2016 IEEE 55th Conference on Decision and Control (CDC)
ARIA Resort & Casino
December 12-14, 2016, Las Vegas, USA

Distributionally Robust Risk-Constrained Optimal Power Flow Using Moment and Unimodality Information

Bowen Li, Ruiwei Jiang, and Johanna L. Mathieu

Abstract— **As we incorporate more random renewable energy into the power grid, power system operators need to ensure physical constraints, such as transmission line limits, are not violated despite uncertainty. Risk-constrained optimal power flow (RCOPF) based on the Conditional Value-at-Risk (CVaR) is a convenient modeling tool, ensuring that these constraints are satisfied with a high probability (e.g., 95%). However, in practice, it is often difficult to perfectly estimate the joint probability distribution of all uncertain variables, including renewable energy production and load consumption. In this paper, we propose a distributionally robust RCOPF approach by considering all possible probability distributions that share the same moment (e.g., mean and covariance) and unimodality properties. Moment and unimodality information can be estimated based on historical data, and so the proposed approach can be applied in a data-driven manner. In view of the computational challenges, we derive a conservative and a relaxed approximation of the problem. We reformulate these approximations as semidefinite programs (SDPs) facilitating the use of highly efficient off-the-shelf optimization solvers (e.g., CVX). We demonstrate the proposed approach based on a modified IEEE 9-bus power network.**

I. INTRODUCTION

Chance-constrained optimization is widely used to solve power system optimal power flow problems with uncertain renewables and loads, e.g., [1]–[8]. This approach enables physical constraints like line limits to be formulated as probabilistic constraints that must be satisfied with high probability, which avoids the over-conservativeness that results from robust formulations. To solve chance constrained optimal power flow problems, researchers have applied the scenario approach proposed in [9], e.g., [4], probabilistically robust methods proposed in [10], e.g., [4], [7], and analytical reformulations, e.g., [5], [6], [8]. The first approach requires large numbers of scenarios and is generally conservative, producing high cost solutions. The second requires a smaller number of scenarios but is more conservative than the first. The third assumes uncertainty distributions are known and, if actual uncertainty distributions do not follow assumed distributions, often results in low solution reliability.

Recent research proposed a distributionally robust chance-constrained optimization formulation [11], which ensures that chance constraints are satisfied at chosen probability levels for any uncertainty distribution with known moments. In [12], tractable semidefinite programming (SDP) approximations for distributionally robust individual and joint chance constraints are proposed using only first and second order moment information. Ref. [13] uses the SDP formulation proposed in [14] to solve the optimal power flow problem, showing that distributionally robust methods provide a good trade-off between performance and computational tractability. In practice, we often know more about the uncertainty distributions than their first two moments, e.g., we may be able to reasonably assume that the distribution is unimodal, and including this information would result in a less conservative solution. Recognizing this, in [15] and [16], analytical distributionally robust transformations are proposed considering univariate structural properties such as unimodality. In [17] and [18], the authors consider multivariate unimodality and derive exact and tractable SDP formulations to evaluate the worst-case probability bound and Conditional Value-at-Risk (CVaR), assuming that the design variables are given and fixed.

Although chance-constrained approaches guarantee a low probability of constraint violation, they do not capture the magnitude of violations. In this paper, we employ CVaR [19]–[21] to control both the probability and magnitude of constraint violations. That is, we study risk-constrained optimal power flow (RCOPF) based on CVaR. We develop a method to incorporate multivariate unimodality into the distributionally robust risk-constrained optimization formulation with adjustable design variables, and we apply our results to the RCOPF problem. To represent multivariate unimodality, we use α-unimodality [22], where the value of α determines the structure and shape of the joint distribution. This concept has been used in worst-case probability and expectation problems [17], [18]. Using our results, we develop two tractable approximate reformulations resulting in SDPs. The computational effort required to solve these problems is similar to that of the original problem without unimodality. We study the impact of the mode location and the value of α on the objective function using a modified IEEE 9-bus system with multiple wind power plants.

The remainder of the paper is organized as follows. The RCOPF formulation is introduced in Section II. Section III defines CVaR [19]–[21], its relationship to distributionally robust optimization, and distributionally robust CVaR (DR-CVaR) constraints [12]. Approximations using the tightest piecewise linear bounding functions on the worst-case expectation problem are provided in Section IV. In Section V, we give two tractable SDP reformulations based on the approximations. In Section VI, we test all the analytical results on the RCOPF problem and briefly introduce an extension to a multi-cut approximation with SDP reformulations. Finally, Section VII summarizes the paper and gives potential future research directions.

This work was supported by the U.S. National Science Foundation (Awards #CCF-1442495 and #CMMI-1555983).

B. Li and J.L. Mathieu are with the Department of Electrical Engineering and Computer Science, University of Michigan at Ann Arbor, libowen@umich.edu; jlmath@umich.edu

R. Jiang is with the Department of Industrial and Operations Engineering, University of Michigan at Ann Arbor, ruiwei@umich.edu

II. Risk-Constrained Optimal Power Flow

In this section, we formulate a simple single-period RCOPF problem, based on [4], [6], assuming two wind power plants.

$$\min_{P_G,d,R^{up/dn}} P_G^T[C_1]P_G + C_2^T P_G + C_3^T(R^{up}+R^{dn}) \quad (1)$$

$$\text{s.t.} - P_l \le A_s P_{inj} \le P_l \quad (2)$$

$$P_{inj} = C_G(P_G+R) + C_W P_W - C_L P_L \quad (3)$$

$$R = -d(w_1+w_2) \quad (4)$$

$$0 \le P_G + R \le \overline{P_G} \quad (5)$$

$$-R^{dn} \le R \le R^{up} \quad (6)$$

$$\mathbf{1}_{1\times N_G} d = 1 \quad (7)$$

$$\mathbf{1}_{1\times N_B}(C_G P_G + C_W P_W^f - C_L P_L) = 0 \quad (8)$$

$$P_G \in \mathbb{R}^+_{N_G}, d \in \mathbb{R}^+_{N_G},\ R^{up,dn} \in \mathbb{R}^+_{N_G}, \quad (9)$$

where $[C_1], C_2$ are cost parameters for generation P_G and C_3 is the unit cost for up/down reserve capacity $R^{up,dn}$. Constraint (2) bounds the line flow, where P_l is the line limit and A_s is a constant matrix that transforms the power injection vector P_{inj}, defined in (3), into the power flow vector. In (3), C_G, C_W, C_L are constant matrices that contain the bus indices for the generators, wind power plants, and loads, respectively; P_L is the load vector (which is assumed known and certain); P_W is the real wind generation vector; and R is the reserve usage, which is defined in (4). In (4), w_1, w_2 are wind forecast errors corresponding to the two wind power plants and d is the distribution vector that specifies the portions of the total supply/demand mismatch that will be provided by each generator. Constraint (5) limits the generator outputs to within $[0, \overline{P_G}]$ and (6) limits the reserve capacity. Constraint (7) is the power balance requirement for reserves where N_G is the number of generators. Equation (8) is the power balance requirement for generation/loads where N_B is the number of buses and P_W^f is the wind forecast vector. Equations (2), (5), and (6) contain wind uncertainties and so we can employ risk constraints to control the constraint violations. We consider a class of risk constraints in this paper and introduce them in the next section.

III. Distributionally Robust Risk Constraints

A. Conditional Value-at-Risk (CVaR)

In this section, we briefly introduce the concept of CVaR [19]–[21]. Consider a random function $L(x,\xi) : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$, where x represents an n-dimensional design variable and ξ represents a k-dimensional random vector defined on a probability space $(\mathbb{R}^k, \mathcal{B}, \mathbb{P})$. In many applications, one wishes to control the random function $L(x,\xi)$ by formulating a constraint in the form

$$L(x,\xi) \le 0. \quad (10)$$

Unfortunately, (10) is subject to uncertainty and so difficult to evaluate. One possible solution is to replace function $L(x,\xi)$ with a proxy that can appropriately evaluates its risk. CVaR is a popular tool used in, e.g., engineering and portfolio management, to measure the risk level of randomness. Intuitively, CVaR evaluates the conditional expectation of $L(x,\xi)$ on a tail part of its distribution.

Definition 3.1: CVaR. Given a random function $L(x,\xi)$ and a constant $\epsilon \in (0,1)$, the CVaR of $L(x,\xi)$ with violation probability ϵ, denoted $\mathrm{CVaR}_\epsilon(L(x,\xi))$, is defined as

$$\mathrm{CVaR}_\epsilon(L(x,\xi)) = \inf_{\beta\in\mathbb{R}} \left\{ \beta + \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}}\left[L(x,\xi)-\beta\right]^+ \right\}, \quad (11)$$

where $\mathbb{E}_{\mathbb{P}}[\cdot]^+ = \mathbb{E}_{\mathbb{P}}[\max(\cdot,0)]$.

In the above definition, the minimizer β^* in (11) is called the Value-at-Risk of $L(x,\xi)$. It can be shown that $\mathrm{CVaR}_\epsilon(L(x,\xi))$ is the conditional expectation of $L(x,\xi)$ on interval $L(x,\xi) \in [\beta^*, \infty)$. An attractive feature of CVaR is that it guarantees the satisfaction of constraint (10) with a high probability, i.e.,

$$\mathrm{CVaR}_\epsilon(L(x,\xi)) \le 0 \ \Rightarrow\ \mathbb{P}\{L(\xi) \le 0\} \ge 1-\epsilon.$$

Hence, the risk constraint $\mathrm{CVaR}_\epsilon(L(x,\xi)) \le 0$ is an appropriate proxy for the random constraint (10).

B. Distributionally Robust CVaR Constraints

In practice, it is often difficult to accurately estimate the probability distribution $\mathbb{P}$ of ξ. Usually, we only have some domain knowledge and historical data of ξ. In this case, it can be biased to estimate $\mathrm{CVaR}_\epsilon(L(x,\xi))$ by assuming that $\mathbb{P}$ belongs to any parametric family of distributions (e.g., normal or log-normal).

In this paper, we propose a distributionally robust approach that does not rely on such parametric assumptions. Instead, we assume that the moment and unimodality information of ξ is available. More precisely, we denote $\mu := \mathbb{E}_{\mathbb{P}}[\xi]$ and $S := \mathbb{E}_{\mathbb{P}}[\xi\xi^T]$, and define the α-unimodality as follows.

Definition 3.2: α-Unimodality [22]. For any fixed $\alpha > 0$, a probability distribution $\mathbb{P}$ on $\mathbb{R}^k$ is called α-unimodal with mode 0 if $t^\alpha \mathbb{P}(B/t)$ is non-decreasing in $t > 0$ for every Borel set $B \in \mathcal{B}$.

We let $\mathcal{P}^\alpha$ represent the set of all α-unimodular probability distributions. Note that the above definition is a generalization of classical unimodality concept on the real line to $\mathbb{R}^k$. Indeed, when $\alpha = k = 1$, it can be shown that α-unimodality coincides with the classical unimodality. Also, ξ can be assumed to have mode 0 without loss of generality, because we can always replace ξ with $\xi - m$ if ξ has a mode $m \ne 0$. In this paper, we consider a set $\mathcal{P}$ of probability distributions that share the moment and unimodality properties, i.e.,

$$\mathcal{P} := \{\mathbb{P} \in \mathcal{P}^\alpha : \mathbb{E}_{\mathbb{P}}[\xi] = \mu, \mathbb{E}_{\mathbb{P}}[\xi\xi^T] = S\}. \quad (12)$$

In practice, μ and S can be estimated based on the historical data of ξ. Meanwhile, α-unimodality can be estimated based on the domain knowledge of ξ. For example, we can set $\alpha = 1$ if each component of vector ξ has a unimodal distribution or $\alpha = k$ for a multivariate unimodality. Based on $\mathcal{P}$, we consider a distributionally robust CVaR (DR-CVaR) constraint

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathrm{CVaR}_\epsilon(L(x,\xi)) \le 0, \quad (13)$$

where risk constraint $\mathrm{CVaR}_\epsilon(L(x,\xi)) \le 0$ is satisfied with regard to all possible probability distributions $\mathbb{P} \in \mathcal{P}$. In this sense, (13) is distributionally robust. Furthermore, note that (13) implies that $\mathbb{P}\{L(\xi) \le 0\} \ge 1-\epsilon$ for all $\mathbb{P} \in \mathcal{P}$.

IV. APPROXIMATIONS

In this section, we describe the solution methodology for handling the DR-CVaR constraint (13). First, we specify the random function $L(x, \xi)$ as a bi-linear function of the design variable x and uncertainty ξ, i.e., $L(x, \xi) := y_0(x) + y(x)^T\xi$, where $y_0(x) : \mathbb{R}^n \to \mathbb{R}$ and $y(x) : \mathbb{R}^n \to \mathbb{R}^k$ represent two linear functions of x. Practical engineering constraints can be reformulated to match the form of this function. Then, by the definition of CVaR (11), we have

$$\begin{aligned}\sup_{\mathbb{P}\in\mathcal{P}} \mathrm{CVaR}_\epsilon(L(x,\xi)) &= \sup_{\mathbb{P}\in\mathcal{P}} \inf_{\beta\in\mathbb{R}} \left\{\beta + \frac{1}{\epsilon}\mathbb{E}_\mathbb{P}\left[L(x,\xi)-\beta\right]^+\right\} \\ &= \inf_{\beta\in\mathbb{R}} \left\{\beta + \frac{1}{\epsilon}\sup_{\mathbb{P}\in\mathcal{P}}\mathbb{E}_\mathbb{P}\left[L(x,\xi)-\beta\right]^+\right\},\end{aligned} \tag{14}$$

where the equality in (14) follows from a stochastic saddle point theorem (see, e.g., [23]). Hence, the computation of the DR-CVaR boils down to evaluating the worst-case expectation $\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_\mathbb{P}\left[L(x,\xi)-\beta\right]^+$. We state the following result for this evaluation.

Lemma 4.1: (Adapted from [18]) Let $\widetilde{\mu} := (\frac{\alpha+1}{\alpha})\mu$ and $\widetilde{S} := (\frac{\alpha+2}{\alpha})S$. For $L(x, \xi) = y_0(x) + y(x)^T\xi$, we define $a := y(x)^T\xi$ and $b := \beta - y_0(x)$. Then,

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_\mathbb{P}\left[L(x,\xi)-\beta\right]^+ = \sup_{\mathbb{P}\in\mathcal{P}(\widetilde{\mu},\widetilde{S})} \mathbb{E}_\mathbb{P}\left[\widetilde{f}(\xi)\right], \tag{15}$$

where $\mathcal{P}(\widetilde{\mu},\widetilde{S}) := \{\mathbb{P} : \mathbb{E}_\mathbb{P}[\xi] = \widetilde{\mu}, \mathbb{E}_\mathbb{P}[\xi\xi^T] = \widetilde{S}\}$ and

$$\widetilde{f}(\xi) := \begin{cases} -\left(\frac{b}{\alpha+1}\right)\left(\frac{b}{a}\right)^\alpha & \text{if } a \le b \\ \left(\frac{\alpha}{\alpha+1}\right)a - b, & \text{otherwise}\end{cases} \tag{16}$$

if $b \le 0$ or, if $b > 0$,

$$\widetilde{f}(\xi) := \begin{cases} 0 & \text{if } a \le b \\ \left(\frac{\alpha}{\alpha+1}\right)a - b + \left(\frac{b}{\alpha+1}\right)\left(\frac{b}{a}\right)^\alpha, & \text{otherwise}\end{cases}. \tag{17}$$

Lemma 4.1 indicates that computation of the DR-CVaR can be difficult because we still need to evaluate the worst-case expectation of a nonlinear function $\widetilde{f}(\xi)$. This renders the worst-case CVaR constraint (13) computationally challenging. We hence resort to replacing $\widetilde{f}(\xi)$ with more concise forms that facilitate tractable reformulations. In this paper, we propose two approximations of $\widetilde{f}(\xi)$ from above (termed $\widetilde{f}_U(\xi)$) and below (termed $\widetilde{f}_L(\xi)$). Both $\widetilde{f}_U(\xi)$ and $\widetilde{f}_L(\xi)$ are convex and have two linear pieces (see Fig. 1). In fact, they are the tightest convex and 2-piece approximations of $\widetilde{f}(\xi)$ one can obtain. Accordingly, we obtain an upper bound of $\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_\mathbb{P}\left[L(x,\xi)-\beta\right]^+$ based on $\widetilde{f}_U(\xi)$, and a lower bound based on $\widetilde{f}_L(\xi)$. We formalize these conclusions in the following two theorems.

Theorem 4.1: Let $\widetilde{f}_U(\xi) = (\frac{\alpha}{\alpha+1})[y(x)^T\xi + y_0(x) - \beta]^+ + (\frac{1}{\alpha+1})[y_0(x) - \beta]^+$. Then,

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_\mathbb{P}\left[L(\xi)-\beta\right]^+ \le \sup_{\mathbb{P}\in\mathcal{P}(\widetilde{\mu},\widetilde{S})} \mathbb{E}_\mathbb{P}\left[\widetilde{f}_U(\xi)\right]. \tag{18}$$

Proof: As in Lemma 4.1, we define $a = y(x)^T\xi$ and $b = \beta - y_0(x)$. Hence, $\widetilde{f}_U(\xi) = (\frac{\alpha}{\alpha+1})(a-b)^+ + (\frac{1}{\alpha+1})(-b)^+$. We show $\widetilde{f}(\xi) \le \widetilde{f}_U(\xi)$ by discussing the following four cases:

1) If $a \le b \le 0$, then $0 \le (b/a) \le 1$ and $-b \ge 0$. It follows that $\widetilde{f}(\xi) = -(b/(\alpha+1))(b/a)^\alpha \le (\frac{1}{\alpha+1})(-b) \le \widetilde{f}_U(\xi)$.
2) If $b \le 0$ and $a > b$, then $(a-b)^+ = a - b$ and $(-b)^+ = -b$. It follows that $\widetilde{f}_U(\xi) = (\frac{\alpha}{\alpha+1})(a-b)^+ + (\frac{1}{\alpha+1})(-b)^+ = (\frac{\alpha}{\alpha+1})a - b = \widetilde{f}(\xi)$.
3) If $b > 0$ and $a \le b$, then it is clear that $\widetilde{f}(\xi) = 0 \le \widetilde{f}_U(\xi)$.
4) If $a > b > 0$, then $0 < (b/a) < 1$ and $(a-b)^+ = a - b$. It follows that $\widetilde{f}(\xi) = (\alpha/(\alpha+1))a - b + (b/(\alpha+1))(b/a)^\alpha \le (\alpha/(\alpha+1))a - b + b/(\alpha+1) = \widetilde{f}_U(\xi)$.

As $\widetilde{f}(\xi) \le \widetilde{f}_U(\xi)$ for all $\xi \in \mathbb{R}^k$, the inequality (18) holds based on Lemma 4.1. ■

Theorem 4.2: Let $\widetilde{f}_L(\xi) = [(\frac{\alpha}{\alpha+1})y(x)^T\xi + y_0(x) - \beta]^+$. Then,

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_\mathbb{P}\left[L(\xi)-\beta\right]^+ \ge \sup_{\mathbb{P}\in\mathcal{P}(\widetilde{\mu},\widetilde{S})} \mathbb{E}_\mathbb{P}\left[\widetilde{f}_L(\xi)\right]. \tag{19}$$

Proof: As in Lemma 4.1, we define $a = y(x)^T\xi$ and $b = \beta - y_0(x)$. Hence, $\widetilde{f}_L(\xi) = [(\frac{\alpha}{\alpha+1})a - b]^+$. We show $\widetilde{f}(\xi) \ge \widetilde{f}_L(\xi)$ by discussing the following four cases:

1) If $a \le b \le 0$, then $0 \le (b/a) \le 1$ and $(-b) \ge 0$. It follows that $\widetilde{f}(\xi) = -(b/(\alpha+1))(b/a)^\alpha \ge 0$. Meanwhile, as $H(a) := -(b/(\alpha+1))(b/a)^\alpha$ is a convex function of a on interval $(-\infty, b]$, we have $H(a) \ge H'(b)(a-b) + H(b)$, i.e.,
$$\begin{aligned}-\left(\frac{b}{\alpha+1}\right)\left(\frac{b}{a}\right)^\alpha \ge \left(\frac{\alpha}{\alpha+1}\right)(a-b) - \frac{b}{\alpha+1} \\ = \left(\frac{\alpha}{\alpha+1}\right)a - b,\end{aligned}$$
where the inequality is because $H'(a) = (\alpha/(\alpha+1))(b/a)^{\alpha+1}$ and $H(b) = (-b/(\alpha+1))$. Hence, $-(b/(\alpha+1))(b/a)^\alpha \ge [(\frac{\alpha}{\alpha+1})a - b]^+$, i.e., $\widetilde{f}(\xi) \ge \widetilde{f}_L(\xi)$.
2) If $b \le 0$ and $a > b$, then $(\frac{\alpha}{\alpha+1})a - b \ge 0$. It follows that $\widetilde{f}_L(\xi) = (\frac{\alpha}{\alpha+1})a - b = \widetilde{f}(\xi)$.
3) If $b > 0$ and $a \le b$, then $(\frac{\alpha}{\alpha+1})a - b < 0$. It follows that $\widetilde{f}_L(\xi) = 0 = \widetilde{f}(\xi)$.
4) If $a > b > 0$, then $(b/a) > 0$. It follows that $\widetilde{f}(\xi) = (\alpha/(\alpha+1))a - b + (b/(\alpha+1))(b/a)^\alpha \ge (\frac{\alpha}{\alpha+1})a - b$. Meanwhile, as $-a < -b < 0$, from Case 1) we have
$$-\left(\frac{-b}{\alpha+1}\right)\left(\frac{-b}{-a}\right)^\alpha \ge \left(\frac{\alpha}{\alpha+1}\right)(-a) - (-b).$$
In other words, $(\alpha/(\alpha+1))a - b + (b/(\alpha+1))(b/a)^\alpha \ge 0$. Hence, $(\alpha/(\alpha+1))a - b + (b/(\alpha+1))(b/a)^\alpha \ge [(\frac{\alpha}{\alpha+1})a - b]^+$, i.e., $\widetilde{f}(\xi) \ge \widetilde{f}_L(\xi)$.

As $\widetilde{f}(\xi) \ge \widetilde{f}_L(\xi)$ for all $\xi \in \mathbb{R}^k$, the inequality (19) holds based on Lemma 4.1. ■

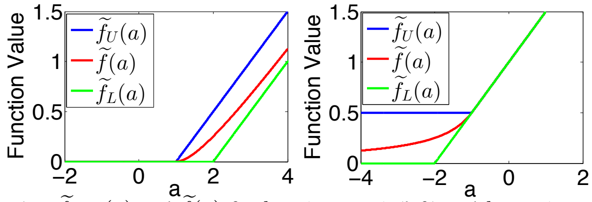

Fig. 1. $\widetilde{f}_{L,U}(a)$ and $\widetilde{f}(a)$ for $b = 1,\ \alpha = 1$ (left) and $b = -1,\ \alpha = 1$ (right).

Figure 1 shows the relationship between functions $\widetilde{f}_{L,U}(\xi)$ and $\widetilde{f}(\xi)$. We can also observe that, as $\alpha \to \infty$, all three curves converge and the approximation errors shrink to zero.

V. SDP Reformulation

In this section, we reformulate the DR-CVaR constraint (13) as an SDP based on the two approximations derived in Section IV. More specifically, we obtain a conservative approximation of (13) based on $\widetilde{f}_U(\xi)$, and a relaxed one based on $\widetilde{f}_L(\xi)$. Before summarizing these results, we review a celebrated reformulation result from the literature.

Lemma 5.1: (Adapted from [12]) For random function $L(\xi) = y_0(x) + y(x)^T\xi$, $\sup_{\mathbb{P}\in\mathcal{P}(\mu,S)} \mathbb{E}_{\mathbb{P}}\left[L(\xi) - \beta\right]^+$ equals the optimal objective value of the following SDP.

$$\begin{aligned} \inf_{M\in\mathbb{S}^+_{k+1}} \quad & \mathrm{Tr}(\Omega\cdot M) \\ \text{s.t.} \quad & M - \begin{bmatrix} 0 & \frac{1}{2}y(x) \\ \frac{1}{2}y(x)^T & y_0(x) - \beta \end{bmatrix} \succeq 0, \end{aligned} \tag{20}$$

where $\mathbb{S}^+_{k+1}$ represents the set of all $(k+1)\times(k+1)$ symmetric and positive semidefinite matrices and $\Omega = \begin{bmatrix} S & \mu \\ \mu^T & 1 \end{bmatrix}$.

We are now ready to state the SDP reformulations based on the approximations in Section IV as follows.

Theorem 5.1: The DR-CVaR constraint (13) is implied by the following SDP constraints:

$$\begin{aligned} \beta + \left(\frac{1}{\epsilon}\right)\left(\frac{\alpha}{\alpha+1}\right)\mathrm{Tr}(\widetilde{\Omega}\cdot M) \ &\le\ 0 \\ \beta + \frac{1}{\epsilon(\alpha+1)}(y_0(x) - \beta) + \\ \left(\frac{1}{\epsilon}\right)\left(\frac{\alpha}{\alpha+1}\right)\mathrm{Tr}(\widetilde{\Omega}\cdot M) \ &\le\ 0 \\ M - \begin{bmatrix} 0 & \frac{1}{2}y(x) \\ \frac{1}{2}y(x)^T & y_0(x) - \beta \end{bmatrix} &\succeq 0, \end{aligned} \tag{21}$$

where $\widetilde{\Omega} = \begin{bmatrix} \widetilde{S} & \widetilde{\mu} \\ \widetilde{\mu}^T & 1 \end{bmatrix}$.

Proof: Based on Theorem 4.1, we have

$$\begin{aligned} & \inf_{\beta\in\mathbb{R}} \left\{ \beta + \frac{1}{\epsilon}\sup_{\mathcal{P}} \mathbb{E}_{\mathbb{P}}\left[L(\xi) - \beta\right]^+ \right\} \\ \le\ & \inf_{\beta\in\mathbb{R}} \Bigg\{ \beta + \frac{1}{\epsilon(\alpha+1)}(y_0(x) - \beta)^+ + \\ & \left(\frac{1}{\epsilon}\right)\left(\frac{\alpha}{\alpha+1}\right) \sup_{\mathbb{P}\in\mathcal{P}(\widetilde{\mu},\widetilde{S})} \mathbb{E}_{\mathbb{P}}\left[L(\xi) - \beta\right]^+ \Bigg\}. \end{aligned}$$

The proof is completed by rewriting the worst-case expectation on the right-hand side using Lemma 5.1. ■

Theorem 5.2: The DR-CVaR constraint (13) implies the following SDP constraints:

$$\begin{aligned} & \beta + \frac{1}{\epsilon}\mathrm{Tr}(\widetilde{\Omega}\cdot M) \le 0 \\ & M - \begin{bmatrix} 0 & \frac{1}{2}\left(\frac{\alpha}{\alpha+1}\right)y(x) \\ \frac{1}{2}\left(\frac{\alpha}{\alpha+1}\right)y(x)^T & y_0(x) - \beta \end{bmatrix} \succeq 0. \end{aligned} \tag{22}$$

Proof: Based on Theorem 4.2, we have

$$\begin{aligned} & \inf_{\beta\in\mathbb{R}} \left\{ \beta + \frac{1}{\epsilon}\sup_{\mathcal{P}} \mathbb{E}_{\mathbb{P}}\left[L(\xi) - \beta\right]^+ \right\} \ \ge \\ & \inf_{\beta\in\mathbb{R}} \left\{ \beta + \frac{1}{\epsilon} \sup_{\mathbb{P}\in\mathcal{P}(\widetilde{\mu},\widetilde{S})} \mathbb{E}_{\mathbb{P}}\left[L_s(\xi) - \beta\right]^+ \right\} \end{aligned} \tag{23}$$

where $L_s(\xi) = y_0(x) + \left(\frac{\alpha}{\alpha+1}\right)y(x)^T\xi$. Based on Lemma 5.1, the right-hand side can be further recast as

$$\begin{aligned} \inf_{\beta\in\mathbb{R}, M\in\mathbb{S}^+_{k+1}} \quad & \beta + \frac{1}{\epsilon}\mathrm{Tr}(\widetilde{\Omega}\cdot M) \\ \text{s.t.} \quad & M - \begin{bmatrix} 0 & \frac{1}{2}\left(\frac{\alpha}{\alpha+1}\right)y(x) \\ \frac{1}{2}\left(\frac{\alpha}{\alpha+1}\right)y(x)^T & y_0(x) - \beta \end{bmatrix} \succeq 0 \end{aligned}$$

The proof is completed. ■

VI. Case Study

A. Simulation Setup

The theoretical results are evaluated on a modified uncongested IEEE 9-bus system [25]. We add two large wind power plants on buses 1 and 2, and assume a wind power forecast at each bus of 100 MW. We increase the total load consumption in the system by 50%. We assume the wind forecast errors at each wind power plant are independent, have standard deviations of 10% of the forecast value, and modes at the origin (i.e., the wind forecast). For cost parameters, we assume C_3 is 10 times C_2, where C_1 and C_2 are defined in [25]. For violation probability, we assume $\epsilon = 5\%$. All optimization problems are solved by CVX with the Mosek solver [26] [27].

B. SDP Results

To check the performance of the reformulations in Theorems 5.1 and 5.2, we compare the objective costs with that of the original problem, specifically the conventional CVaR problem without the assumption of α-unimodality. The computational effort required by each approach is similar. For systematic comparison, we fix the covariance matrix $\Gamma = S - \mu\mu^T$ and sweep over feasible values for μ, which corresponds to the means of the wind forecast errors w_1 and w_2. Feasibility can be ensured if $\widetilde{\Omega} \succeq 0$ [17].

Figure 2 shows that the approximation in Theorem 5.1 provides conservative results when the means of w_1 and w_2 are both close to the origin. However, in this same case, the approximation in Theorem 5.2 provides a lower bound, which is close to the original cost. This means that, when the mean is

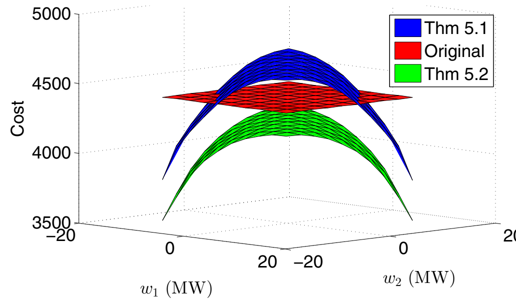

Fig. 2. Objective value comparison between the approximations in Theorems 5.1 and 5.2 with $\alpha = 1$, and the original case without the unimodality assumption.

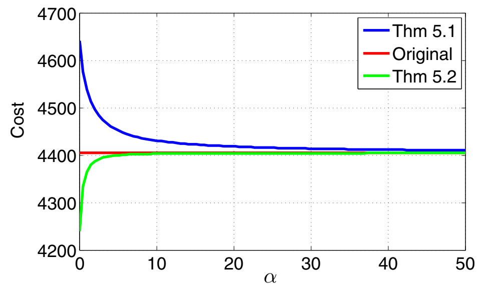

Fig. 3. When $\mu = [0; 0]$ and with increasing α, the original case without the unimodality assumption and the approximations in Theorems 5.1 and 5.2 converge.

close to the mode, the mode does not have a large effect on the worst-case distribution, which determines the cost. When the means are further from the mode, both Theorems 5.1 and 5.2 give better objective values compared to the original cost. The reason is that the mode at the origin restricts the worst-case distribution to be less extreme than the worst-case distribution without the unimodality assumption.

Figures 3 and 4 show that as α increases, the results of the three approaches converge because the mode has less effect on the cost. Figure 4 shows that when the mean and mode are far apart and $\alpha \leq k$, the mode has a strong effect on the worst-case distribution, resulting in lower cost. However, when $\alpha > k$, the effect of the mode becomes weaker. The reason is that when α increases, the feasible region $\widetilde{\Omega} \succeq 0$ will expand and the same mean and mode will have less of an effect on the solution. As $\alpha \to \infty$, $\widetilde{\Omega} \to \Omega$ and, as we only fix the covariance matrix Γ, the feasible region for μ becomes $\mathbb{R}^k$. However, Theorem 5.1 can still provide better solution than that of the original case when α is large if the mean is far from the mode.

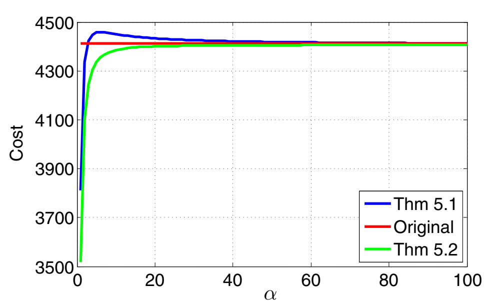

Fig. 4. When $\mu = [12; 12]$ MW and with increasing α, the original case without the unimodality assumption and the approximations in Theorems 5.1 and 5.2 converge.

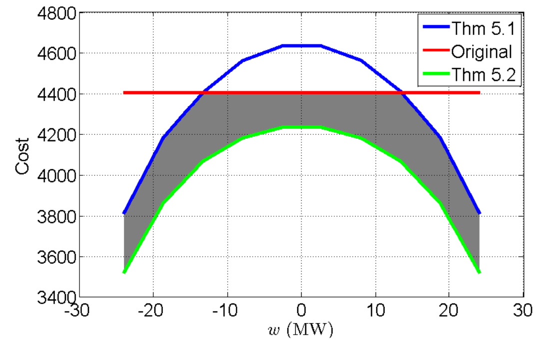

Fig. 5. Univariate objective cost comparison between the results of Theorems 5.1 and 5.2 with $\alpha = 1$, and the original case without the unimodality assumption.

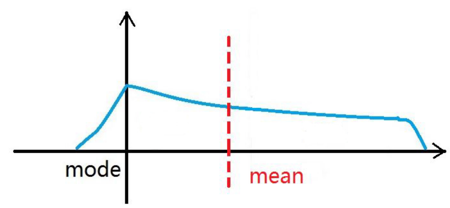

Fig. 6. Distribution with mean of w far from the mode.

1) Univariate Interpretation: From (4), another interpretation would be to treat $w = w_1 + w_2$ as a univariate distribution. Then, we use the following theorem.

Theorem 6.1: (Adapted from [22]). If random vector $\xi \in \mathbb{R}^k$ has a α-unimodal distribution with some fixed α, and A is a linear mapping from $\mathbb{R}^k$ to $\mathbb{R}^q$, then $\widetilde{\xi} = A\xi$ has a α-unimodal distribution with the same α.

Then, we have the following property.

$$\inf_{\mathbb{P} \in \mathcal{P}^{\alpha}_{\mu,S}} \mathbb{P}(L(A\xi) \leq 0) = \inf_{\mathbb{P} \in \mathcal{P}^{\alpha}_{A\mu,ASA^T}} \mathbb{P}(L(\widetilde{\xi}) \leq 0),$$

where $\mathcal{P}^{\alpha}_{\mu,S}$ denotes the set of α-unimodal probability distributions with first two moments μ and S, respectively. In our case, $A = [1\ 1]$ and so we can directly obtain results for the univariate case. Figure 5 shows that when the mean of w is further from the mode, the approximations in Theorems 5.1 and 5.2 provide better objective costs than the original case. Additionally, the approximations become closer to one another. With Theorems 5.1 and 5.2, the DR-CVaR is bounded within the shaded region in Fig. 5. In the univariate case with increasing α, similar results are observed to those in Figs. 3 and 4.

An empirical explanation for the influence of the mode is as follows. In Fig. 6, assume we have a positive mean and a mode at the origin. Using the location of the mean, we know the cumulative probability below that will be limited. To choose R^{up}, we want to estimate the largest possible magnitude of negative total wind error. Due to the existence of the mode, the tail region on the negative axis will be more restrictive compared to the case without the mode assumption. Hence, we require less up reserve capacity. This restriction will become more severe as the mean becomes further from mode, as shown in our previous results.

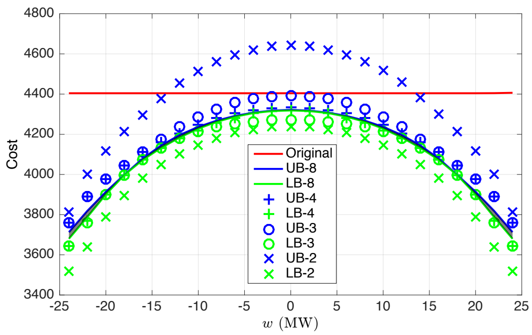

Fig. 7. Upper bounds (UB) and lower bounds (LB) for different numbers of as number of cuts (e.g., UB-8 refers to the upper bound with 8 cuts), compared with original case without the unimodality assumption.

C. A Multi-Cut Approximation

We extend the 2-piece approximation described in Section IV to a multi-cut one by approximating function $\tilde{f}(\xi)$ with multiple linear pieces. By increasing the number of approximation pieces, the approximation error reduces, leading to a tighter confidence region for the real optimal value. The multi-cut approximations can also be exactly reformulated as SDP problems. The resulting SDP formulation and its proof are similar to those of the 2-piece approximation and so omitted here due to the space limit. Figure 7 shows that as we increase the number of linear pieces in the multi-cut approximation, better upper and lower bounds can be achieved. Additionally, by comparing with Fig. 5, we observe that the approximation quickly converges to the real optimal value as we increase the number of linear pieces.

VII. Conclusions and Future Work

In this paper, we included generalized α-unimodality into the distributionally robust risk-constrained optimization problem assuming known moment information. We proposed two tractable approximate reformulations, an upper bound and a lower bound, that we expressed as SDPs. While the computational effort for the problems is similar to the original problem, including the unimodality assumption results in less-conservative, lower-cost solutions when the mode is far from the mean. We also observe that the solutions generated under the assumption of α-unimodality will converge to the solution generated without assuming unimodality as α increases. Finally, we presented the results of a multi-cut approximation which provides tighter bounds on the objective cost.

In the future, we will attempt to derive an exact reformulation of the RCOPF. We will also develop approaches to derive joint DR-CVaR constraints with unimodality, include other properties like bounded supports and moment ambiguity, and test the results on more realistic systems with higher uncertainty dimensions and correlations to determine if the performance of the theorems is case-dependent.

References

[1] H. Zhang and P. Li, "Probabilistic analysis for optimal power flow under uncertainty," *IET Generation, Transmission and Distribution*, vol. 4(5), pp. 553-561, May. 2010.

[2] H. Zhang and P. Li, "Chance Constrained Programming for Optimal Power Flow under Uncertainty", *IEEE Trans. Power Syst.*, Vol. 26, No. 4, Nov. 2011.

[3] R. A. Jabr, "Adjustable robust OPF with renewable energy sources," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 47424751, Nov. 2013.

[4] M. Vrakopoulou, K. Margellos, J. Lygeros, and G. Andersson, "A Probabilistic Framework for Reserve Scheduling and N-1 Security Assessment of Systems With High Wind Power Penetration," *IEEE Trans. Power Syst.*, 28(4): 3885-3896, 2013.

[5] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson, "Analytical reformulation of security constrained optimal power flow with probabilistic constraints," *Proceedings of PowerTech*, 2013.

[6] D. Bienstock, M. Chertkov, and S. Harnett, "Chance-Constrained Optimal Power Flow: Risk-Aware Network Control under Uncertainty," *SIAM Review*, 56(3): 461-495, 2014.

[7] M. Vrakopoulou, J.L. Mathieu, and G. Andersson, "Stochastic Optimal Power Flow with Uncertain Reserves from Demand Response," *Proceedings of HICSS*, pp. 2353-2362, 2014.

[8] B. Li and J.L. Mathieu. "Analytical Reformulation of Chance-Constrained Optimal Power Flow with Uncertain Load Control," *Proceedings of the PowerTech*, Eindhoven, Netherlands, June 2015.

[9] Marco C. Campi, Simone Garatti, Maria Prandini, "The scenario approach for systems and control design," *Annual Reviews in Control*, 33(2): 149-157, 2009.

[10] K. Margellos, P. Goulart, and J. Lygeros, "On the Road Between Robust Optimization and the Scenario Approach for Chance Constrained Optimization Problems," *IEEE Transactions on Automatic Control*, 59(8): 2258-2263, 2014.

[11] E. Delage and Y. Ye, "Distributionally Robust Optimization under Moment Uncertainty with Application to Data-driven problems," *Operations Research*, vol. 58, no. 3, pp. 595-612, 2010

[12] S. Zymler, D. Kuhn and B. Rustem, "Distributionally Robust Joint Chance Constraints with Second-order Moment Information," *Mathematical Programming* 137(1-2), 167-198, 2013.

[13] Y. Zhang, S. Shen, and J.L. Mathieu, "Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads," *IEEE Trans. Power Syst. (in press)*, 2016.

[14] R. Jiang and Y. Guan, "Data-driven Chance Constrained Stochastic Program," *Mathematical Programming*, pp1-37, 2015.

[15] L. Roald, F. Oldewurtel, B.V. Parys, and G. Andersson, "Security Constrained Optimal Power Flow with Distributionally Robust Chance Constraints," Arxiv, http://arxiv.org/abs/1508.06061, 2015.

[16] T. Summers, J. Warrington, M. Morari, J. Lygeros, "Stochastic optimal power flow based on conditional value at risk and distributional robustness," *International Journal of Electrical Power & Energy Systems*, vol. 72, pp. 116-125, 2015.

[17] B. Van Parys, P. Goulart and D. Kuhn, "Generalized Gauss Inequalities via Semidefinite Programming," *Mathematical Programming*, pp 1-32, 2015.

[18] B. Van Parys, P. Goulart and M. Morari, "Distributionally Robust Expectation Inequalities for Structured distributions," *Optimization Online* 2015.

[19] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, "Coherent Measures of Risk," *Mathematical Finance*, 9: 203228, 1999.

[20] R.T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking and Finance*, 26(7):1443-1471, 2002.

[21] R.T. Rockfellar and S. Uryasev, "Optimization of conditional value-at-risk," J.Risk 2, 21-41, 2002.

[22] S.W. Dharmadhikari and K. Joat-Dev, "Unimodality, Convexity and Application," *Probability and Mathematical Statistics*, vol. 27. Academic Press, 1988.

[23] A. Shapiro and A.J. Kleywegt, "Minimax analysis of stochastic problems," *Optim. Methods Softw.* 17(3), 523-542, 2002.

[24] I. Popsecu, "A Semidefinite Programming Approach to Optimal Moment Bounds for Convex Classes of Distributions," *Mathematics of Operations Research* 30(3), 632-657, 2005.

[25] R.D. Zimmerman, C.E. Murillo-Sanchez, and R.J. Thomas. "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education." *IEEE Trans. Power Syst.*, 26(1):12-19, 2011.

[26] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.0 beta," http://cvxr.com/cvx, September 2013.

[27] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs, Recent Advances in Learning and Control," V. Blondel, S. Boyd, and H. Kimura, editors, pages 95-110, Lecture Notes in Control and Information Sciences, Springer, 2008.