (1), by which the cost differences may be larger. In addition, the consideration of renewable generation uncertainty does affect the reliability of the obtained solutions, as shown in Table 2. Combining the results from Figs. 3, 4, and 5, we can observe that, in order to gain a solution with high reliability under the Right-skewed dataset, TCC-ACOPF-K2 is preferred because it accurately captures the possibility of electricity shortage, while inducing higher yet relatively similar costs.

Table 3
Computational Time (seconds).

Model	ϵ (%)				
	20	10	5	1	0.5
ACOPF	2.30	1.54	2.08	2.33	1.95
OCC-ACOPF	0.35	0.38	0.38	0.36	0.36
TCC-ACOPF-K1	2.19	2.82	2.42	2.52	2.92
TCC-ACOPF-K2	4.97	6.24	5.28	5.49	5.93

Table 4
Algorithm 1 vs. Uni-distance: # Pieces Required.

δ	0.05	0.01	0.005	0.002	0.001	0.0005
Algorithm 1	3	6	7	10	14	19
Uni-distance	4	8	11	17	23	33

4.2.2. Computational efficiency

We demonstrate the computational efficiency of our proposed models and other benchmark models using the Right-skewed dataset on the modified IEEE-118 bus system. We first report the computational times of the four models, i.e., deterministic ACOPF (1), OCC-ACOPF, TCC-ACOPF-K1, and TCC-ACOPF-K2, with respect to different values of ϵ, as shown in Table 3. We observe that all four models can be solved very efficiently (within 10 seconds). The slight computational time increase for TCC-ACOPF-K2 is due to the relatively high number of linear segments in the PWL approximation, i.e., $M = 10$.

Next, we focus on the effect of PWL approximation accuracy on the efficiency of solution obtained from TCC-ACOPF-K2. As we mentioned in Section 3.2, the PWL approximation error is controllable through the number of interpolation points and their positioning in the PWL approximation, and we proved that Algorithm 1 obtains the minimum number of interpolation points required for an approximation error δ when approximating the Gaussian CDF. To numerically demonstrate the significance of using Algorithm 1, we compare it with the typical uni-distance algorithm used for PWL approximation. Different from Algorithm 1, the uni-distance algorithm positions $t_0, \ldots, t_M$ such that $t_i - t_{i-1} = t_j - t_{j-1}$ for any $i, j \in [M]$ and $i \neq j$. We first report the number of pieces required by each algorithm to reach the approximation error δ in Table 4. We have the following observations. (i) By increasing the approximation quality, i.e., reducing δ, both algorithms rationally require more linear pieces. (ii) For given δ, Algorithm 1 requires significantly fewer linear pieces, as compared to the uni-distance algorithm. For instance, when $\delta = 0.001$, Algorithm 1 requires around 40% fewer pieces than the uni-distance algorithm. The demand of fewer pieces leads to significant reduction in computational time because fewer constraints are involved.

We then perform sensitivity analyses to investigate how the approximation error δ affects the optimal value (represented by "Opt. Val. ($)" in Table 5) and computational time (represented by "CPU (seconds)" in Table 5) of model TCC-ACOPF-K2. In Table 5, it is clear that more accurate PWL approximations (i.e., smaller values of δ) lead to higher-quality solutions for TCC-ACOPF-K2, while such increasing quality comes at the price of longer computational times. In practice, system operators can choose an appropriate value of δ based on practical considerations to reach a balance between solution quality and computational time.

Table 6
Hourly wind power forecast (MW).

Farm g	1	2	3	4	5
$r_g^{\mathrm{p}}(0)$	70	147	102	150	160

4.2.3. Scalability

We evaluate the scalability of our proposed approach under various penetration levels of renewable energy. We compare the performance of the four models mentioned above: deterministic ACOPF (1), OCC-ACOPF, TCC-ACOPF-K1, and TCC-ACOPF-K2, over the IEEE 30-bus test system. In this system, we consider $|\mathcal{R}| = 3, 4$, and 5 wind farms, respectively, leading to 20%, 30%, and 40% of renewable penetration levels. The forecast power outputs of all the five wind farms are described in Table 6.

To represent the uncertainty, we follow the process described in Section 4.1 to generate synthetic data samples with Left-skewed distributions, where the skewness parameter $\varpi = 0.9$. We generate 100 data instances, where each instance is generated following the process in Section 4.1 and has $N = 20,000$ data samples and $|\mathcal{R}|$ variates. Given one of the four models above, we follow the experiments in Section 4.2.1 to first obtain its in-sample optimal value and out-of-sample violation probability for each of the 100 data instances, and then report the average result in Table 7. In Table 7, the column "Problem Setting" describes how the wind farms are installed, the columns "Opt. Val. ($)" and "Vio. Prob." have the same meaning with those in Table 2, and the column "CPU (seconds)" represents the average computational time for the in-sample tests.

We have the following observations. First, as the renewable penetration level increases, the total generation cost of the dispatchable thermal generators decreases for all the four models because the thermal generators need to produce less electricity when more renewable energy is available. Second, all the models can be efficiently solved in less than one second, and the computational times are similar for all the models under different penetration levels. Third, when the penetration level is fixed, the optimal objective values of the four models are similar. Fourth, with a slightly higher generation cost and longer computational time, the model TCC-ACOPF-K2 can generate solutions that meet the violation probability requirement (i.e., $\leq \epsilon = 0.2$) under different penetration levels. Specifically, it is the only model that can do so. Meanwhile, we observe that the model TCC-ACOPF-K2 performs better than the model TCC-ACOPF-K1, further indicating the benefits of adopting a GM distribution based on our proposed TCC models. Overall, the results demonstrate the effectiveness of our proposed approach in dealing with the increasing renewable energy penetration.

4.2.4. Real case studies

We perform case studies on the modified IEEE 118-bus test system using existing historical data in practice to further compare the computational performance of four models: deterministic ACOPF (1), OCC-ACOPF, TCC-ACOPF-K1, and TCC-ACOPF-K2. Specifically, we evaluate how the solutions to the four models perform under different uncertainty realizations. To that end, we first take 5,000 data samples from the real-world dataset for each wind farm location, fit them to the corresponding distribution of the uncertain forecast error for each model, and obtain certain parameters

Table 5
Effect of PWL Function on the Solution Quality of TCC-ACOPF-K2.

δ	0.05	0.01	0.005	0.002	0.001	0.0005
Opt. Val. ($)	98120.3	93263.8	93018.2	93003.3	92956.8	92951.9
CPU (seconds)	3.8	4.1	4.4	4.8	8.9	14.2

Table 7
Performance on Scalability with $\epsilon = 0.2$.

Problem Setting	Model	Numerical Results		
		Opt. Val. ($)	Vio. Prob.	CPU (seconds)
3 Farms: {1, 2, 3} at buses: {2, 7, 30}	ACOPF	484.80	–	0.14
	OCC-ACOPF	484.80	0.24	0.06
	TCC-ACOPF-K1	484.82	0.21	0.16
	TCC-ACOPF-K2	484.82	0.19	0.42
4 Farms: {1, 2, 3, 4} at buses: {2, 7, 17, 30}	ACOPF	413.50	–	0.14
	OCC-ACOPF	413.58	0.22	0.05
	TCC-ACOPF-K1	413.77	0.21	0.19
	TCC-ACOPF-K2	414.15	0.19	0.52
5 Farms: {1, 2, 3, 4, 5} at buses: {2, 7, 17, 24, 30}	ACOPF	340.60	–	0.17
	OCC-ACOPF	340.65	0.25	0.04
	TCC-ACOPF-K1	343.47	0.24	0.24
	TCC-ACOPF-K2	345.45	0.18	0.46

(e.g., mean, variance, and weighted means) of the distribution for each model. Then, we solve the four models to obtain the optimal solution to each model. Given an optimal solution obtained from each model, denoted by $(\boldsymbol{p}^*, \boldsymbol{q}^*, \boldsymbol{v}^*, \boldsymbol{\theta}^*, \boldsymbol{\alpha}^*, \boldsymbol{\gamma}^*)$, we evaluate the out-of-sample performance of this solution over another 2,000 data samples. Specifically, given one of the 2,000 data samples, i.e., a given uncertainty realization $\hat{\boldsymbol{\xi}}$, we perform the out-of-sample evaluation by solving the following re-dispatch model, where the obtained solution $(\boldsymbol{p}^*, \boldsymbol{q}^*, \boldsymbol{v}^*, \boldsymbol{\theta}^*, \boldsymbol{\alpha}^*, \boldsymbol{\gamma}^*)$ is given as input parameters, to evaluate the feasibility of this obtained solution.

$$\min \sum_{(i,j)\in\mathcal{L}} \bar{s}^{\mathrm{L}}_{ij} + \sum_{g\in\mathcal{G}} (\underline{s}^{\mathrm{p}}_g + \bar{s}^{\mathrm{p}}_g + \underline{s}^{\mathrm{q}}_g + \bar{s}^{\mathrm{q}}_g) + \sum_{i\in\mathcal{B}} (\underline{s}^{\mathrm{v}}_i + \bar{s}^{\mathrm{v}}_i) \tag{23a}$$

$$\text{s.t. } \underline{P}_g - \underline{s}^{\mathrm{p}}_g \le p_g(\hat{\boldsymbol{\xi}}) \le \bar{P}_g + \bar{s}^{\mathrm{p}}_g, \qquad \forall\, g \in \mathcal{G}, \tag{23b}$$

$$\underline{Q}_g - \underline{s}^{\mathrm{q}}_g \le q_g(\hat{\boldsymbol{\xi}}) \le \bar{Q}_g + \bar{s}^{\mathrm{q}}_g, \qquad \forall\, g \in \mathcal{G}, \tag{23c}$$

$$\underline{v}_i - \underline{s}^{\mathrm{v}}_i \le v_i(\hat{\boldsymbol{\xi}}) \le \bar{v}_i + \bar{s}^{\mathrm{v}}_i, \qquad \forall\, i \in \mathcal{B}, \tag{23d}$$

$$\left(e^{\mathrm{p}}_{ij}(\hat{\boldsymbol{\xi}})\right)^2 + \left(e^{\mathrm{q}}_{ij}(\hat{\boldsymbol{\xi}})\right)^2 \le \left(\bar{I}_{ij} + \bar{s}^{\mathrm{L}}_{ij}\right)^2, \qquad \forall\, (i,j) \in \mathcal{L}, \tag{23e}$$

(1f), (1g)

$$r^{\mathrm{q}}_g(\hat{\boldsymbol{\xi}}) = \gamma^*_g r^{\mathrm{p}}_g(\hat{\boldsymbol{\xi}}), \qquad \forall g \in \mathcal{R}, \tag{23f}$$

$$p_g(\hat{\boldsymbol{\xi}}) = p^*_g - \alpha^* \sum_{g\in\mathcal{R}} \xi_g \qquad \forall\, g \in \mathcal{G}, \tag{23g}$$

$$q_g(\hat{\boldsymbol{\xi}}) = q^*_g, \qquad \forall g \in \mathcal{G}_i, i \in \mathcal{B}_{pq}, \tag{23h}$$

$$v_i(\hat{\boldsymbol{\xi}}) = v^*_i, \qquad \forall\, i \in \mathcal{B}_{pv} \cup \mathcal{B}_{v\theta}, \tag{23i}$$

$$\underline{s}^{\mathrm{p}}_g, \bar{s}^{\mathrm{p}}_g, \underline{s}^{\mathrm{q}}_g, \bar{s}^{\mathrm{q}}_g, \underline{s}^{\mathrm{v}}_i, \bar{s}^{\mathrm{v}}_i, \bar{s}^{\mathrm{L}}_{ij} \ge 0. \tag{23j}$$

The above re-dispatch model is essentially a modified nonlinear ACOPF (1), in which nonnegative slack variables $(\underline{s}^{\mathrm{p}}_g, \bar{s}^{\mathrm{p}}_g)$, $(\underline{s}^{\mathrm{q}}_g, \bar{s}^{\mathrm{q}}_g)$, $(\underline{s}^{\mathrm{v}}_i, \bar{s}^{\mathrm{v}}_i)$, and $\bar{s}^{\mathrm{L}}_{ij}$ are added to measure constraint violations and the original variables in (1) are fixed to their optimal values through the affine response control policies (see Section 2.2). Here the objective function (23a) is to minimize the total amount of violations (i.e., the summation of all slack variables). Constraints (23b), (23c), (23d), and (23e) measure the violation of two-sided bounds through slack variables. Constraints (23g), (23h), and (23i) fix the values of the explicit decision variables at the given optimal solution $(\boldsymbol{p}^*, \boldsymbol{q}^*, \boldsymbol{v}^*, \boldsymbol{\alpha}^*, \boldsymbol{\gamma}^*)$. For instance, (23g) fixes the amount of active power generation from each thermal generator according to the AGC power response policy mentioned in (5). Note that, to evaluate the solution to the deterministic ACOPF (1), which does not have the participation factor $\boldsymbol{\alpha}$ as decision variable, we manually set $\alpha^*_g = \mathcal{U}(g)/\sum_{g\in\mathcal{G}} \mathcal{U}(g)$ for each $g \in \mathcal{G}$, where $\mathcal{U}(g) = \max\{\bar{P}_g - P_g, P_g - \underline{P}_g\}$, representing unused capacity of generator $g \in \mathcal{G}$.

When the re-dispatch model leads to an optimal value at zero, i.e., no violation, then the given solution $(\boldsymbol{p}^*, \boldsymbol{q}^*, \boldsymbol{v}^*, \boldsymbol{\theta}^*, \boldsymbol{\alpha}^*, \boldsymbol{\gamma}^*)$ is feasible to the corresponding given uncertainty realization $\hat{\boldsymbol{\xi}}$ under the existing dispatchable resources. Otherwise, when some slack variables take positive values, the given solution is infeasible to the given uncertainty realization. We accordingly introduce four "imbalance metrics" to measure the violations corresponding to each of (1b)–(1f): $I_p = \sum_{g\in\mathcal{G}} (\bar{s}^p_g + \underline{s}^p_g)$, $I_q = \sum_{g\in\mathcal{G}} (\bar{s}^q_g + \underline{s}^q_g)$, $I_v = \sum_{i\in\mathcal{B}} (\bar{s}^v_i + \underline{s}^v_i)$, and $e_{ij} = \sum_{ij\in\mathcal{L}} \bar{s}^l_{ij}$. The results are reported in Table 8, where we report the average value and standard deviation (i.e., the number within a parentheses) of each imbalance metric over the 2,000 data samples. From the table, we can observe that while the chance-constrained models outperform the deterministic one and the TCC-based models outperform the OCC-based one, the TCC-ACOPF-K2 model based on our proposed approaches provides the smallest values over all of the four metrics. That is, TCC-ACOPF-K2 is exceptionally robust by providing solutions feasible to more uncertainty realizations. Such high reliability demonstrates the significance of our proposed model and approaches, which consider two-sided chance constraints for the ACOPF problem under uncertainty and include more accurate distributional information to represent the uncertainty by adopting GM distributions.

Table 8
Out-of-Sample Performance over Real Data.

Model	I_p		I_q		I_v		I_l	
ACOPF	6.9	(31.2)	93.5	(285.2)	0.6	(1.4)	58.6	(185.3)
OCC-ACOPF	2.1	(13.1)	26.7	(109.1)	0.2	(0.5)	15.5	(56.3)
TCC-ACOPF-K1	0.6	(2.6)	7.2	(33.7)	0.2	(0.4)	8.9	(37.1)
TCC-ACOPF-K2	0.2	(1.8)	6.4	(24.2)	0.1	(0.2)	5.1	(25.3)

5. Conclusion

As higher levels of renewable electricity penetrate the power system, the increasing uncertainty in the power system can cause adverse power interruptions, power outages, and network instabil-

ity. To secure the smooth operation of the power system, more accurate attention to these uncertainties is necessary. In this paper, we propose a fully two-sided chance-constrained AC optimal power flow formulation. This TCC-ACOPF guarantees (with a predefined probability) that both upper and lower bounds on active and reactive power generations, voltage, and power flows simultaneously hold under uncertainty. In our formulation and to model the effect of uncertainty, we adapt a GM distribution to represent the forecast errors. Hence, we address the forecast errors much more accurately, especially in cases where the data are skewed and the common normality assumption fails. This novel TCC-ACOPF problem, however, is nonlinear and nonconvex; hence, we propose an SOC tractable approximation for it. To do so, we first provide a convex approximation of a TCC under GM distribution, which is also exact when a sufficient condition is satisfied. The resulting convex formulation is nonlinear; hence, it is next efficiently approximated by a set of SOC constraints using PWL approximation of the CDF function. We prove that the resulting SOC formulation enjoys asymptotic convergence properties. Moreover, the resulting tractable formulation becomes more accurate if a high number of well-positioned segments construct the PWL function. On the other hand, a higher number of segments also leads to computational difficulties. Therefore, to speed up our solution procedure, we also provide an algorithm to optimally select the PWL segments. A proof of optimality of the algorithm is also included.

In case studies on modified IEEE 30-bus and IEEE 118-bus test systems, we show that our TCC-ACOPF formulations achieve higher quality optimal solutions compared to their OCC-ACOPF and deterministic OPF benchmarks. Moreover, it is shown that our formulations are significantly more robust against uncertainty, especially under high penetration levels of renewable energy, while being computationally tractable. In particular, we observe that the utilization of GM distribution with two-sided chance constraints in TCC-ACOPF-K2 leads to maximum robustness both on synthetic and real historical datasets. We also show that our proposed PWL $\delta-$approximation algorithm can successfully speed up our methodology by the efficient selection of PWL segments, hence, making our methodology suitable for large-scale real-world applications.

This research can be extended in different directions. First, the objective function of our generic ACOPF model (1) includes only the generation cost of dispatchable thermal generators. It would be interesting to consider more cost components (e.g., ancillary services costs) in the objective function, as additional practical cost considerations may be included in the industry. Second, our current experiments consider the value of K at 1 and 2. More results that demonstrate the benefits of adopting GM distributions can be obtained by increasing the value of K. Indeed, a larger K helps better represent the distribution of the uncertainty. We leave these for future research.

Acknowledgments

The authors thank the editor and three anonymous referees for their sincere and constructive suggestions, which have helped improve the quality of this paper significantly. Jianqiang Cheng was supported in part by the Office of Naval Research [Grant N00014-20-1-2154] and in part by the National Science Foundation [Grant ECCS-2143679]. Kai Pan, also affiliated with the Hong Kong Polytechnic University Shenzhen Research Institute, was supported in part by the National Natural Science Foundation of China [Grant 72001185] and in part by the Research Grants Council of Hong Kong [Grant 15501920]. Boshi Yang was supported in part by the Office of Naval Research [Grant N00014-20-1-2154].

References

Aigner, K.-M., Clarner, J.-P., Liers, F., & Martin, A. (2022). Robust approximation of chance constrained DC optimal power flow under decision-dependent uncertainty. *European Journal of Operational Research, 301*(1), 318–333.
Ardestani-Jaafari, A., & Delage, E. (2016). Robust optimization of sums of piecewise linear functions with application to inventory problems. *Operations Research, 64*(2), 474–494.
Baker, K., Dall'Anese, E., & Summers, T. (2016). Distribution-agnostic stochastic optimal power flow for distribution grids. In *2016 North American power symposium (NAPS)* (pp. 1–6). IEEE.
Benaglia, T., Chauveau, D., Hunter, D. R., & Young, D. S. (2010). mixtools: An R package for analyzing mixture models. *Journal of Statistical Software, 32*, 1–29.
Bertholon, H., Monfort, A., & Pegoraro, F. (2007). Pricing and inference with mixtures of conditionally normal processes,.
Bienstock, D., Chertkov, M., & Harnett, S. (2014). Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Review, 56*(3), 461–495.
Bloomenergy (2021). California power outage map. Accessed March 18 https://www.bloomenergy.com/bloom-energy-outage-map.
Borkowska, B. (1974). Probabilistic load flow. *IEEE Transactions on Power Apparatus and Systems*, (3), 752–759.
Cabrera-Tobar, A., Bullich-Massagué, E., Aragüés-Peñalba, M., & Gomis-Bellmunt, O. (2019). Active and reactive power control of a PV generator for grid code compliance. *Energies, 12*(20), 3872.
Capitanescu, F. (2016). Critical review of recent advances and further developments needed in AC optimal power flow. *Electric Power Systems Research, 136*, 57–68.
Carpentier, J. (1962). Contribution a l'etude du dispatching economique. *Bulletin de la Societe Francaise des Electriciens, 3*(1), 431–447.
Combettes, P. L. (2016). Perspective functions: Properties, constructions, and examples. *Set-Valued and Variational Analysis, 26*, 247–264.
Dall'Anese, E., Baker, K., & Summers, T. (2017). Chance-constrained AC optimal power flow for distribution systems with renewables. *IEEE Transactions on Power Systems, 32*(5), 3427–3438.
Dattatreya, G., & Kanal, L. N. (1990). Estimation of mixing probabilities in multiclass finite mixtures. *IEEE Transactions on Systems, Man, and Cybernetics, 20*(1), 149–158.
De Rubira, T. T., & Hug, G. (2016). Adaptive certainty-equivalent approach for optimal generator dispatch under uncertainty. In *2016 European control conference (ECC)* (pp. 1215–1222). IEEE.
Draxl, C., Hodge, B., Clifton, A., & McCaa, J. (2015). Overview and meteorological validation of the wind integration national dataset toolkit. *Technical Report*. Golden, CO (United States): National Renewable Energy Lab. (NREL).
Dunning, I., Huchette, J., & Lubin, M. (2017). Jump: A modeling language for mathematical optimization. *SIAM Review, 59*(2), 295–320.
Farivar, M., & Low, S. H. (2013). Branch flow model: Relaxations and convexification-Part I. *IEEE Transactions on Power Systems, 28*(3), 2554–2564.
Filabadi, M. D., & Azad, S. P. (2020). Robust optimisation framework for SCED problem in mixed AC-HVDC power systems with wind uncertainty. *IET Renewable Power Generation, 14*(14), 2563–2572.
Fu, W., & McCalley, J. D. (2001). Risk based optimal power flow. In *2001 IEEE porto power tech proceedings (Cat. No. 01ex502): vol. 3* (pp. 6–pp). IEEE.
Hamann, B., & Chen, J.-L. (1994). Data point selection for piecewise linear curve approximation. *Computer Aided Geometric Design, 11*(3), 289–301.
Hanasusanto, G. A., Roitch, V., Kuhn, D., & Wiesemann, W. (2017). Ambiguous joint chance constraints under mean and dispersion information. *Operations Research, 65*(3), 751–767.
Hodge, B.-M., Lew, D., Milligan, M., Holttinen, H., Sillanpaa, S., Gómez-Lázaro, E., … Giebel, G., et al., (2012). Wind power forecasting error distributions: An international comparison. *Technical Report*. Golden, CO (United States): National Renewable Energy Lab. (NREL).
Hodge, B.-M., & Milligan, M. (2011). Wind power forecasting error distributions over multiple timescales. In *2011 IEEE power and energy society general meeting* (pp. 1–8). IEEE.
Hojjat, M., & Javidi, M. H. (2015). Chance-constrained programming approach to stochastic congestion management considering system uncertainties. *IET Generation, Transmission & Distribution, 9*(12), 1421–1429.
Jaleeli, N., VanSlyck, L. S., Ewart, D. N., Fink, L. H., & Hoffmann, A. G. (1992). Understanding automatic generation control. *IEEE Transactions on Power Systems, 7*(3), 1106–1122.
Kong, L., & Maravelias, C. T. (2020). On the derivation of continuous piecewise linear approximating functions. *INFORMS Journal on Computing, 32*(3), 531–546.
Kothari, D. P., & Nagrath, I. (2003). *Modern power system analysis*. Tata McGraw-Hill Education.
Kundur, P., Paserba, J., Ajjarapu, V., Andersson, G., Bose, A., Canizares, C., … Taylor, C., et al., (2004). Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions. *IEEE Transactions on Power Systems, 19*(3), 1387–1401.
Kuryatnikova, O., Ghaddar, B., & Molzahn, D. K. (2021). Adjustable robust two-stage polynomial optimization with application to AC optimal power flow. arXiv preprint arXiv:2104.03107.
Lange, M. (2005). On the uncertainty of wind power predictions-analysis of the forecast accuracy and statistical distribution of errors. *Journal of Solar Energy Engineering, 127*(2), 177–184.
Li, B., Vrakopoulou, M., & Mathieu, J. L. (2017). Chance constrained reserve schedul-

ing using uncertain controllable loads Part II: Analytical reformulation. *IEEE Transactions on Smart Grid, 10*(2), 1618–1625.
Lin, J., Magnago, F., & Alemany, J. M. (2018). Optimization methods applied to power systems: Current practices and challenges. In *Classical and recent aspects of power system optimization* (pp. 1–18). Elsevier.
Lubin, M., Bienstock, D., & Vielma, J. P. (2015). Two-sided linear chance constraints and extensions. arXiv preprint arXiv:1507.01995.
Lubin, M., Dvorkin, Y., & Roald, L. (2019). Chance constraints for improving the security of AC optimal power flow. *IEEE Transactions on Power Systems, 34*(3), 1908–1917.
Nemirovski, A., & Shapiro, A. (2007). Convex approximations of chance constrained programs. *SIAM Journal on Optimization, 17*(4), 969–996.
NREL (2021). Electric disturbance events (OE-417) annual summaries. Accessed March 18 https://www.oe.netl.doe.gov/OE417_annual_summary.aspx
Overbye, T. J., Cheng, X., & Sun, Y. (2004). A comparison of the AC and DC power flow models for LMP calculations. In *37th Annual Hawaii international conference on system sciences, 2004. Proceedings of the* (p. 9). IEEE.
Parzen, E. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics, 33*(3), 1065–1076.
Paudyal, S., Canizares, C. A., & Bhattacharya, K. (2011). Three-phase distribution OPF in smart grids: Optimality versus computational burden. In *2011 2nd IEEE PES international conference and exhibition on innovative smart grid technologies* (pp. 1–7). IEEE.
Pena-Ordieres, A., Molzahn, D., Roald, L., & Waechter, A. (2019). DC optimal power flow with joint chance constraints. arXiv preprint arXiv:1911.12439.
Rebennack, S., & Krasko, V. (2020). Piecewise linear function fitting via mixed-integer linear programming. *INFORMS Journal on Computing, 32*(2), 507–530.
Reynolds, D. A. (2009). Gaussian mixture models. *Encyclopedia of Biometrics, 741*, 659–663.
Roald, L., & Andersson, G. (2017). Chance-constrained AC optimal power flow: Reformulations and efficient algorithms. *IEEE Transactions on Power Systems, 33*(3), 2906–2918.
Roald, L. A. (2016). *Optimization methods to manage uncertainty and risk in power systems operation*. ETH Zurich Ph.D. thesis..
Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics, 27*(3), 832–837.
Skolfield, J. K., & Escobedo, A. R. (2022). Operations research in optimal power flow: A guide to recent and emerging methodologies and applications. *European Journal of Operational Research, 300*(2), 387–404.
Stott, B., & Alsaç, O. (2012). Optimal power flow: Basic requirements for real-life problems and their solutions. In *SEPOPE XII symposium, Rio de Janeiro, Brazil: vol. 11*.
Tomek, I. (1974). Two algorithms for piecewise-linear continuous approximation of functions of one variable. *IEEE Transactions on Computers, 100*(4), 445–448.
Venzke, A., Halilbasic, L., Markovic, U., Hug, G., & Chatzivasileiadis, S. (2017). Convex relaxations of chance constrained AC optimal power flow. *IEEE Transactions on Power Systems, 33*(3), 2829–2841.
Vrakopoulou, M., Katsampani, M., Margellos, K., Lygeros, J., & Andersson, G. (2013). Probabilistic security-constrained AC optimal power flow. In *2013 IEEE grenoble conference* (pp. 1–6). IEEE.
Wächter, A., & Biegler, L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming, 106*(1), 25–57.
WirfsBrock, J. (2014). Power outages on the rise across the U.S.http://insideenergy.org/2014/08/18/power-outages-on-the-rise-across-the-u-s/
Wood, A. J., Wollenberg, B. F., & Sheblé, G. B. (2013). *Power generation, operation, and control*. John Wiley & Sons.
Xie, W., & Ahmed, S. (2017). Distributionally robust chance constrained optimal power flow with renewables: A conic reformulation. *IEEE Transactions on Power Systems, 33*(2), 1860–1867.
Zhang, H., & Li, P. (2011). Chance constrained programming for optimal power flow under uncertainty. *IEEE Transactions on Power Systems, 26*(4), 2417–2424.
Zhang, Y., Shen, S., & Mathieu, J. L. (2016). Distributionally robust chance–constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads. *IEEE Transactions on Power Systems, 32*(2), 1378–1388.
Zhuang, X., Huang, Y., Palaniappan, K., & Zhao, Y. (1996). Gaussian mixture density modeling, decomposition, and applications. *IEEE Transactions on Image Processing, 5*(9), 1293–1302.
Zimmerman, R. D., Murillo-Sánchez, C. E., & Thomas, R. J. (2010). MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems, 26*(1), 12–19.
Zohrizadeh, F., Josz, C., Jin, M., Madani, R., Lavaei, J., & Sojoudi, S. (2020). A survey on conic relaxations of optimal power flow problem. *European Journal of Operational Research, 287*(2), 391–409.