Tight Conic Approximations for Chance-Constrained AC Optimal Power Flow

Abolhassan Mohammadi Fathabad, Jianqiang Cheng
Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721, USA, amohammadi@email.arizona.edu, jqcheng@arizona.edu

Kai Pan
Department of Logistics and Maritime Studies, Faculty of Business, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, kai.pan@polyu.edu.hk

Boshi Yang
School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC, USA, boshiy@clemson.edu

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. As 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 118-bus system. We show that our formulations provide significantly more robust solutions (about 60% reduction in constraint violation) without additional computational costs, as compared to other state-of-art ACOPF formulations.

Key words: Stochastic programming, two-sided chance constraint, AC optimal power flow, second-order cone programming, piecewise linear approximation.

1. Introduction

The penetration of renewables such as wind turbines and solar panels is increasing 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 a large number of renewable energy sources, it is well known that the 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, which can lead to significant challenges from voltage fluctuations to overloaded branches and can shut down the entire or parts of the power system (De Rubira and Hug 2016, Kundur et al. 2004, Filabadi and Azad 2020). Such complexities raise the need for fast and reliable optimization methods that can both 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 stability of the system. It guarantees the secure operation of the system 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 complexity of the equations 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 et al. 2018). 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. Indeed, the deterministic DC models are still being used in many commercial and industrial applications (Stott and 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 et al. 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 and 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 and Shapiro 2007, Paudyal et al. 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 and foremost, as discussed above, the underlying physical characteristics of the power system include nonconvex and nonlinear AC power flow equations; thus, most works in literature approximate the AC power equations with an imprecise DC model to reduce the complexity (Bienstock et al. 2014, Overbye et al. 2004, Vrakopoulou et al. 2013, Xie and Ahmed 2017, Zhang et al. 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. 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 et al. (2017) distinctly models linearized AC power flow equations around a given voltage profile. Others (Hojjat and Javidi 2015, Zhang and Li 2011) have considered full linearization of the responses around expected values of the random variables using linear decision rules. 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. In this study, similar to Fu and McCalley (2001), Roald and Andersson (2017) and Lubin et al. (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 and Andersson 2017). This also further allows for the development of analytical chance-constrained reformulations.

Secondly, to facilitate the development of analytical chance-constrained reformulations, the 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, Roald and Andersson 2017, Lubin et al. 2019), 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 et al. 2016, Bienstock et al. 2014, Zhang and Li 2011), it is 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 et al. (2015), Pena-Ordieres et al. (2019), and Xie and Ahmed (2017) are the only known work treating lower and upper bounds simultaneously. Among them, the distributionally robust results in Pena-Ordieres et al. (2019), Xie and Ahmed (2017) depend on an inaccurate DC approximation, and the analytical results in Lubin et al. (2015) only considers 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, Roald and Andersson 2017, Lubin et al. 2019, 2015, Li et al. 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 error distribution have found that the forecast error distribution differs greatly from the commonly assumed normal distribution (Hodge and 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) studies historical data to analyze the underlying probability distributions of forecast errors, and it realizes that forecast error distributions are skewed in many cases. It further shows 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 et al. 2007). In fact, any distribution can be approximated by a Gaussian mixture distribution since the well-known normal kernel density estimator may be considered as a particular GM distribution (Dattatreya and Kanal 1990, Reynolds 2009, Zhuang et al. 1996). Furthermore, GM is stable by convolution and easy to simulate. Hence, employing it

helps us develop OPF solutions that are significantly robust against forecast error uncertainties, and adds a huge deal of accuracy and flexibility to our model.

The resulting two-sided chance-constrained ACOPF (TCC-ACOPF) problem under GM, 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. 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 higher computational costs. 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 and 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 and Krasko (2020) and Kong and 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 monotonocity 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 (40% reduction in number of pieces). 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 more computational costs, 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 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 Subsections 2.1 and 2.2, respectively.

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:

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

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

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

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

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

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

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

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 et al. 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 law are enforced by (1f). Specifically, the active and reactive power flows are functions of the voltage magnitudes v and voltage angels θ, as illustrated in the following equations:

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

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

where 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 power flow balance equations can be represented by the following equations:

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

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

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_i^{p} (resp. D_i^{q}) 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_g^{p} (resp. r_g^{q}) 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_{ij}^{p} and e_{ij}^{q} are uniquely determined. Thus, e^{p} and e^{q} are not involved in the function $\varphi(p,q,v,\theta) = 0$ 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 (Roald 2016, Kothari and Nagrath 2003).

The deterministic ACOPF (1) assumes that the renewable generation outputs (i.e., r_g^{p} and r_g^{q} 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 advanded decision making approaches under uncertainty for the ACOPF problem.

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, we denote this generator's uncertain active power generation amount by $r_g^{\mathrm{p}}(\xi) = r_g^{\mathrm{p}} + \xi_g$, where r_g^{p} 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 ξ denotes the vector of ξ_g's with any $g \in \mathcal{R}$. Due to such active power generation variation, the power system has to adjust its responses with respect to other parameters, i.e., $r_g^{\mathrm{q}}(\xi), p_g(\xi), q_g(\xi), v_i(\xi), \theta_i(\xi)$, 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. In this paper the response policies are intuitively selected as follows. First, the reactive

power output of a renewable generator g changes following the active power output variations according to

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

where γ_g is a decision variable that depends on operational requirements (Roald 2016, Cabrera-Tobar et al. 2019). Second, according to the automatic generation control (AGC) policy (Borkowska 1974), we set

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

where α_g is the participation factor for this generator 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 ξ, leading to the following four sets of chance constraints, respectively:

$$\mathbb{P}\left(\underline{P}_g \le p_g(\xi) \le \bar{P}_g\right) \ge 1-\epsilon_p, \qquad \forall\, g \in \mathcal{G}, \tag{6a}$$

$$\mathbb{P}\left(\underline{Q}_g \le q_g(\xi) \le \bar{Q}_g\right) \ge 1-\epsilon_q, \qquad \forall\, g \in \mathcal{G}, \tag{6b}$$

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

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

Here, $\mathbb{P}$ denotes the probability of an event with respect to the distribution of ξ. 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 $1-\epsilon_p$. 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.

Note that constraints (6d) have quadratic dependency on ξ, while the other constraints have linear dependency on ξ. Although there is no known tractable reformulation of (6d), it is known that (6d) is convex when $\epsilon_l < 1/2$ and it can be inner approximated by the following constraints (Lubin et al. 2015):

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

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

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

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 two-sided chance constraints with linear dependency on ξ and (7c) is a deterministic convex quadratic constraint.

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

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

$$\text{s.t.} \quad (6a),(6b),(6c),(7a),(7b),(7c),(1g),$$

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

where $\mathbb{E}$ denotes the expectation with respect to the distribution of ξ, objective function (8a) minimizes the expected total operating cost, and constraints (8b) describe the linearized version of power balance equations (3). Specifically, constraints (8b) are obtained from a Taylor expansion around a feasible solution (denoted by $(\hat{p},\hat{q},\hat{v},\hat{\theta})$) to the deterministic ACOPF model (1) (see e.g., Roald and Andersson (2017) and Lubin et al. (2019)), and they ensure that the solution of (8) satisfies the power balance equations under all possible realizations of uncertainty. Here, the uncertain active and reactive power injections at each bus $i \in \mathcal{B}$ are calculated as $\sum_{g\in\mathcal{G}_i} p_g(\xi) - D_i^{\mathrm{p}}(\xi) + \sum_{g\in\mathcal{R}_i} r_g^{\mathrm{p}}(\xi)$ and $\sum_{g\in\mathcal{G}_i} q_g(\xi) - D_i^{\mathrm{q}}(\xi) + \sum_{g\in\mathcal{R}_i} r_g^{\mathrm{q}}(\xi)$, respectively. Note that, (6a) - (6c) and (7a)-(7b) represent all the TCCs in our model that are nonlinear and nonconvex in general.

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:

$$\mathbb{P}\left(l_b \le h_1(x)^\top \xi + h_0(x) \le u_b\right) \ge 1-\epsilon, \tag{9}$$

where $h_1(x)$ and $h_0(x)$ are affine functions of a vector of decision variables x, l_b and u_b are lower and upper bounds of the TCC, respectively, and ξ represents a vector of random variables.

In general, TCCs are very difficult to solve. As a result, many existing studies, including Bienstock et al. (2014), Roald and Andersson (2017), and Dall'Anese et al. (2017), approximate the TCC (9) by two one-sided chance constraints as follows:

$$\mathbb{P}\left(h_1(x)^\top \xi + h_0(x) \le u_b\right) \ge 1-\hat{\epsilon}, \tag{10a}$$

$$\mathbb{P}\left(h_1(x)^\top \xi + h_0(x) \ge l_b\right) \ge 1-\hat{\epsilon}. \tag{10b}$$

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 (Nemirovski and Shapiro 2007, Hanasusanto et al. 2017).

The approximation (10) is inexact and often very weak (Xie and 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 ξ 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 true 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 ξ follows a GM distribution with K components as follows:*

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

where component $k \in [K]$ is a Gaussian distribution with mean μ_k, 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 ξ and accordingly have $F(\xi) = \sum_{k=1}^{K} w_k F_k(\xi)$, where F_k denotes the corresponding CDF of each component $k \in [K]$. Note that the covariance of each component is a multiplication of $\Sigma \succeq 0$, which is the covariance matrix base for all components. Moreover, if the distribution of ξ 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. For ease of exposition, we define $\mu'_k := h_1(x)^\top \mu_k + h_0(x)$ and $\Sigma' := \sqrt{h_1(x)^\top \Sigma h_1(x)}$ in 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:*

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

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

$$\Sigma' \le \lambda, \tag{12c}$$

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,* x *satisfies* (9) *if and only if there exists* $\lambda \in \mathbb{R}$ *such that* (x, λ) *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). Its perspective function $\hat{\Phi}(z, \lambda) := \lambda\, \Phi(z/\lambda)$ with domain $\{(z, \lambda) | z \ge 0, \lambda > 0\}$ is also concave (Boyd et al. 2004). Since affine substitution and nonnegative weighted summation maintain concavity, (12a) is a convex constraint. 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). Specifically, under Assumption 1, we have

$$\begin{aligned}
\mathbb{P}_F\left(l_b \le h_1(x)^\top \xi + h_0(x) \le u_b\right) &= \sum_{k=1}^K w_k \mathbb{P}_{F_k}\left(l_b \le h_1(x)^\top \xi + h_0(x) \le u_b\right) \\
&= \sum_{k=1}^K w_k \mathbb{P}_{F_k}\left(\frac{l_b - \mu'_k}{\sqrt{\eta_k \Sigma'}} \le \frac{h_{(x,\xi)} - \mu'_k}{\sqrt{\eta_k \Sigma'}} \le \frac{u_b - \mu'_k}{\sqrt{\eta_k \Sigma'}}\right) \\
&= \sum_{k=1}^K w_k \left(\Phi\left(\frac{u_b - \mu'_k}{\sqrt{\eta_k \Sigma'}}\right) - \Phi\left(\frac{l_b - \mu'_k}{\sqrt{\eta_k \Sigma'}}\right)\right) \\
&= \sum_{k=1}^K w_k \left(\Phi\left(\frac{u_b - \mu'_k}{\sqrt{\eta_k \Sigma'}}\right) + \Phi\left(\frac{l_b - \mu'_k}{\sqrt{\eta_k \Sigma'}}\right)\right) - 1,
\end{aligned}$$

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

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

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

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

which, together with (12a), implies (13).

Finally, we show the equivalence between (12) and (13) when $\epsilon \le (1/2)\min\{w_1, \ldots, w_K\}$. Suppose that x satisfies constraint (13). We claim that for all $k \in [K]$,

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

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

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

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

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

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

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

Remark 2. 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 $1 - \epsilon \geq 1 - (1/(2K))$. That means a mixture of 4 Gaussians with $1 - \epsilon \geq 0.875$ 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, given $t \in \mathbb{R}^{M+1}$ with components $t_0 = 0$ and $t_1, \ldots, t_M \in (0, \infty)$, we define $\hat{\Phi}^t_M(z) = \min_{m\in[M+1]}\{a_m z + b_m\}$ the PWL function with $M + 1$ segments such that

$$\begin{cases} a_m t_{m-1} + b_m = \Phi(t_{m-1}), \\ a_m t_m + b_m = \Phi(t_m), \end{cases} \qquad \forall\, m \in [M],$$

$a_{M+1} = 0$, and $b_{M+1} = \Phi(t_M)$, as illustrated in Figure 1. We further let $\Delta_m := t_m - t_{m-1}$ for each $m \in [M]$ and $E(t) := \sup_{z\geq 0}\{\Phi(z) - \hat{\Phi}^t_M(z)\}$. The following lemma about $\hat{\Phi}^t_M(\cdot)$ is straightforward due to the concavity of $\Phi(\cdot)$, and the proof is omitted.

Lemma 1. *For any $M \geq 1$ and $t \in \mathbb{R}^{M+1}$, $\hat{\Phi}^t_M(\cdot)$ is a concave PWL function and $\hat{\Phi}^t_M(z) \leq \Phi(z)$ for all $z \in [0, \infty)$. Moreover, $E(t) \to 0$ for any sequence of t such that $\Delta_m \to 0$, $m \in [M]$ and $t_M \to \infty$.*

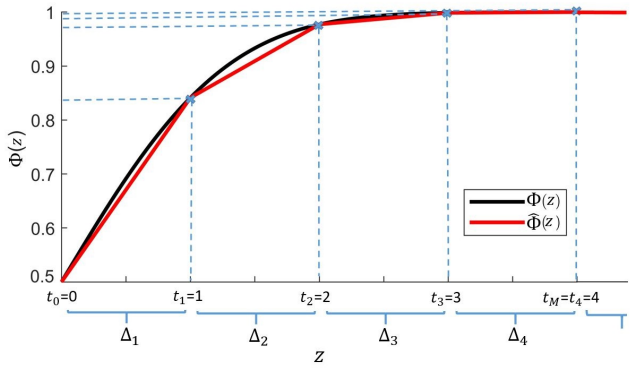

Figure 1 **An example of PWL fitting of the standard normal CDF ($M = 4$).**

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

$$\lambda\left(\sum_k w_k\left(\hat{\Phi}^t_M\left(\frac{u_b-\mu'_k}{\sqrt{\eta_k}\lambda}\right)+\hat{\Phi}^t_M\left(\frac{\mu'_k-l_b}{\sqrt{\eta_k}\lambda}\right)\right)\right)\geq\lambda(2-\epsilon). \tag{14}$$

Moreover, when $E(t)\to 0$, constraint (14) converges to (12a) in the following sense.

PROPOSITION 2. *Suppose that $f(\cdot,\cdot)$ is a continuous function of (x,λ). Let (x^*,λ^*) be an optimal solution of*

$$\inf\{f(x,\lambda)|(x,\lambda)\ \textit{satisfies}\ (12)\},$$

and let (x^*_t,λ^*_t) be an optimal solution of

$$\inf\{f(x,\lambda)|(x,\lambda)\ \textit{satisfies}\ (14),(12b),(12c)\}.$$

Then, we have $f(x^*_t,\lambda^*_t)\leq f(x^*,\lambda^*)$ for arbitrary t. Moreover, if there exists $(\bar{x},\bar{\lambda})$ such that (12) are satisfied and (12a) is strictly satisfied, then

$$\lim_{E(t)\to 0} f(x^*_t,\lambda^*_t)=f(x^*,\lambda^*).$$

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

LEMMA 2. *Let $X\subseteq\mathbb{R}^n$ be a nonempty convex set, $\Psi:\mathbb{R}^n\to\mathbb{R}$ be a concave function, and $f:\mathbb{R}^n\to\mathbb{R}$ be a continuous function. Suppose that $\{\Psi_\delta\}_{\delta>0}$ is a sequence of functions from $\mathbb{R}^n$ to $\mathbb{R}$ such that for any $y\in X$,*

$$\Psi_\delta(y)\leq\Psi(y)\quad\textit{and}\quad\lim_{\delta\to 0}\Psi_\delta(y)=\Psi(y).$$

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

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

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

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

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

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

$$\lim_{\delta \to 0} f(y_\delta^*) \leq f(y^*). \tag{15}$$

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

$$f(y_\delta^*) - f(y^*) \geq 0. \tag{16}$$

By combining (15) and (16), we complete the proof. □

Lemma 2 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 $\{\Psi_\delta\}_{\delta > 0}$) 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. Let

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

and

$$\Psi_\delta(x, \lambda) := \lambda \left(\sum_k w_k \left(\hat{\Phi}_M^t\left(\frac{u_b - \mu_k'}{\sqrt{\eta_k}\lambda} \right) + \hat{\Phi}_M^t\left(\frac{\mu_k' - l_b}{\sqrt{\eta_k}\lambda} \right) \right) + \epsilon - 2 \right),$$

where t is any set of interpolation points such that $E(t) \leq \delta$. Also let $X := \{(x, \lambda) | (x, \lambda) \text{ satisfies (12b), (12c)}\}$. By Proposition 1 and the existence of $(\bar{x}, \bar{\lambda})$, we have that X is a

nonempty convex set, $\Psi(\cdot,\cdot)$ is a concave function, and there exists $(\bar{x},\bar{\lambda}) \in X$ such that $\Psi(\bar{x},\bar{\lambda}) > 0$. Moreover, by Lemma 1, it is easy to check that

$$\Psi_\delta(x,\lambda) \le \Psi(x,\lambda) \quad \text{and} \quad \lim_{\delta\to 0} \Psi_\delta(x,\lambda) = \Psi(x,\lambda)$$

for all $(x,\lambda) \in X$. Therefore, the proof is complete by a direct application of Lemma 2. □

REMARK 3. The assumption about the existence of $(\bar{x},\bar{\lambda})$ in Proposition 2 is easy to satisfy in practice. In particular, if there exists $\bar{x}$ such that $\mathbb{P}(l_b \le h_1(\bar{x})^\top \xi + h_0(\bar{x}) \le u_b) > 1-\epsilon$, then under Assumption 1, (13) is satisfied strictly at $\bar{x}$, and thus $(\bar{x}, \sqrt{h_1(\bar{x})^\top \Sigma h_1(\bar{x})})$ 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 (x,λ) that satisfies (12b) and (12c). Then, (x,λ) satisfies (14) if and only if there exists $\pi \in \mathbb{R}^{2K}$ such that (x,λ,π) satisfies the following linear constraints:*

$$\begin{cases} a_m(u_b - \mu_k') + b_m\sqrt{\eta_k}\lambda \ge \sqrt{\eta_k}\pi_{k,1}, \\ a_m(\mu_k' - l_b) + b_m\sqrt{\eta_k}\lambda \ge \sqrt{\eta_k}\pi_{k,2}, \end{cases} \qquad \forall\, m \in [M+1],\ \forall\, k \in [K], \tag{17a}$$

$$\sum_{k=1}^{K} w_k(\pi_{k,1} + \pi_{k,2}) \ge \lambda(2-\epsilon). \tag{17b}$$

Proof. First, if (x,λ,π) satisfies (17) for some $\pi \in \mathbb{R}^{2K}$, then

$$\begin{aligned} \lambda\hat{\Phi}_M^t\left(\frac{u_b-\mu_k'}{\sqrt{\eta_k}\lambda}\right) &= \lambda \min_{m\in[M+1]} \left\{ a_m\left(\frac{u_b-\mu_k'}{\sqrt{\eta_k}\lambda}\right) + b_m \right\} \\ &= \frac{1}{\sqrt{\eta_k}} \min_{m\in[M+1]} \{a_m(u_b-\mu_k') + b_m\sqrt{\eta_k}\lambda\} \ge \pi_{k,1}, \end{aligned} \tag{18}$$

where the inequality holds due to (17a). With similar arguments, we have

$$\lambda\hat{\Phi}_M^t\left(\frac{\mu_k'-l_b}{\sqrt{\eta_k}\lambda}\right) \ge \pi_{k,2}. \tag{19}$$

Combining (17b) and (18) - (19), we see that (14) is satisfied by (x,λ).

Next, if (x,λ) satisfies (14), then we can set

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

for $k \in [K]$. It follows that (x,λ,π) satisfies (17). □

For any continuous function $f(\cdot)$, let $f^* := \inf\{f(x) | x \text{ satisfies (9)}\}$ and $f_t^* := \inf\{f(x) | x \text{ satisfies (17), (12b), (12c) with some } \lambda,\pi\}$. We summarize our results in this section in the following theorem.

THEOREM 1. *Under Assumption 1, TCC (9) can be inner approximated by linear and SOC constraints (17), (12b), and (12c). In addition, if* $\epsilon \le (1/2)\min\{w_1, \dots, w_K\}$ *and (9) is strictly satisfied by some* $\bar{x}$*, the inner approximation converges to (9) in the sense that* $f_t^* \to f^*$ *as* $E(t) \to 0$.

Using the results from Theorem 1, one can provide an SOCP approximation of the TCC-ACOPF (8) — each of the TCCs (6a)-(6c) and (7a)-(7b) can be replaced with its counterpart as in (17), (12b), and (12c).

3.2. Quality of the Approximation

In Section 3.1, we see that the quality of our SOCP approximation highly depends on the maximum error of the PWL approximation $E(t) = \max_{z\in[0,\infty)}\{\Phi(z) - \hat{\Phi}_M^t(z)\}$. The maximum error depends on two factors: the number of interpolation points M 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 (17) and thus more computational costs. In this section, we discuss how to choose the minimum number of interpolation points for a given maximum error δ. First, we provide the following lemma to help characterize the maximum error of a linear interpolation.

LEMMA 3. *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*

$$\max_{z\in[a,b]} \left\{\Phi(z) - \hat{\Phi}_{a,b}(z)\right\} = \Phi\left(\sqrt{-\ln(2\pi s^2)}\right) - s\left(\sqrt{-\ln(2\pi s^2)} - a\right) - \Phi(a), \tag{20}$$

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 and* $\bar{E}_b(\cdot)$ *is a decreasing continuous function.*

Proof. The linear interpolation of $\Phi(\cdot)$ on $[a, b]$ can be expressed as $\hat{\Phi}_{a,b}(z) = \Phi(a) + s(z-a)$. Therefore, 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 (20).

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 $z \in (a, b]$,

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

As a result,

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

That is, $E_a(\cdot)$ is an increasing function. Similarly, we can show that $\bar{E}_b(\cdot)$ is decreasing. By (20), it is clear that both $E_a(\cdot)$ and $\bar{E}_b(\cdot)$ are continuous. □

Algorithm 1: Interpolation points positioning

Data: approximation tolerance $\delta > 0$.

Result: the smallest number M and $t_0, \ldots, t_M$ such that $\Phi(z) - \hat{\Phi}^t_M(z) \leq \delta$ for all $z \geq 0$.

$m \leftarrow 0$, $t_0 \leftarrow 0$;

while $\Phi(t_m) < 1 - \delta$ **do**

  use line search to find $t_{m+1} \in [t_m, \infty)$ such that $E_{t_m}(t_{m+1}) = \delta$;

  $m \leftarrow m + 1$;

$M \leftarrow m$;

Next, based on Lemma 3, we propose Algorithm 1 to calculate the positions of interpolation points. Given a tolerance δ, Algorithm 1 finds a PWL approximation of $\Phi(\cdot)$ with the least interpolation points, with theoretical demonstration provided in the following theorem.

THEOREM 2. *For $\delta > 0$, let $t_0, \ldots, t_M$ be the output of Algorithm 1. Let $\hat{\Phi}^{t'}_{M'}(\cdot)$ be a linear interpolation of $\Phi(\cdot)$ defined by interpolation points $t'_0, \ldots, t'_{M'}$ with $t'_0 = 0$. If $\Phi(z) - \hat{\Phi}^{t'}_{M'}(z) \leq \delta$ for all $z \geq 0$, then $M' \geq M$.*

Proof. It suffices to show that $t'_m \leq t_m$ for all $m \in [M]$. We prove the statement by induction on m. Since $\Phi(z) - \hat{\Phi}^{t'}_{M'}(z) \leq \delta$ for all $z \geq 0$, we have

$$E_{t'_{m-1}}(t'_m) \leq \delta, \ \forall\, m \in [M']. \tag{21}$$

For $m = 1$, $t'_1 \leq t_1$ because $E_0(\cdot)$ is increasing and $E_0(t_1) = \delta$. Now suppose that $t'_m \leq t_m$ for some $m \in [M - 1]$. If $t'_{m+1} > t_{m+1}$, then by the monotonicity of $E_{t'_m}(\cdot)$ and $\bar{E}_{t_{m+1}}(\cdot)$, we have

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

which contradicts to (21). Therefore, $t'_m \leq t_m$ for $m \in [M]$. □

It is worthwhile 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 a modified IEEE 118-bus test system. First, in Section 4.1, we introduce this test system, 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 were performed on a PC with an Intel Core i7-7700 CPU and 16 GB RAM. We used JuMP in Julia (Dunning et al. 2017) to implement all of the models. Ipopt solver (Wächter and 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 et al. 2010), which includes 118 buses, 54 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. In our computational experiments, we set $\epsilon_p = \epsilon_q = \epsilon_v = \epsilon_l$ in (8) and use a single risk parameter ϵ to denote all of them thereafter.

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

We consider the renewable power generation forecast error 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) were analyzed to obtain historical data on forecast errors (Draxl et al. 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 $N = 20,000$ and 11 variates (for the 11 wind farms). 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 11×11 correlation matrix ρ.

- Second, we create Group 1 data samples with size of $N/2$. Each data sample follows a multivariate normal distribution $\mathcal{N}(\check{\mu}_1, 0.1\rho^\top\rho)$ with $\check{\mu}_1 \in [-0.05, 0.05]$. That is, each data sample is 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 $\mathcal{N}(\check{\mu}_2, \rho^\top\rho)$ with $\check{\mu}_2 \in [-0.15, -0.05]$. That is, each data sample is below zero.
- Fourth, we create Group 3 data samples with size of $(N/2)(1-\varpi)$. Each data sample follows a multivariate normal distribution $\mathcal{N}(\check{\mu}_3, \rho^\top\rho)$ with $\check{\mu}_3 \in [0.05, 0.15]$. That is, each data sample is above zero.

In this paper, 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 synthetic dataset created above, we can fit it with a Gaussian distribution ($K=1$) and a GM distribution with two components ($K=2$) separately, both of which will be considered in our computational experiments, thereby obtaining necessary parameters therein. It leads to "Guassian Fit" and "GM Fit" correspondingly for each dataset. For the Gaussian Fit, the mean vectors and the 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 et al. 2009). We take a renewable generator as an example to show the fitting results based on the histogram of the data visually in Figure 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.

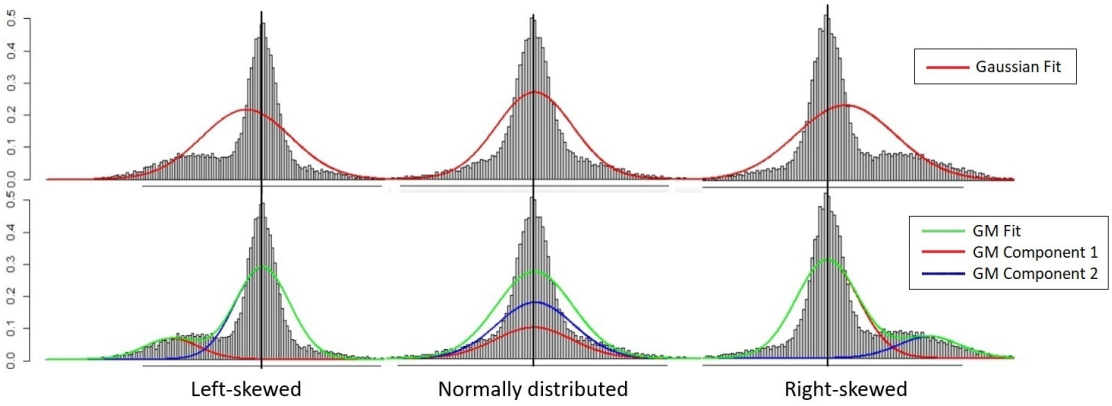

Figure 2 **Synthetic data alongside Gaussian and GM fits for forecast error of a wind farm.**

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) assumed 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; and (iii) 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 analyses in practice, as shown in Section 4.2.3.

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 with respect to their in-sample optimal values and out-of-sample violation probabilities. We set $\epsilon = 0.2$ and select 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 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 $\hat{\xi}$, an optimal solution (p_g^*, α_g^*), and a generator $g \in \mathcal{G}$, we check whether $\underline{P}_g \leq p_g(\hat{\xi}) - \alpha_g^* \sum_{r \in R} \hat{\xi}_r \leq \bar{P}_g$ 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

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

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 even higher than the pre-set risk 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 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 $\epsilon = 0.2$.

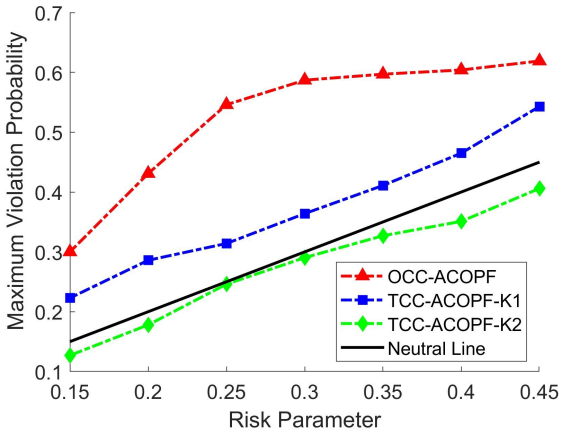

Figure 3 Sensitivity analyses on the maximum violation probability with respect to ϵ

Next, we use the Right-skewed dataset to perform sensitivity analyses with respect to the value of ϵ, with results illustrated in Figures 3 - 5. Figure 3 shows that the 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 rosbut and may lead to unstable operations in practice. The blue curve representing TCC-ACOPF-K1 stays lower than the 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 and TCC-ACOPF-K2 further confirms the advantages of adopting GM distributions over simple Gaussian distributions, as illustrated in Figure 2.

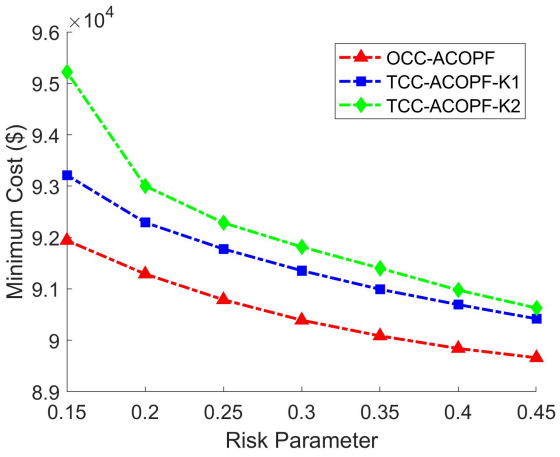

Figure 4 **Sensitivity analyses on the total cost with respect to ϵ**

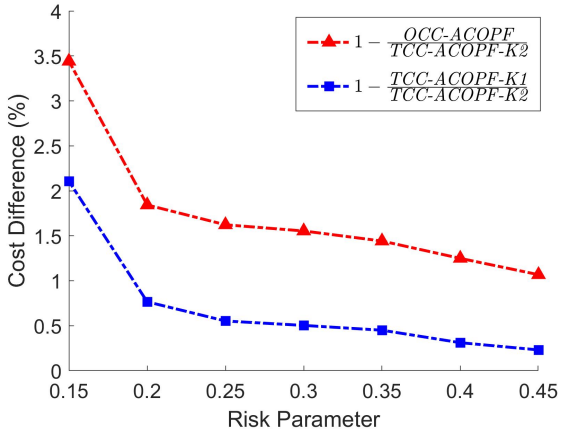

Figure 5 **Sensitivity analyses on cost difference with respect to ϵ**

Furthermore, Figure 4 shows the total cost of the models, while the cost difference between TCC-ACOPF-K2 and OCC-ACOPF (in blue) and that between TCC-ACOPF-K2 and TCC-ACOPF-K1 (in red) are shown in Figure 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, such uncertainty consideration does affect the reliability of the obtained solution, as shown in Table 2. In order to gain a solution with high reliability (e.g., by TCC-ACOPF-K2) under the Right-skewed dataset, TCC-ACOPF-K2

induces the highest cost when accurately capturing the possibility of electricity shortage, as shown in Figure 4.

4.2.2. Computational Efficiency In this section, we demonstrate the computational efficiency of our proposed models and other benchmark models using the Right-skewed dataset. 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., $M = 10$.

Table 3 Computational Time (seconds)

Model	ϵ (%)				
	20	10	5	1	0.5
ACOPF	1.91	2.89	2.45	2.44	2.18
OCC-ACOPF	1.66	2.66	3.10	2.88	4.10
TCC-ACOPF-K1	2.33	2.22	2.45	2.70	2.82
TCC-ACOPF-K2	4.83	4.55	5.85	5.02	5.31

Next, we focus on the effect of PWL approximation accuracy on the efficiency of solution 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 PWL approximation, and we proved that Algorithm 1 obtains the minimum number of interpolation points for an approximation error δ. To numerically demonstrate the significance of using Algorithm 1, we compared 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 $\delta = 0.01$, Algorithm 1 requires 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.

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

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 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

4.2.3. Real Case Studies We perform case studies over 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 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 (e.g., mean, variance, and weighted means) of the distribution for each model. Given an optimal solution obtained from each model, denoted by $(p^*, q^*, v^*, \theta^*, \alpha^*)$, we then evaluate the out-of-sample performance of this solution over another 2,000 data samples. Note that, to evaluate the solution of deterministic ACOPF (1), which does not have the participation factor α 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.

For each given uncertainty realization $\hat{\xi}$, the out-of-sample evaluation with respect to a given solution is performed by solving a re-dispatch model. The re-dispatch model is essentially a modified nonlinear ACOPF (1), in which the implicit variables are fixed to their optimal values and also slack variables $(\underline{s}_g^p, \bar{s}_g^p), (\underline{s}_g^q, \bar{s}_g^q), (\underline{s}_i^v, \bar{s}_i^v)$, and $\bar{s}_{ij}^l$ are added to measure constraint violations. The variables are fixed through the response policies. The detailed mathematical formulation for such re-dispatch is provided as follows:

$$\min \sum_{(i,j)\in\mathcal{L}} \bar{s}_{ij}^l + \sum_{g\in\mathcal{G}}(\underline{s}_g^p + \bar{s}_g^p + \underline{s}_g^q + \bar{s}_g^q) + \sum_{i\in\mathcal{B}}(\underline{s}_i^v + \bar{s}_i^v) \tag{22a}$$

$$\text{s.t. } \underline{P}_g - \underline{s}_g^p \le p_g(\hat{\xi}) \le \bar{P}_g + \bar{s}_g^p, \qquad \forall\, g \in \mathcal{G}, \tag{22b}$$

$$\underline{Q}_g - \underline{s}_g^q \le q_g(\hat{\xi}) \le \bar{Q}_g + \bar{s}_g^q, \qquad \forall\, g \in \mathcal{G}, \tag{22c}$$

$$\underline{v}_i - \underline{s}_i^v \le v_i(\hat{\xi}) \le \bar{v}_i + \bar{s}_i^v, \qquad \forall\, i \in \mathcal{B}, \tag{22d}$$

$$\left(e^{\mathrm{p}}_{ij}(\hat{\xi})\right)^2 + \left(e^{\mathrm{q}}_{ij}(\hat{\xi})\right)^2 \leq \left(\bar{I}_{ij} + \bar{s}^l_{ij}\right)^2, \qquad \forall\,(i,j) \in \mathcal{L}, \tag{22e}$$

(1f), (1g)

$$p_g(\hat{\xi}) = p^*_g - \alpha^* \sum_{g \in \mathcal{R}} \xi_g \qquad \forall\, g \in \mathcal{G}, \tag{22f}$$

$$q_g(\hat{\xi}) = q^*_g, \qquad \forall g \in \mathcal{G}_i, i \in \mathcal{B}_{pq}, \tag{22g}$$

$$v_i(\hat{\xi}) = v^*_i, \qquad \forall\, i \in \mathcal{B}_{pv} \cup \mathcal{B}_{v\theta}. \tag{22h}$$

The objective function (22a) is to minimize the total amount of violations (i.e., the summation of all slack variables). Constraints (22b), (22c), (22d), and (22e) measure the violation of two-sided bounds through slack variables. Constraints (22f), (22g) and (22h) fix the values of the explicit decision variables at the given optimal solution $(p^*, q^*, v^*, \theta^*, \alpha^*)$. For instance, (22f) fixes the amount of active power generation from each generator according to the AGC power response policy mentioned in (5).

Table 6 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)

If the re-dispatch model leads to an optimal value at zero, i.e., no violation, then the given solution is feasible to the corresponding given uncertainty realization under the existing dispatchable resources. Otherwise, i.e., some slack variables take positive values, then 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 resulting results are reported in Table 6, 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 outperforms 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. 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 distributional information to represent the uncertainty by adopting GM distributions.

5. Conclusion

As higher levels of renewable electricity penetrate the power system, the uncertainty in the power system can cause adverse power interruptions, power outages, and network instability. To secure the smooth operation of the power system, more accurate attention to these uncertainties is necessary. In this paper, we proposed 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 adapted 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 proposed 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 proved 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 provided an algorithm to optimally select the PWL segments. A proof of optimality of the algorithm is also included.

In a case study on a modieifed IEEE 118-bus test system, we showed that our TCC-ACOPF formulations achieve higher quality optimal solutions as compared to their OCC-ACOPF and deterministic OPF benchmarks. Moreover, it is shown that our formulations are significantly more robust against uncertainty while being computationally tractable. In particular, we observed 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 showed that our proposed PWL $\delta-$approximation algorithm can successfully speed up our methodology by the efficient selection of PWL segments, hence, making our methodology appropriate for large-scale real-world applications.

References

Baker, K., Dall'Anese, E., and Summers, T. (2016). Distribution-agnostic stochastic optimal power flow for distribution grids. In *2016 North American Power Symposium (NAPS)*, pages 1–6. IEEE.

Benaglia, T., Chauveau, D., Hunter, D., and Young, D. (2009). mixtools: An R package for analyzing finite mixture models.

Bertholon, H., Monfort, A., and Pegoraro, F. (2007). Pricing and inference with mixtures of conditionally normal processes.

Bienstock, D., Chertkov, M., and 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.

Boyd, S., Boyd, S. P., and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge university press.

Cabrera-Tobar, A., Bullich-Massagué, E., Aragüés-Peñalba, M., and 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.

Dall'Anese, E., Baker, K., and 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. and 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. and Hug, G. (2016). Adaptive certainty-equivalent approach for optimal generator dispatch under uncertainty. In *2016 European Control Conference (ECC)*, pages 1215–1222. IEEE.

Draxl, C., Hodge, B., Clifton, A., and McCaa, J. (2015). Overview and meteorological validation of the wind integration national dataset toolkit. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Dunning, I., Huchette, J., and Lubin, M. (2017). Jump: A modeling language for mathematical optimization. *SIAM Review*, 59(2):295–320.

Filabadi, M. D. and 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. and McCalley, J. D. (2001). Risk based optimal power flow. In *2001 IEEE Porto Power Tech Proceedings (Cat. No. 01EX502)*, volume 3, pages 6–pp. IEEE.

Hamann, B. and 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., and 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., Scharff, R., Soder, L., Larsén, X. G., Giebel, G., et al. (2012). Wind power forecasting error distributions: An international comparison. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Hodge, B.-M. and Milligan, M. (2011). Wind power forecasting error distributions over multiple timescales. In *2011 IEEE power and energy society general meeting*, pages 1–8. IEEE.

Hojjat, M. and Javidi, M. H. (2015). Chance-constrained programming approach to stochastic congestion management considering system uncertainties. *IET Generation, Transmission & Distribution*, 9(12):1421–1429.

Kong, L. and Maravelias, C. T. (2020). On the derivation of continuous piecewise linear approximating functions. *INFORMS Journal on Computing*.

Kothari, D. P. and Nagrath, I. (2003). *Modern power system analysis*. Tata McGraw-Hill Education.

Kundur, P., Paserba, J., Ajjarapu, V., Andersson, G., Bose, A., Canizares, C., Hatziargyriou, N., Hill, D., Stankovic, A., 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.

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., and Mathieu, J. L. (2017). Chance constrained reserve scheduling using uncertain controllable loads part ii: Analytical reformulation. *IEEE Transactions on Smart Grid*, 10(2):1618–1625.

Lin, J., Magnago, F., and Alemany, J. M. (2018). Optimization methods applied to power systems: current practices and challenges. In *Classical and Recent Aspects of Power System Optimization*, pages 1–18. Elsevier.

Lubin, M., Bienstock, D., and Vielma, J. P. (2015). Two-sided linear chance constraints and extensions. *arXiv preprint arXiv:1507.01995*.

Lubin, M., Dvorkin, Y., and 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. and 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., and 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*, pages 9–pp. IEEE.

Paudyal, S., Canizares, C. A., and 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*, pages 1–7. IEEE.

Pena-Ordieres, A., Molzahn, D., Roald, L., and Waechter, A. (2019). DC optimal power flow with joint chance constraints. *arXiv preprint arXiv:1911.12439*.

Rebennack, S. and 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. and 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*. PhD thesis, ETH Zurich.

Stott, B. and Alsaç, O. (2012). Optimal power flow: Basic requirements for real-life problems and their solutions. In *SEPOPE XII Symposium, Rio de Janeiro, Brazil*, volume 11.

Tomek, I. (1974). Two algorithms for piecewise-linear continuous approximation of functions of one variable. *IEEE Transactions on Computers*, 100(4):445–448.

Vrakopoulou, M., Katsampani, M., Margellos, K., Lygeros, J., and Andersson, G. (2013). Probabilistic security-constrained AC optimal power flow. In *2013 IEEE Grenoble Conference*, pages 1–6. IEEE.

Wächter, A. and 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., and Sheblé, G. B. (2013). *Power Generation, Operation, and Control*. John Wiley & Sons.

Xie, W. and 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. and 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., and 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., and 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., and 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.