Analyzing and Quantifying the Intrinsic Distributional Robustness of CVaR Reformulation for Chance Constrained Stochastic Programs

Author(s): Cao, Yang; Wei, Wei; Mei, Shengwei; Shafie-khah, Miadreza; Catalao, Joao P.S.

Title: Analyzing and Quantifying the Intrinsic Distributional Robustness of CVaR Reformulation for Chance Constrained Stochastic Programs

Year: 2020

Version: Accepted version

Copyright ©2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Please cite the original version:

https://doi.org/10.1109/TPWRS.2020.3021285
Analyzing and Quantifying the Intrinsic Distributional Robustness of CVaR Reformulation for Chance Constrained Stochastic Programs

Yang Cao, Student Member, IEEE, Wei Wei, Senior Member, IEEE, Shengwei Mei, Fellow, IEEE
Miaadreza Shafie-khah, Senior Member, IEEE, João P. S. Catalão, Senior Member, IEEE

Abstract—Chance constrained program (CCP) is a popular stochastic optimization method in power system planning and operation problems. Conditional Value-at-Risk (CVaR) provides a convex approximation for chance constraints which are non-convex. Although CCP assumes an exact empirical distribution, and the optimum of a stochastic programming model is thought to be sensitive in the designated probability distribution, this letter discloses that CVaR reformulation of chance constraint is intrinsically robust. A pair of indices are proposed to quantify the maximum tolerable perturbation of the probability distribution, and can be computed from a computationally-cheap dichotomy search. An example on the coordinated capacity optimization of energy storage and transmission line for a remote wind farm validates the main claims. The above results demonstrate that stochastic optimization methods are not necessarily vulnerable to distributional uncertainty, and justify the positive effect of the conservatism brought by the CVaR reformulation.

Index Terms—chance constraint, conditional-value-at-risk, distributional robustness, stochastic optimization, uncertainty

I. INTRODUCTION

WITH the proliferation of renewable energy resources, the uncertainty in power system is growing rapidly. Chance constrained program (CCP) is a popular stochastic optimization method in power system planning and operation related problems. A CCP is generally formulated as

$$\min_{x \in X} f(x)$$

subject to

$$\Pr \{ h(x, \xi) - \lambda \leq 0 \} \geq 1 - \alpha$$

where $x$ is the vector of decision variables and $X$ represents physical operating constraints; $\xi$ denotes uncertain parameters; $h(x, \xi)$ is a loss function, and $\lambda$ is the cap of loss. Chance constraint (1b) requires the probability of event $h(x, \xi) - \lambda \leq 0$ evaluated at the empirical probability distribution function (PDF) $P_0$ of $\xi$ should be at least $1 - \alpha$.

Two difficulties exist in CCP model (1). The first one is that the empirical PDF $P_0$ to validate chance constraint (1b) is generally inexact. Therefore, for the optimum $x^*$ of model (1), the probability $Pr \{ h(x^*, \xi) - \lambda \leq 0 \}$ evaluated at the real PDF in practice can be either greater or smaller than $1 - \alpha$. The other is that chance constraint (1b) does not have a closed-form expression and is generally non-convex in decision variable $x$, making a CCP hard to solve. A chance constraint on $x$ is equivalent to imposing an upper bound on a Value-at-Risk (VaR) index [1], which can be further transformed to a mixed-integer linear program (MILP) based on sampling average approximation [2].

This work was supported in part by NSFC 51807101 and in part by the energy storage project from State Grid Corporation of China under Special Grant 522800180003. Corresponding to: W. Wei and S. Mei.

M. Shafie-khah is with the School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland (e-mail: miaadrea@gmail.com).
J.P.S. Catalão is with the Faculty of Engineering and INESC TEC, University of Porto, Porto 4200-465, Portugal (e-mail: catalao@fe.up.pt).
In this process, the range of PDF perturbation is unknown, and the maximal range is to be determined. For any strictly feasible \( x^* \) other than the optimal one, the proposed method can give two indices reflecting its robustness for the chance constraint. So the result of evaluation is decision-dependent.

**II. QUANTIFYING THE DISTRIBUTATIONAL ROBUSTNESS**

In risk theory, the quantile for the probability of loss \( h(x, \xi) \) no greater \( 1 - \alpha \) is defined as the Value-at-Risk (VaR)

\[
(1 - \alpha)\text{-Var}(x, \xi, P_0) = \min \left\{ z \in \mathbb{R} \mid \int_{h(x, \xi) \leq z} p_0(\xi) d\xi \geq 1 - \alpha \right\}
\]

where \( p_0(\xi) \) is the probability density function of \( P_0 \). By this definition, constraint

\[
(1 - \alpha)\text{-Var}(x, \xi, P_0) \leq \lambda
\]

is equivalent to the non-convex chance constraints (1b) for \( x \). So constraint (3) is also non-convex.

The most widely used risk measure for approximating a chance constraint is CVaR, which is defined as the conditional expectation of loss no less than the VaR, i.e.,

\[
(1 - \alpha)\text{-CVaR}(x, \xi, P_0) = \frac{1}{\alpha} \int_{h(x, \xi) \geq (1 - \alpha)\text{-Var}(x, \xi, P_0)} h(x, \xi) p_0(\xi) d\xi
\]

The loss less than VaR does not contribute to CVaR, hence

\[
(1 - \alpha)\text{-Var}(x, \xi, P_0) < (1 - \alpha)\text{-CVaR}(x, \xi, P_0)
\]

The advantage of CVaR is its convexity [3]. In this regard, chance constraint (1b) can be conservatively approximated by

\[
(1 - \alpha)\text{-CVaR}(x, \xi, P_0) \leq \lambda
\]

Conservatism means if (6) holds, then (3) must hold because of (5), and so does (1b), given the equivalence between (1b) and (3). In other words, the feasible set of \( x \) defined by (6) is smaller than that induced by (3). CVaR constraint (6) usually comes down to linear equalities in most practical optimization problems [3] after performing sampling average approximation. In the following, we discuss how much distributionally robustness is brought by the CVaR reformulation.

To characterize the distributional robustness, it is essential to define the distance between two PDFs. Suppose \( p_0(\xi)/p(\xi) \) is the probability density function of \( P_0 \) and \( P \), then the Kullback-Leibler (KL) divergence defined by

\[
D_{KL}(P \mid \mid P_0) = \int_{\Omega} p(\xi) \log \frac{p(\xi)}{p_0(\xi)} d\xi
\]

is a measure on the distance between \( P \) and \( P_0 \). For discrete distributions on a set of representative scenarios \( \{\pi_n\}_{n=1}^N \), the KL divergence has the form of

\[
D_{KL}(P \mid \mid P_0) = \sum_{n=1}^N \pi_n \log \frac{\pi_n}{\pi_0^n}
\]

where \( \pi_n/\pi_0^n \) is the probability of \( \xi_n \) under distribution \( P/P_0 \). With this divergence measure, we can explain the meaning of distributional robustness.

Consider \( x^* \) satisfying constraint (6). Evaluate the probability of event \( h(x^*, \xi) - \lambda \leq 0 \) at \( P_0 \), the result is:

\[
\Pr_{\{h(x^*, \xi) - \lambda \leq 0\}} = 1 - \alpha^*
\]

where \( \alpha^* \) must be smaller than \( \alpha \), because of the conservatism introduced by (5). Denoted by \( \bar{\alpha} = \alpha - \alpha^* \) the margin of reliability. Furthermore, if the actual distribution varies in

\[
W = \{P|D_{KL}(P \mid \mid P_0) \leq d_M\}
\]

Then, \( \Pr_{\{h(x^*, \xi) - \lambda \leq 0\}} \) evaluated at distribution \( P \) may fluctuate around \( 1 - \alpha^* \) but is still no less than \( 1 - \alpha \) due to the margin \( \bar{\alpha} \). Let \( d_M \) be the maximum value such that

\[
\inf_{P \in W} \Pr_{\{h(x^*, \xi) - \lambda \leq 0\}} = 1 - \alpha
\]

The pair \((\bar{\alpha}, d_M)\) quantifies the conservatism and distributional robustness brought by the CVaR reformulation.

To compute the proposed indices \((\bar{\alpha}, d_M)\), we have to find the relationship among \( \alpha \), \( \alpha^* \) and \( d_M \), which can be derived from Theorem 1 and Proposition 4 in [7]. In brief, if (11) holds with fixed \( \alpha \) and \( d_M \), then the probability of event \( h(x^*, \xi) - \lambda \leq 0 \) evaluated at \( P_0 \) is \( 1 - \alpha^* \), where

\[
\alpha^* = \max \left\{0, 1 - \inf_{z \in (0,1)} \left\{ \frac{e^{-d_M(z^{1-\alpha}-1)}}{z-1} \right\} \right\}
\]

where the univariate function \( \psi(z) = (e^{-d_M(z^{1-\alpha}-1)})/(z-1) \) is convex in \( z \) over the open interval \((0,1)\) [7], so the minimum can be easily found using one-dimensional search. In the algorithmic implementation, we use the renowned golden section search algorithm which entails only function value comparison and is very efficient.

Equation (12) predicts \( \alpha^* \) from the known values of \( \alpha \) and \( d_M \), while in the proposed method, both of \( \alpha \) and \( d_M \) are unknown. Based on relation (12), a dichotomy algorithm is proposed to compute the indices \( \bar{\alpha} \) and \( d_M \). The flowchart is given in Algorithm 1. The robustness quantification is performed by steps 3–6; the solution \( x^* \) is offered by the CVaR reformulation model in step 1. Because no complex optimization problem is solved in the main loop of steps 3–6, Algorithm 1 is actually very fast.

**III. COORDINATED CAPACITY PLANNING PROBLEM**

Consider a remote wind farm shown in Fig. 1, and its capacity is known. The power grid company plans to build a local energy storage unit and a transmission line connecting the wind farm with the power grid. The target is to minimize
Algorithm 1

1. Replace chance constraint (1b) with CVaR constraint (6), solve the following problem, the optimal solution is $x^*$:

$$\min \{ f(x) (1-\alpha)\text{-CVaR}(x, \xi, P_0) \leq \lambda \}$$

2. Evaluate probability $P_0 \{ h(x^*, \xi - \lambda) \leq 0 \}$ at $P_0$, and the result is $1-\alpha$; then $\alpha = \alpha - \alpha^*$;

3. Choose a convergence tolerance $\varepsilon > 0$; Initialize the upper and lower bounds $d_M^{\text{min}}, d_M^{\text{max}}$;

4. Set $d_M = (d_M^{\text{max}} + d_M^{\text{min}})/2$ and solve

$$\beta = \max \left\{ 0, 1 - \inf_{z \in (0,1)} \left\{ \frac{e^{-d_M z(1-\alpha - 1)}}{z - 1} \right\} \right\}$$

5. If $\beta > \alpha$, update $d_M^{\text{min}} = d_M$; else update $d_M^{\text{max}} = d_M$.

6. If $|d_M^{\text{max}} - d_M^{\text{min}}| < \varepsilon$, terminate and report $d_M$ as the final solution; else go to step 4.

the total investment cost while ensuring that wind power curtailment rate does not exceed a threshold. Here, we first give the definition of wind power curtailment rate $h(x, \xi)$, which is the optimal value of the following problem

$$h(x, \xi) = \min \left\{ \sum_{t=1}^T (\xi_t - p_t^w) \right\}$$

s.t. $p_t^l = p_t^w + p_t^d - p_t^c$

$$0 \leq p_t^w \leq \xi_t, 0 \leq p_t^l \leq x_t$$

$$W_t^E = W_t^E + \left( p_t^c \eta_c^E - p_t^d \eta_d^E \right) \Delta t$$

$$0 \leq [p_t^c, p_t^d, W_t^E] \leq [R_c, x_t, R_d, x_t]$$

where $T$ is the number of periods; $x = [x_t(MWh), x_t(MW)]$ are the capacity of energy storage unit and transmission line; $\xi_t$ is the maximal wind power output depending on wind speed, which is uncertain. Decision variables include the dispatched wind power $p_t^w$, the transported power $p_t^l$, the charging/discharging power $p_t^c/p_t^d$, and the energy $W_t^E$ in the storage unit. (13b) prescribes power balance; (13c) limits the dispatched wind power and transmitted power; (13d) represents the charging dynamics of storage unit; (13e) imposes bounds on storage charging power, discharging power, and storage level; $1 - \mu_t^E$ is the self-discharge rate, $\eta_c^E/\eta_d^E$ is the charging/discharging efficiency, and $R_c$ and $R_d$ are constants.

In the coordinated planning problem, we aim to maintain the wind power curtailment rate $h(x, \xi)$ within a threshold $\lambda$, which can be interpreted as an event $h(x, \xi) - \lambda \leq 0$. The probability of this event depends on the distribution of $\xi$, yielding the following CCP

$$\min C_s x_s + C_l x_l$$

s.t. $x_s \geq 0, x_l \geq 0$

$$P_0 \{ h(x, \xi) - \lambda \leq 0 \} \geq 1 - \alpha$$

where $C_s(S/MWh)/C_l(S/MW)$ is the unit capacity cost of the energy storage unit and transmission line. Chance constraint (14c) requires that under the empirical distribution $P_0$, the probability of event $h(x, \xi) - \lambda \leq 0$ is at least $1-\alpha$. According to the discussions in Section II, the chance constraint (14c) is approximated by the following CVaR constraint

$$(1-\alpha)-\text{CVaR}(x, \xi, P_0) \leq \lambda$$

Suppose we have $N$ typical scenarios $\xi^1, \xi^2, \ldots, \xi^N$ with probabilities $\pi^1, \pi^2, \ldots, \pi^N$, respectively. Decision variables and constraints in (13b)-(13e) are duplicated for each scenario with a script $n$. Afterwards, the CVaR constraint (15) comes down to the following linear forms [3]

$$\sum_{t=1}^T (\xi_t - p_t^{w,n}) - \lambda + \gamma \leq \sigma^n, \forall n$$

$$\sum_{n=1}^N \pi^n \sigma^n \leq \gamma \alpha, \sigma^n \geq 0, \forall n$$

where $\sigma^n (n = 1, 2, \ldots, N)$ and $\gamma$ are auxiliary variables.

Replacing (14c) with (16), which is known as the sampling average approximation of CVaR constraint (15), the coordinated planning model (14) finally comes down to

$$\min C_s x_s + C_l x_l$$

s.t. $x_s \geq 0, x_l \geq 0$

$$p_t^{l,n} = p_t^{w,n} + p_t^{d,n} - p_t^{c,n}$$

$$0 \leq p_t^{w,n} \leq \xi_t, 0 \leq p_t^{l,n} \leq x_t$$

$$W_t^{E,n} = W_t^{E,n} \mu_t + \left( p_t^{c,n} \eta_c^E - p_t^{d,n} \eta_d^E \right) \Delta t$$

$$0 \leq [p_t^{c,n}, p_t^{d,n}, W_t^{E,n}] \leq [R_c, x_t, R_d, x_t]$$

$$\sum_{t=1}^T (\xi_t - p_t^{w,n}) - \lambda + \gamma \leq \sigma^n$$

$$\sum_{n=1}^N \pi^n \sigma^n \leq \gamma \alpha, \sigma^n \geq 0$$

Problem (17) is a linear program, which can be efficiently solved by off-the-shelf solvers.

IV. SIMULATIONS RESULTS

Consider the scene in Fig. 1 with a 100MW wind farm. Parameters of energy storage unit and transmission line are shown in I. Based on the real weather data from wind farms in Qinghai province, China, the predicted wind power outputs in four typical days are extracted from spring, summer, autumn and winter to represent a whole year. Assuming the forecast errors obey Gaussian distribution with zero mean, and the standard deviation is 0.2 multiplying the forecast values. Next, 5000 scenarios are generated via Monte Carlo method to construct the empirical PDF $P_0$. The wind power output data is available at [10]. In the numerical experiments, the cap of curtailment rate $\lambda = 0.05$ and the risk level $\alpha = 10\%$. All optimization models are established in MATLAB 2018b and solved by CPLEX 12.8 with YALMIP interface.

The optimal solution of model (17) is $C_s = 33.37$ MWh and $C_l = 58.08$ MW, which is denoted by $x^*$. The curtailment rate requirement $h(x^*, \xi) - \lambda \leq 0$ holds with a probability $1-\alpha^* = 96.01\%$ under the empirical distribution $P_0$, so the reliability margin is $\bar{\alpha} = 6.01\%$. By executing Algorithm 1, the value of
The time for calculating the pair of indices \( x_1 \) is no less than \( d \), VaR reformation does not warrant distributional robustness. \( 1 \) is less conservative, and the distributional robustness is weaker. level increase of \( 1 \) in Table II. It is observed that \( D_\alpha \) is 2.04s, demonstrating the efficiency of Algorithm 1.

To validate our conclusion, we generate data sets consisting of 10000 scenarios whose distribution is very different from the empirical distribution \( P_0 \), while maintaining \( D_{KL}(P||P_0) = 0.0335 \). Then we test the probability \( \Pr(h(x^*, \xi) - \lambda \leq 0) \) under this unfavorable distribution \( P \); the result is 90.27%, which is slightly higher than the anticipated value \( 1 - \alpha = 90\% \), validating that the proposed index \( d_M \) is accurate. For comparison, the VaR reformation of CCP model (14), which is an MILP [2], is solved; the optimal solution \( x' \) gives \( C_s = 31.96 \text{ MWh} \) and \( C_I = 56.21 \text{ MW} \). Both of them are smaller than those in \( x^* \). Then we test the probability \( \Pr(h(x', \xi) - \lambda \leq 0) \) under the aforementioned unfavorable distribution \( P \); the result is 81.92%, which is much lower than the anticipated value \( 1 - \alpha = 90\% \). This fact means that the optimal solution of VaR model is not robust and unable to protect system security with a predefined reliability level if the distribution used in the model is inexact.

The risk level \( \alpha \) remarkably influences the optimal solution of the CCP model. In the following tests, we investigate the impact of \( \alpha \) on the distributional robustness. First, the value of \( 1 - \alpha^* \), \( \bar{\alpha} \) and \( d_M \) is calculated from Algorithm 1 with different values of \( \alpha \). Next, a series of unfavorable distributions with corresponding KL-divergence distances to \( P_0 \) is generated. We test \( \Pr(h(x^*, \xi) - \lambda \leq 0)/\Pr(h(x', \xi) - \lambda \leq 0) \), termed as the test probability for CVaR/VaR. Results are summarized in Table II. It is observed that \( \bar{\alpha} \) and \( d_M \) decrease with the increase of \( 1 - \alpha \), implying that when the targeted reliability level \( 1 - \alpha \) grows higher, the CVaR reformulation becomes less conservative, and the distributional robustness is weaker. The test probability for CVaR is always slightly higher than \( 1 - \alpha \), demonstrating the effectiveness of the proposed characterization on distributional robustness. In contrast, the test probability for VaR is always lower than \( 1 - \alpha \), because the VaR reformation does not warrant distributional robustness.

Finally, we investigate the impact of curtailment rate cap \( \lambda \) on the distributional robustness indices. Similar to the above analysis, we calculate the values of \( 1 - \alpha^* \), \( \bar{\alpha} \), \( d_M \) and test probabilities with different \( \lambda \) while remaining \( 1 - \alpha = 90\% \). The results are summarized in Table III. It is observed that the values of \( 1 - \alpha^* \), \( \bar{\alpha} \), \( d_M \) and test probability for CVaR and VaR almost keep unchanged with different values of \( \lambda \), which means that the distributional robustness of CVaR reformulation are barely affected by the cap of curtailment rate \( \lambda \).

### Table I

<table>
<thead>
<tr>
<th>Parameter Settings</th>
<th>( C_s ) (S/MWh)</th>
<th>( C_l ) (S/MWh)</th>
<th>( \mu^E )</th>
<th>( \eta_c )</th>
<th>( \eta_d )</th>
<th>( R_c )</th>
<th>( R_d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 2 \times 10^5 )</td>
<td>( 1.2 \times 10^5 )</td>
<td>0.99</td>
<td>0.95</td>
<td>0.95</td>
<td>0.25</td>
<td>0.25</td>
<td></td>
</tr>
</tbody>
</table>

### Table II

<table>
<thead>
<tr>
<th>RESULTS WITH DIFFERENT RISK LEVEL ( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - ( \alpha )</td>
</tr>
<tr>
<td>1 - ( \alpha^* )</td>
</tr>
<tr>
<td>( \bar{\alpha} )</td>
</tr>
<tr>
<td>( d_M )</td>
</tr>
</tbody>
</table>

Test probabilities for CVaR: 80.20% 85.24% 90.27% 95.21%  
Test probabilities for VaR: 81.94% 81.92% 82.20% 82.03%

### Table III

<table>
<thead>
<tr>
<th>RESULTS WITH DIFFERENT CURTAILMENT RATE CAP ( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda )</td>
</tr>
<tr>
<td>1 - ( \alpha^* )</td>
</tr>
<tr>
<td>( \bar{\alpha} )</td>
</tr>
<tr>
<td>( d_M )</td>
</tr>
</tbody>
</table>

Test probabilities for CVaR: 90.26% 90.27% 90.24% 90.26%  
Test probabilities for VaR: 81.94% 81.92% 82.20% 82.03%

**REFERENCES**


