Convex Reformulation for Two-sided Distributionally Robust Chance Constraints with Inexact Moment Information

Lun Yang, Yinliang Xu, *Senior Member*, *IEEE*, Zheng Xu, and Hongbin Sun, *Fellow*, *IEEE*

Abstract—**Constraints on each node and line in power systems generally have upper and lower bounds, denoted as two-sided constraints. Most existing power system optimization methods with the distributionally robust (DR) chance-constrained program treat the two-sided DR chance constraint separately, which is an inexact approximation. This letter derives an equivalent reformulation for the generic two-sided DR chance constraint under the interval moment based ambiguity set, which does not require the exact moment information. The derived reformulation is a second-order cone program (SOCP) formulation and is then applied to the optimal power flow (OPF) problem under uncertainty. Numerical results on several IEEE systems demonstrate the effectiveness of the proposed SOCP formulation and show the differences with other DR chance-constrained OPF approaches.**

Index Terms—**Two-sided chance constraint, distributionally robust, conic reformulation, interval moment, optimal power flow.**

I. Introduction

DISTRIBUTIONALLY robust (DR) chance-constrained program is an efficient approach for decision-making in uncertain environments [1] and has received much attention in power system planning [2] and operations, especially for optimal power flow (OPF) problems [3]-[5]. There exists a major concern in most of existing works on DR chance-constrained formulations. It is commonly known that buses and transmission lines in power systems have both lower and upper bounds. However, most works (e.g., [3], [4]) manage the two-sided DR chance constraint separately, which is an inaccurate approximation and may cause unpredictable errors. To make this issue clear, taking the line power flow limits as an illustration, the line chance constraint is given as:

$$\mathbb{P}(-\bar{p}_l \leq p_l \leq \bar{p}_l) \geq 1-\varepsilon \tag{1}$$

where p_l and $\bar{p}_l$ are the power flow across line l and its capacity, respectively; ε is the allowable violation probability; and $\mathbb{P}(\cdot)$ is the probability function.

Let A_1 and A_2 denote the events of $\{p_l \leq \bar{p}_l\}$ and $\{-\bar{p}_l \leq p_l\}$, respectively. The two-sided chance constraint (1) can be expressed as:

$$\mathbb{P}(A_1 \cap A_2) \geq 1-\varepsilon \tag{2}$$

In fact, the two-sided chance constraint (1) is a joint chance constraint including two individual chance constraints. The common choice treats (2) separately, i.e., using two individual chance constraints given in (3) and (4) to approximate (2).

$$\mathbb{P}(A_1) = \mathbb{P}(p_l \leq \bar{p}_l) \geq 1-\varepsilon \tag{3}$$

$$\mathbb{P}(A_2) = \mathbb{P}(-\bar{p}_l \leq p_l) \geq 1-\varepsilon \tag{4}$$

Obviously, $\mathbb{P}(A_1 \cap A_2) < \mathbb{P}(A_1) + \mathbb{P}(A_2)$, which indicates that the common treatment mentioned above is an inexact approximation. An inspiring published work [5] treats the lower and upper bounds in the two-sided DR chance constraint simultaneously in a single-period OPF problem and provides an exact second-order cone program (SOCP) reformulation under a specific moment-based ambiguity set. Later, the derived results in [5] are extended to a multi-period optimal power-gas flow problem with two-sided DR chance constraints for an electricity-gas coupled system [6]. In [5] and [6], the moment-based ambiguity set describing the wind power forecasting errors is built on the exact mean and covariance, and the setup of mean is zero. However, although the mean and covariance could be generally estimated from the available historical data, the exact moment information is very difficult to determine in reality since the accuracy of the estimation related to the quality and quantity of historical data is unascertainable. Reference [4] considers an inexact moment-based ambiguity set, in which the mean and covariance are bounded by ellipsoid and semi-definite cone, to deal with the uncertain moment information. Then, the single-sided DR chance constraint is reformulated as a semi-definite program while the computational complexity is significantly raised in contrast to the SOCP formulation under the exact moment-based ambiguity set [4]. Based on the above literature analysis, the key question herein is how to obtain a tractable yet efficient reformulation for the generic two-sided

Manuscript received: March 10, 2021; revised: May 25, 2021; accepted: August 13, 2021. Date of CrossCheck: August 13, 2021. Date of online publication: October 1, 2021.

This work was supported by the Natural Science Foundation of Guangdong Province (No. 2021A1515012450).

L. Yang, Y. Xu (corresponding author), and Z. Xu are with the Tsinghua-Berkeley Shenzhen Institute, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China (e-mail: l-yang19@mails.tsinghua.edu.cn; xu.yinliang@sz.tsinghua.edu.cn; xu-z17@mails.tsinghua.edu.cn).

H. Sun is with the Department of Electrical Engineering, State Key Laboratory of Power Systems, Tsinghua University, Beijing 100084, China (e-mail: shb@tsinghua.edu.cn).

DOI: 10.35833/MPCE.2021.000160

DR chance constraint without the assumption of exact moment information. This letter seeks to answer this question. Compared with acquiring the exact mean and covariance from historical data, their upper and lower bounds are relatively easier to obtain [7]. Inspired by this point, this letter introduces an interval moment based ambiguity set to handle the uncertain moment information and then derives a tractable SOCP formulation for a generic two-sided DR chance constraint. The derived result only relies on the historical data of uncertainty for supporting the estimation of interval moment information, which can be directly applied to power system optimization problems involving two-sided DR chance constraints.

II. Convex Reformulation of Two-sided DR Chance Constraint

To tackle the issue of uncertain moment information, an interval moment based ambiguity set $\mathcal{P}$ defined in (5) is introduced to characterize the random variable vector $\boldsymbol{\xi}$.

$$\begin{aligned}\mathcal{P}=\Big\{&\mathbb{P}\colon \mathbb{P}(\boldsymbol{\xi}\in\mathbb{R}^K)=1,\\ &\mathbb{E}_{\mathbb{P}}(\boldsymbol{\xi})=\boldsymbol{\mu}, \mathbb{E}_{\mathbb{P}}((\boldsymbol{\xi}-\boldsymbol{\mu})(\boldsymbol{\xi}-\boldsymbol{\mu})^{\mathrm{T}})=\boldsymbol{\Sigma},\\ &\underline{\boldsymbol{\mu}}\leqslant\boldsymbol{\mu}\leqslant\bar{\boldsymbol{\mu}}, \underline{\boldsymbol{\Sigma}}\leqslant\boldsymbol{\Sigma}\leqslant\overline{\boldsymbol{\Sigma}}\Big\}\end{aligned} \tag{5}$$

where the first line describes that $\boldsymbol{\xi}$ is constrained within a support set $\mathbb{R}^K$, and K is the dimension of $\boldsymbol{\xi}$; the second line describes the moment information of $\boldsymbol{\xi}$, and $\mathbb{E}_{\mathbb{P}}(\boldsymbol{\xi})$ is the expectation function; and the third line suggests that the mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ lie in a box region specified by upper and lower bounds, and $\overline{(\cdot)}$ and $\underline{(\cdot)}$ denote the upper and lower bounds, respectively.

In accordance with two-sided DR chance constraints in power system optimization models, for ease of illustration, a generic two-sided DR chance constraint is defined as:

$$\inf_{\mathbb{P}\in\mathcal{P}}\mathbb{P}_{\xi}\big(l(\boldsymbol{x})\leq\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\xi}+b(\boldsymbol{x})\leq u(\boldsymbol{x})\big)\geq 1-\varepsilon \tag{6}$$

where $l(\boldsymbol{x})$, $\boldsymbol{a}(\boldsymbol{x})$, $b(\boldsymbol{x})$, and $u(\boldsymbol{x})$ are all affine mappings in $\boldsymbol{x}$.

Theorem 1: supposing the ambiguity set is defined in (5), the generic two-sided DR chance constraint (6) is equivalent to the following SOCP:

$$\min_{y,z,\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_3,\boldsymbol{\lambda}_4}\ \{y^2+\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\overline{\boldsymbol{\Sigma}}\boldsymbol{a}(\boldsymbol{x})\leq\varepsilon(T_1(\boldsymbol{x})-z)^2\} \tag{7}$$

s.t.

$$\boldsymbol{\lambda}_1\bar{\boldsymbol{\mu}}-\boldsymbol{\lambda}_2\underline{\boldsymbol{\mu}}+(b(\boldsymbol{x})-T_2(\boldsymbol{x}))\leq y+z \tag{8}$$

$$\boldsymbol{\lambda}_3\bar{\boldsymbol{\mu}}-\boldsymbol{\lambda}_4\underline{\boldsymbol{\mu}}-(b(\boldsymbol{x})-T_2(\boldsymbol{x}))\leq y+z \tag{9}$$

$$\begin{cases}\boldsymbol{a}(\boldsymbol{x})=\boldsymbol{\lambda}_1-\boldsymbol{\lambda}_2\\ -\boldsymbol{a}(\boldsymbol{x})=\boldsymbol{\lambda}_3-\boldsymbol{\lambda}_4\end{cases} \tag{10}$$

$$\begin{cases}0\leq z\leq T_1(\boldsymbol{x})\\ y\geq 0\\ \boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_3,\boldsymbol{\lambda}_4\geq\boldsymbol{0}\end{cases} \tag{11}$$

Proof: see Appendix A.

III. Case Study

A. Simulation Setup

Consider a power system where the sets of buses, lines, generators, and wind farms are denoted as $\mathcal{B}$, $\mathcal{L}$, $\mathcal{G}$, and $\mathcal{W}$, respectively. Each bus $i\in\mathcal{B}$ has load d_i. For each $i\in\mathcal{W}$, the uncertain wind power is modeled by $p_i+\xi_i$, where p_i is the forecasting value, and ξ_i is the uncertain forecasting error. To compensate for the total forecasting deviations of wind power, each generator adjusts its output using the affine policy similar to [4]-[6]. The actual generation output is modeled by $g_i-\alpha_i(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{\xi})$, where g_i is the generation output without consideration of forecasting deviations; the term $-\alpha_i(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{\xi})$ represents the activated reserve of generator i for uncertainty mitigation; α_i is the participation factor; and $\boldsymbol{e}$ denotes a vector of all ones. Following [5], a DR chance-constrained OPF problem is formulated as:

$$\min\Big\{\sum_{i\in\mathcal{G}}c_{1,i}g_i^2+c_{2,i}g_i+c_i^{\mathrm{R}}(r_i^{up}+r_i^{dn})\Big\} \tag{12}$$

$$\sum_{i\in\mathcal{G}}g_i+\sum_{i\in\mathcal{W}}p_i-\sum_{i\in\mathcal{B}}d_i=0 \tag{13}$$

$$\sum_{i\in\mathcal{G}}\alpha_i=1 \quad \alpha_i\geq 0 \tag{14}$$

$$\inf_{\mathbb{P}\in\mathcal{P}}\mathbb{P}\Big(\underline{g}_i\leq g_i-\alpha_i(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{\xi})\leq\bar{g}_i\Big)\geq 1-\varepsilon \quad \forall i\in\mathcal{G} \tag{15}$$

$$\inf_{\mathbb{P}\in\mathcal{P}}\mathbb{P}\Big(-r_i^{dn}\leq-\alpha_i(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{\xi})\leq r_i^{up}\Big)\geq 1-\varepsilon \quad \forall i\in\mathcal{G} \tag{16}$$

$$\inf_{\mathbb{P}\in\mathcal{P}}\mathbb{P}(-p_l\leq\boldsymbol{\Psi}_{\mathrm{G},l}(\boldsymbol{g}-\boldsymbol{\alpha}(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{\xi}))+\boldsymbol{\Psi}_{\mathrm{W},l}(\boldsymbol{p}+\boldsymbol{\xi})-\boldsymbol{\Psi}_{\mathrm{B},l}\boldsymbol{d}\leq p_l)\geq 1-\varepsilon \quad \forall l\in\mathcal{L} \tag{17}$$

The objective function (12) minimizes the total operation cost including generation cost and reserve cost, where $c_{1,i}$ and $c_{2,i}$ are the generation cost coefficients, and c_i^{R} is the reserve cost coefficient. Constraints (13) and (14) enforce power balance without and with wind power forecasting errors, respectively. Constraint (15) imposes the generation output within its limits $[\underline{g}_i,\bar{g}_i]$. Constraint (16) restricts the activated reserve $-\alpha_i(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{\xi})$ by upward and downward reserve capacities r_i^{up} and r_i^{dn}. Constraint (17) restricts the line power flow within its capacity limit p_l, where $\boldsymbol{\Psi}_{\mathrm{G},l}$, $\boldsymbol{\Psi}_{\mathrm{W},l}$, and $\boldsymbol{\Psi}_{\mathrm{B},l}$ are power transfer distribution factor vectors mapping generators, wind farms, and loads to line l, respectively. Constraints (15)-(17) are two-sided DR chance constraints. With the derived result (7)-(11), problem (12)-(17) can be reformulated as an SOCP presented in Appendix B.

We test our approach on several IEEE systems whose data are obtained from MATPOWER 3.1 [8]. The following three methods are compared. ① M1: DR chance-constrained OPF method in [4], which treats two-sided DR chance constraints separately under the exact moment-based ambiguity set. ② M2: DR chance-constrained OPF method in [5], which treats two-sided DR chance constraints simultaneously under the exact moment-based ambiguity set. ③ M3: the proposed DR chance-constrained OPF method based on the derived result (7)-(11) under the interval moment-based ambiguity set described in Section II. All the models are solved by Gurobi 9.1.0.

To ensure the reproducibility of all case studies, we herein consider a unified setup on IEEE test systems. We assume that each generator bus connects a wind farm. The forecasting value of wind power is set to be 10% of the capacity of the local generator. For the moment-based ambiguity set in M1 and M2, the mean is set to be 0 in consistent with that in [5], and the covariance matrix is diagonal with the diagonal element (i.e., variance) being 625. For the interval moment-based ambiguity set in M3, the upper and lower bounds of the mean are set to ±5 MW deviating from 0, and the upper and lower bounds on the diagonal element of covariance matrix are set to deviate ±5% of that in M1 and M2.

B. Simulation Results

1) Reliability Comparisons Among M1, M2, and M3

We generate 50000 samples from uncertain wind power forecasting errors obeying Gaussian distributions with mean and covariance given above and check the joint violation probability of solutions, which is defined as the percentage of samples for which any chance constraint is violated. In Fig. 1, the joint violation probability of solutions in M3 is evidently less than those in M1 and M2 under all tested risk parameters in the IEEE 39-bus system. Furthermore, with the growth of risk parameter, joint violation probabilities in M1 and M2 increase notably while that in M3 rises moderately. Even when the risk parameter is 0.3, the joint violation probability in M3 is just 0.05. These observations show that the proposed method (M3) holds a higher solution reliability and is robust to the risk parameter. As observed in Fig. 2, the total operation costs provided in M1-M3 decrease with the increase of the risk parameter. Therefore, the conservativeness of these methods can be adjusted by the risk parameter. Another observation is that the total operation cost in M3 is higher than those in M1 and M2, which indicates that M3 is more conservative. This is because the ambiguity set in M3 just uses the interval moment information while those in M1 and M2 require the exact moment information.

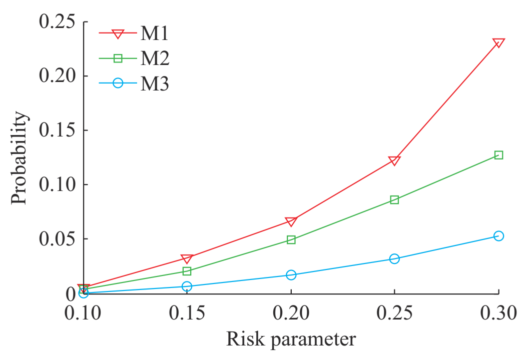

Fig. 1. Joint violation probabilities in M1, M2, and M3 in IEEE 39-bus system.

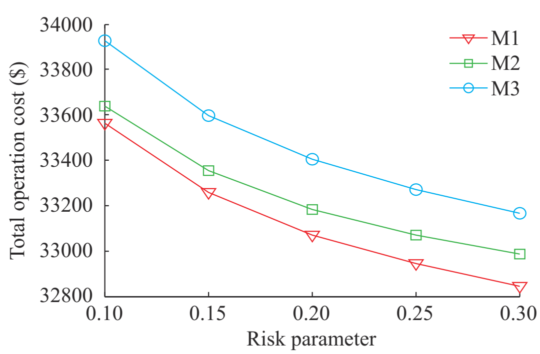

Fig. 2. Total operation costs in M1, M2, and M3 in IEEE 39-bus system.

2) Influence of Different Interval Sizes of Moments

The key feature of the proposed interval moment-based ambiguity set is the inclusion of uncertain moments described by the interval. Hence, this part investigates the influence of different interval sizes on the solutions in the IEEE 39-bus system. The interval sizes of mean and covariance in M3 are enlarged gradually as follows: ① the upper and lower bounds of the mean are set to $\pm\alpha_\mu$ MW deviating from 0; ② the upper and lower bounds of diagonal element of covariance are set to $\pm\alpha_\Sigma$% deviating from those in M1 and M2, where α_μ and α_Σ increase from 1 to 5 with step size 1. The risk parameter is set to be 0.1 in these tests. The total operation cost in M3 under different interval sizes is depicted in Fig. 3. It is clear that the total operation cost increases when the interval sizes expand. Besides, the interval size of mean poses a higher impact on the total operation cost than that of covariance. Therefore, in practice, system operators can adjust the total operation cost by changing the interval sizes of moments according to the obtained interval estimated from the historical data.

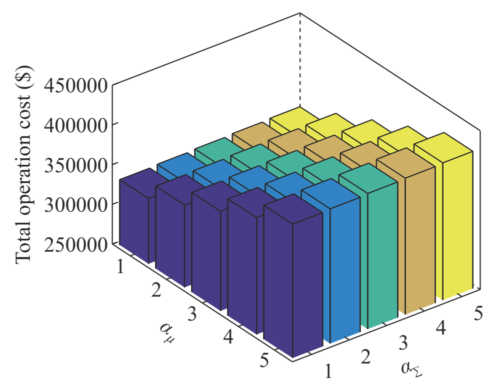

Fig. 3. Total operation costs under different interval sizes of moments in IEEE 39-bus system.

3) Computation Time

This part compares the computation time of M1, M2, and M3 by solving different MATPOWER cases, i.e., case9, case24, case39, and case118 corresponding to IEEE 9-bus, 24-bus, 39-bus, and 118-bus systems, respectively, on a computer with Inter Core i5 2.5 GHz CPU and 24 GB memory. As shown in Table I, all methods can be solved efficiently within 30 s even for the large-scale IEEE 118-bus system (case118). Another observation is that the computation time of M3 is more than that of M1 and M2 since M3 requires at most $(4|\mathcal{L}|+8|\mathcal{G}|)|\mathcal{W}|$ more variables than M2 and $(6|\mathcal{L}|+12|\mathcal{G}|)|\mathcal{W}|$ more than M1, but M3 still holds a decent computational efficiency.

TABLE I
COMPUTATION TIME OF M1, M2, AND M3

Case	Data				Time (s)		
	$\lvert\mathcal{B}\rvert$	$\lvert\mathcal{L}\rvert$	$\lvert\mathcal{G}\rvert$	$\lvert\mathcal{W}\rvert$	M1	M2	M3
case9	9	9	3	3	0.297	0.453	0.360
case24	24	38	33	33	0.529	0.798	0.929
case39	39	46	10	10	0.366	0.351	0.479
case118	118	186	54	54	19.291	13.719	25.717

IV. Conclusion

This letter derives a tractable SOCP formulation for the generic two-sided DR chance constraint with interval moment information and then applies this result to a DR chance-constrained OPF problem. The derived formulation does not rely on the assumption of exact moment information. Numerical results show that the proposed SOCP formulation can be solved efficiently and the obtained solutions hold a higher reliability with a lower violation probability and are more robust to the risk parameter compared with the existing methods on two-sided DR chance constraints, i.e., the inexact approximation by two single-sided DR chance constraints and the conic reformulation under the ambiguity set built on exact moment information.

Appendix A

Proof*:* it is clear that (6) can be transformed to a generic symmetrical two-sided DR chance constraint:

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}_{\xi}(-T_1(\boldsymbol{x})\le \boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\xi}+b(\boldsymbol{x})-T_2(\boldsymbol{x})\le T_1(\boldsymbol{x}))\ge 1-\varepsilon \tag{A1}$$

where $T_1(\boldsymbol{x})=(u(\boldsymbol{x})-l(\boldsymbol{x}))/2$; and $T_2(\boldsymbol{x})=(u(\boldsymbol{x})+l(\boldsymbol{x}))/2$. Constraint (A1) can be then reformulated as:

$$\sup_{(\boldsymbol{\mu},\boldsymbol{\Sigma})\in\mathcal{U}}\inf_{\mathbb{P}\in\mathcal{P}_1} \mathbb{P}_{\xi}(-T_1(\boldsymbol{x})\le \boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\xi}+b(\boldsymbol{x})-T_2(\boldsymbol{x})\le T_1(\boldsymbol{x}))\ge 1-\varepsilon \tag{A2}$$

$$\mathcal{P}_1=\{\mathbb{P}: \mathbb{P}(\boldsymbol{\xi}\in\mathbb{R}^K)=1, \mathbb{E}_{\mathbb{P}}(\boldsymbol{\xi})=\boldsymbol{\mu}, \mathbb{E}_{\mathbb{P}}((\boldsymbol{\xi}-\boldsymbol{\mu})(\boldsymbol{\xi}-\boldsymbol{\mu})^{\mathrm{T}})=\boldsymbol{\Sigma}\} \tag{A3}$$

$$\mathcal{U}=\left\{(\boldsymbol{\mu},\boldsymbol{\Sigma}): \underline{\boldsymbol{\mu}}\le\boldsymbol{\mu}\le\overline{\boldsymbol{\mu}}\ ;\ \underline{\boldsymbol{\Sigma}}\le\boldsymbol{\Sigma}\le\overline{\boldsymbol{\Sigma}}\right\} \tag{A4}$$

Now in (A2), we first conduct the reformulation derivation for the inner infimum problem under $\mathcal{P}_1$ and then outer supremum problem under $\mathcal{U}$.

Let $\omega=\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\xi}+b(\boldsymbol{x})-T_2(\boldsymbol{x})$, the inner infimum problem in the left-hand side of (A2) is equivalent to $\inf_{\mathbb{P}\in\mathcal{P}_2}\mathbb{P}_{\omega}(|\omega|\le T_1(\boldsymbol{x}))$, where $\mathcal{P}_2$ is expressed as:

$$\begin{aligned}\mathcal{P}_2=\{\mathbb{P}: \mathbb{P}(\omega\in\mathbb{R})=1, \mathbb{E}_{\mathbb{P}}(\omega)=\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}),\\ \mathbb{E}_{\mathbb{P}}((\omega-\mathbb{E}_{\mathbb{P}}(\omega))(\omega-\mathbb{E}_{\mathbb{P}}(\omega))^{\mathrm{T}})=\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\}\end{aligned} \tag{A5}$$

The $\inf_{\mathbb{P}\in\mathcal{P}_2}\mathbb{P}_{\omega}(|\omega|\le T_1(\boldsymbol{x}))$ is equivalently unfolded as:

$$\min_{f(\omega)}\int_{\mathbb{R}}\mathbb{I}_{\mathcal{A}}(\omega)f(\omega)\mathrm{d}\omega \tag{A6}$$

s.t.

$$\int_{\mathbb{R}}f(\omega)\mathrm{d}\omega=1 \tag{A7}$$

$$\int_{\mathbb{R}}\omega f(\omega)\mathrm{d}\omega=\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}) \tag{A8}$$

$$\int_{\mathbb{R}}\omega^2 f(\omega)\mathrm{d}\omega=\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})+(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))^2 \tag{A9}$$

where $\mathbb{I}_{\mathcal{A}}(\omega)$ is an indicator function, which is 1 if $\omega\in\mathcal{A}$, $\mathcal{A}=\{\omega: |\omega|\le T_1(\boldsymbol{x})\}$ and 0 otherwise. By conic duality [5], problem (A6)-(A9) can be reformulated as:

$$\begin{aligned}\max_{\lambda,\gamma,\delta}\ \{\lambda+(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\mu+b(\boldsymbol{x})-T_2(\boldsymbol{x}))\gamma+\\ [\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})+(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\mu+b(\boldsymbol{x})-T_2(\boldsymbol{x}))^2]\delta\}\end{aligned} \tag{A10}$$

s.t.

$$\lambda+\omega\gamma+\omega^2\delta\le 1\quad \forall\omega\in\mathbb{R} \tag{A11}$$

$$\lambda+\omega\gamma+\omega^2\delta\le 0\quad \forall\omega\ge T_1(\boldsymbol{x}) \tag{A12}$$

$$\lambda+\omega\gamma+\omega^2\delta\le 0\quad \forall\omega\le -T_1(\boldsymbol{x}) \tag{A13}$$

where λ, γ, and δ are the dual variables for constraints (A7)-(A9). Note that the feasible region of (A10)-(A13) is nonempty when $\delta\le 0$ and $|\gamma/(2\beta)|\le T_1(\boldsymbol{x})$. Otherwise, $\sup_{\omega}(\lambda+\omega\gamma+\omega^2\delta)$ holds and leads to $\mathbb{E}_{\mathbb{P}}(\lambda+\omega\gamma+\omega^2\delta)=\lambda+(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))\gamma+(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})+(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))^2)\delta\le 0$, which contradicts $\inf_{\mathbb{P}\in\mathcal{P}_2}\mathbb{P}_{\omega}(|\omega|\le T_1(\boldsymbol{x}))\ge 1-\varepsilon>0$.

Owing to $\delta\le 0$ and $|\gamma/(2\delta)|\le T_1(\boldsymbol{x})$, problem (A10)-(A13) is equivalent to:

$$\begin{aligned}\max_{\lambda,\gamma,\delta}\ \{\lambda+(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))\gamma+\\ [\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})+(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))^2]\delta\}\end{aligned} \tag{A14}$$

s.t.

$$\lambda-\gamma^2/(4\delta)\le 1 \tag{A15}$$

$$\lambda+|T_1(\boldsymbol{x})|\gamma+T_1^2(\boldsymbol{x})\delta\le 0\quad \delta\le 0 \tag{A16}$$

It is clear that, to maximize the objective function in (A14), the optimal γ must have the same sign as $\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})$. Thus, $\inf_{\mathbb{P}\in\mathcal{P}_2}\mathbb{P}_{\omega}(|\omega|\le T_1(\boldsymbol{x}))\ge 1-\varepsilon$ is equivalent to:

$$\begin{aligned}\max_{\lambda,\gamma,\delta}\ \{\lambda+|(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))||\gamma|+\\ (\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})+(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))^2)\delta\}\ge 1-\varepsilon\end{aligned} \tag{A17}$$

s.t.

$$\lambda-\gamma^2/(4\delta)\le 1 \tag{A18}$$

$$\lambda+|T_1(\boldsymbol{x})|\gamma+T_1(\boldsymbol{x})^2\delta\le 0\quad \delta\le 0 \tag{A19}$$

Let $\theta=-1/\delta, \hat{\gamma}=-|\gamma|/\delta, \hat{\lambda}=-\lambda/\delta$, problem (A17)-(A19) is equivalent to:

$$\begin{aligned}\max_{\hat{\lambda},\hat{\gamma},\theta}\ \{\hat{\lambda}+(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))|\hat{\gamma}|-\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})-\\ (\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))^2\}\ge(1-\varepsilon)\theta\end{aligned} \tag{A20}$$

s.t.

$$\hat{\lambda}+\hat{\gamma}^2/4\le\theta \tag{A21}$$

$$\hat{\lambda}+|T_1(\boldsymbol{x})|\hat{\gamma}-T_1^2(\boldsymbol{x})\le 0\quad \hat{\gamma}\le 0 \tag{A22}$$

Using Fourier-Motzkin procedure to eliminate $\hat{\lambda}$ and θ in (A20)-(A22), we have:

$$\begin{cases}\min_{\hat{\gamma}}\ \{(|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|-\hat{\gamma}/2)^2+\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\}\le\varepsilon(T_1(\boldsymbol{x})-\hat{\gamma}/2)^2\\ \text{s.t.}\ \ \hat{\gamma}\ge 0\end{cases} \tag{A23}$$

Problem (A23) can be unfolded as:

$$\begin{cases}\min_{\hat{\gamma}}\left\{\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})-(|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|-\varepsilon T_1(\boldsymbol{x}))\hat{\gamma}+\dfrac{1-\varepsilon}{4}\hat{\gamma}^2+\right.\\ \left.\qquad(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))^2\right\}\le\varepsilon T_1(\boldsymbol{x})^2\\ \text{s.t.}\ \ \hat{\gamma}\ge 0\end{cases} \tag{A24}$$

The optimal solution $\hat{\gamma}$ in (A24) is distinguished by the following two cases.

1) Case 1: if $|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|\le\varepsilon T_1(\boldsymbol{x})$, then $\hat{\gamma}^*=0$. Problem (A24) can be reformulated as:

$$(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))^2+\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\leq\varepsilon T_1(\boldsymbol{x})^2 \tag{A25}$$

$$|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|\leq\varepsilon T_1(\boldsymbol{x}) \tag{A26}$$

2) Case 2: if $|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|\geq\varepsilon T_1(\boldsymbol{x})$, then $\hat{\gamma}^*=2(|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|-\varepsilon T_1(\boldsymbol{x}))/(1-\varepsilon)$. Problem (A24) can be reformulated as:

$$\frac{1-\varepsilon}{\varepsilon}\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\leq(T_1(\boldsymbol{x})-|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|)^2 \tag{A27}$$

$$|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|\geq\varepsilon T_1(\boldsymbol{x}) \tag{A28}$$

It follows that $X=X_1\cup X_2$, where X, X_1, and X_2 denote the sets of $\boldsymbol{x}$ defined by (A23), (A25) and (A26), and (A27) and (A28), respectively.

From optimizing $\hat{\gamma}$ in the above two cases (Case 1 and Case 2), it can be observed that the optimal $\hat{\gamma}$ in (A23) must be less than $|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|$. Otherwise, the derived formulation from (A23) will yield a smaller restriction on $\boldsymbol{x}$. Let $z=\hat{\gamma}/2$, problem (A23) is thus equivalent to:

$$\min_z\ \{(|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|-z)^2+\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\}\leq\varepsilon(T_1(\boldsymbol{x})-z)^2 \tag{A29}$$

s.t.

$$0\leq z\leq|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})| \tag{A30}$$

Now we claim that the problem described in (A29) and (A30) is equivalent to:

$$\min_{z,y}\ \{y^2+\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\}\leq\varepsilon(T_1(\boldsymbol{x})-z)^2 \tag{A31}$$

s.t.

$$|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|\leq y+z \tag{A32}$$

$$\begin{cases}0\leq z\leq T_1(\boldsymbol{x})\\ y\geq 0\end{cases} \tag{A33}$$

The claim is proven as below. Let X' denote the set of $\boldsymbol{x}$ defined by (A31)-(A33).

1) First, we show $X\subseteq X'$. Given $\boldsymbol{x}\in X$, there exists a z such that $(\boldsymbol{x},z)$ meets (A29) and (A30). Let $y=|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|-z$, it is clear that $(\boldsymbol{x},z,y)$ satisfies (A31)-(A33). Thus, $\boldsymbol{x}\in X'$, which implies $X\subseteq X'$.

2) Then, we show $X'\subseteq X$. Given $\boldsymbol{x}\in X'$, there exists a (z,y) such that $(\boldsymbol{x},z,y)$ meets (A31)-(A33). Two cases are discussed as follows.

① If $0\leq z\leq|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|$, then we have $y\geq|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|-z\geq 0$. Together with (A31), we further have $(|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|-z)^2+\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\leq y^2+\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\leq\varepsilon(T_1(\boldsymbol{x})-z)^2$. Thus, $(\boldsymbol{x},z)$ meets (A29) and (A30), which implies $\boldsymbol{x}\in X$.

② If $|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|\leq z\leq T_1(\boldsymbol{x})$, together with (A31), we have:

$$\begin{aligned}\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})&\leq y^2+\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\leq\varepsilon(T_1(\boldsymbol{x})-z)^2\leq\\ &\varepsilon(T_1(\boldsymbol{x})-|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|)^2\end{aligned} \tag{A34}$$

Now in (A34), we analyze the following two situations:

Situation 1: if $|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|\leq z\leq T_1(\boldsymbol{x})$, we have

$$\begin{aligned}\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})&\leq\varepsilon(T_1(\boldsymbol{x})-|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|)^2\leq\\ &\varepsilon T_1(\boldsymbol{x})^2-(2-\varepsilon)(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))^2\leq\\ &\varepsilon T_1(\boldsymbol{x})^2-(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))^2\end{aligned} \tag{A35}$$

where the first inequality holds due to (A34), the second inequality holds due to $|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|\leq\varepsilon T_1(\boldsymbol{x})$, and the third inequality holds due to $\varepsilon\in(0,1)$. It follows that $x\in X_1\subset X$.

Situation 2: if $|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|\geq\varepsilon T_1(\boldsymbol{x})$, (A34) implies that $[(1-\varepsilon)/\varepsilon]\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\leq(T_1(\boldsymbol{x})-|\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})|)^2$ since $(1-\varepsilon)/\varepsilon\leq 1/\varepsilon$. Thus, it follows that $\boldsymbol{x}\in X_2\subset X$.

Summarizing the analysis in the above situations, we have $X'\subseteq X$. Thus $X'=X$, i.e., the problem described in (A29) and (A30) is equivalent to (A31)-(A33). The proof of the claim is completed.

Recalling problem (A2), we know problem (A31)-(A33) under uncertainty set $\mathcal{U}$ and the robust counterpart in problem (A2) can be reformulated as:

$$\min_{x,y,z}\left\{y^2+\max_{\boldsymbol{\Sigma}}\ \boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})\right\}\leq\varepsilon(T_1(\boldsymbol{x})-z)^2\quad\forall\boldsymbol{\Sigma}:\underline{\boldsymbol{\Sigma}}\preccurlyeq\boldsymbol{\Sigma}\preccurlyeq\overline{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}\succcurlyeq\boldsymbol{0} \tag{A36}$$

s.t.

$$\max_{\boldsymbol{\mu}}\ \{\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x})\}\leq y+z\quad\forall\boldsymbol{\mu}:\underline{\boldsymbol{\mu}}\leq\boldsymbol{\mu}\leq\overline{\boldsymbol{\mu}} \tag{A37}$$

$$\max_{\boldsymbol{\mu}}\ \{-(\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\mu}+b(\boldsymbol{x})-T_2(\boldsymbol{x}))\}\leq y+z\quad\forall\boldsymbol{\mu}:\underline{\boldsymbol{\mu}}\leq\boldsymbol{\mu}\leq\overline{\boldsymbol{\mu}} \tag{A38}$$

$$\begin{cases}0\leq z\leq T_1(\boldsymbol{x})\\ y\geq 0\end{cases} \tag{A39}$$

Note that in (A36), since $\boldsymbol{\Sigma}\preccurlyeq\overline{\boldsymbol{\Sigma}}$, $\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}(\overline{\boldsymbol{\Sigma}}-\boldsymbol{\Sigma})\boldsymbol{a}(\boldsymbol{x})\geq 0$ holds. Thus, we have $\max\limits_{\underline{\boldsymbol{\Sigma}}\preccurlyeq\boldsymbol{\Sigma}\preccurlyeq\overline{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}\succcurlyeq\boldsymbol{0}}\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})=\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\overline{\boldsymbol{\Sigma}}\boldsymbol{a}(\boldsymbol{x})$. For the max form in the left-hand side of (A37) and (A38) with respect to $\boldsymbol{\mu}$, their dual forms are as follows: $\max\limits_{\underline{\boldsymbol{\Sigma}}\preccurlyeq\boldsymbol{\Sigma}\preccurlyeq\overline{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}\succcurlyeq\boldsymbol{0}}\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a}(\boldsymbol{x})=\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\overline{\boldsymbol{\Sigma}}\boldsymbol{a}(\boldsymbol{x})$, s.t. $\boldsymbol{a}(\boldsymbol{x})=\boldsymbol{\lambda}_1-\boldsymbol{\lambda}_2$; $\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2\geq\boldsymbol{0}$, and $\min\limits_{\boldsymbol{\lambda}_3,\boldsymbol{\lambda}_4}\boldsymbol{\lambda}_3\overline{\boldsymbol{\mu}}-\boldsymbol{\lambda}_4\underline{\boldsymbol{\mu}}$, s.t. $-\boldsymbol{a}(\boldsymbol{x})=\boldsymbol{\lambda}_3-\boldsymbol{\lambda}_4$; $\boldsymbol{\lambda}_3,\boldsymbol{\lambda}_4\geq\boldsymbol{0}$. Thus, problem (A36)-(A39) can be equivalently reformulated as:

$$\min_{y,z,\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_3,\boldsymbol{\lambda}_4}\ \{y^2+\boldsymbol{a}(\boldsymbol{x})^{\mathrm{T}}\overline{\boldsymbol{\Sigma}}\boldsymbol{a}(\boldsymbol{x})\}\leq\varepsilon(T_1(\boldsymbol{x})-z)^2 \tag{A40}$$

s.t.

$$\boldsymbol{\lambda}_1\overline{\boldsymbol{\mu}}-\boldsymbol{\lambda}_2\underline{\boldsymbol{\mu}}+(b(\boldsymbol{x})-T_2(\boldsymbol{x}))\leq y+z \tag{A41}$$

$$\boldsymbol{\lambda}_3\overline{\boldsymbol{\mu}}-\boldsymbol{\lambda}_4\underline{\boldsymbol{\mu}}-(b(\boldsymbol{x})-T_2(\boldsymbol{x}))\leq y+z \tag{A42}$$

$$\begin{cases}\boldsymbol{a}(\boldsymbol{x})=\boldsymbol{\lambda}_1-\boldsymbol{\lambda}_2\\ -\boldsymbol{a}(\boldsymbol{x})=\boldsymbol{\lambda}_3-\boldsymbol{\lambda}_4\end{cases} \tag{A43}$$

$$\begin{cases}0\leq z\leq T_1(\boldsymbol{x})\\ y\geq 0\\ \boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_3,\boldsymbol{\lambda}_4\geq\boldsymbol{0}\end{cases} \tag{A44}$$

Appendix B

The SOCP reformulation of the DR chance-constrained OPF problem (12)-(17) is presented as:

$$\min\left\{\sum_{i\in\mathcal{G}}c_{1,i}g_i^2+c_{2,i}g_i+c_i^{\mathrm{R}}(r_i^{up}+r_i^{dn})\right\} \tag{B1}$$

$$\sum_{i\in\mathcal{G}}g_i+\sum_{i\in\mathcal{W}}p_i-\sum_{i\in\mathcal{B}}d_i=0 \tag{B2}$$

$$\sum_{i\in\mathcal{G}}\alpha_i=1\quad\alpha_i\geq 0 \tag{B3}$$

$$y_{g,i}^2 + (-\alpha_i \boldsymbol{e}^{\mathrm{T}})^{\mathrm{T}} \overline{\boldsymbol{\Sigma}} (-\alpha_i \boldsymbol{e}^{\mathrm{T}}) \le \varepsilon \left(\frac{\bar{g}_i - \underline{g}_i}{2} - z_{g,i} \right)^2 \quad \forall i \in \mathcal{G} \tag{B4}$$

$$\boldsymbol{\lambda}_{1,g,i} \bar{\boldsymbol{\mu}} - \boldsymbol{\lambda}_{2,g,i} \underline{\boldsymbol{\mu}} + \left(g_i - \frac{\bar{g}_i + \underline{g}_i}{2} \right) \le y_{g,i} + z_{g,i} \quad \forall i \in \mathcal{G} \tag{B5}$$

$$\boldsymbol{\lambda}_{3,g,i} \bar{\boldsymbol{\mu}} - \boldsymbol{\lambda}_{4,g,i} \underline{\boldsymbol{\mu}} - \left(g_i - \frac{\bar{g}_i + \underline{g}_i}{2} \right) \le y_{g,i} + z_{g,i} \quad \forall i \in \mathcal{G} \tag{B6}$$

$$-\alpha_i \boldsymbol{e}^{\mathrm{T}} = \boldsymbol{\lambda}_{1,g,i} - \boldsymbol{\lambda}_{2,g,i} \quad \forall i \in \mathcal{G} \tag{B7}$$

$$-(-\alpha_i \boldsymbol{e}^{\mathrm{T}}) = \boldsymbol{\lambda}_{3,g,i} - \boldsymbol{\lambda}_{4,g,i} \quad \forall i \in \mathcal{G} \tag{B8}$$

$$0 \le z_{g,i} \le \frac{\bar{g}_i - \underline{g}_i}{2} \quad \forall i \in \mathcal{G} \tag{B9}$$

$$\begin{cases} y_{g,i} \ge 0 \\ \boldsymbol{\lambda}_{1,g,i}, \boldsymbol{\lambda}_{2,g,i}, \boldsymbol{\lambda}_{3,g,i}, \boldsymbol{\lambda}_{4,g,i} \ge \boldsymbol{0} \end{cases} \quad \forall i \in \mathcal{G} \tag{B10}$$

$$y_{r,i}^2 + (-\alpha_i \boldsymbol{e}^{\mathrm{T}})^{\mathrm{T}} \overline{\boldsymbol{\Sigma}} (-\alpha_i \boldsymbol{e}^{\mathrm{T}}) \le \varepsilon \left(\frac{r_i^{up} + r_i^{dn}}{2} - z_{r,i} \right)^2 \quad \forall i \in \mathcal{G} \tag{B11}$$

$$\boldsymbol{\lambda}_{1,r,i} \bar{\boldsymbol{\mu}} - \boldsymbol{\lambda}_{2,r,i} \underline{\boldsymbol{\mu}} + \left(- \frac{r_i^{up} - r_i^{dn}}{2} \right) \le y_{r,i} + z_{r,i} \quad \forall i \in \mathcal{G} \tag{B12}$$

$$\boldsymbol{\lambda}_{3,r,i} \bar{\boldsymbol{\mu}} - \boldsymbol{\lambda}_{4,r,i} \underline{\boldsymbol{\mu}} - \left(- \frac{r_i^{up} - r_i^{dn}}{2} \right) \le y_{r,i} + z_{r,i} \quad \forall i \in \mathcal{G} \tag{B13}$$

$$(-\alpha_i \boldsymbol{e}^{\mathrm{T}}) = \boldsymbol{\lambda}_{1,r,i} - \boldsymbol{\lambda}_{2,r,i} \quad \forall i \in \mathcal{G} \tag{B14}$$

$$-(-\alpha_i \boldsymbol{e}^{\mathrm{T}}) = \boldsymbol{\lambda}_{3,r,i} - \boldsymbol{\lambda}_{4,r,i} \quad \forall i \in \mathcal{G} \tag{B15}$$

$$0 \le z_{r,i} \le \frac{r_i^{up} + r_i^{dn}}{2} \quad \forall i \in \mathcal{G} \tag{B16}$$

$$\begin{cases} y_{r,i} \ge 0 \\ \boldsymbol{\lambda}_{1,r,i}, \boldsymbol{\lambda}_{2,r,i}, \boldsymbol{\lambda}_{3,r,i}, \boldsymbol{\lambda}_{4,r,i} \ge \boldsymbol{0} \end{cases} \quad \forall i \in \mathcal{G} \tag{B17}$$

$$y_{l,i}^2 + (\boldsymbol{\Psi}_{\mathrm{G},l}(-\boldsymbol{\alpha e}^{\mathrm{T}}) + \boldsymbol{\Psi}_{\mathrm{W},l})^{\mathrm{T}} \overline{\boldsymbol{\Sigma}} (\boldsymbol{\Psi}_{\mathrm{G},l}(-\boldsymbol{\alpha e}^{\mathrm{T}}) + \boldsymbol{\Psi}_{\mathrm{W},l}) \le \varepsilon (p_l - z_{l,i})^2 \quad \forall l \in \mathcal{L} \tag{B18}$$

$$\boldsymbol{\lambda}_{1,l,i} \bar{\boldsymbol{\mu}} - \boldsymbol{\lambda}_{2,l,i} \underline{\boldsymbol{\mu}} + (\boldsymbol{\Psi}_{\mathrm{G},l} \boldsymbol{g} + \boldsymbol{\Psi}_{\mathrm{P},l} \boldsymbol{p} - \boldsymbol{\Psi}_{\mathrm{B},l} \boldsymbol{d}) \le y_{l,i} + z_{l,i} \quad \forall l \in \mathcal{L} \tag{B19}$$

$$\boldsymbol{\lambda}_{3,l,i} \bar{\boldsymbol{\mu}} - \boldsymbol{\lambda}_{4,l,i} \underline{\boldsymbol{\mu}} - (\boldsymbol{\Psi}_{\mathrm{G},l} \boldsymbol{g} + \boldsymbol{\Psi}_{\mathrm{P},l} \boldsymbol{p} - \boldsymbol{\Psi}_{\mathrm{B},l} \boldsymbol{d}) \le y_{l,i} + z_{l,i} \quad \forall l \in \mathcal{L} \tag{B20}$$

$$(\boldsymbol{\Psi}_{\mathrm{G},l}(-\boldsymbol{\alpha e}^{\mathrm{T}}) + \boldsymbol{\Psi}_{\mathrm{W},l}) = \boldsymbol{\lambda}_{1,l,i} - \boldsymbol{\lambda}_{2,l,i} \quad \forall l \in \mathcal{L} \tag{B21}$$

$$-(\boldsymbol{\Psi}_{\mathrm{G},l}(-\boldsymbol{\alpha e}^{\mathrm{T}}) + \boldsymbol{\Psi}_{\mathrm{W},l}) = \boldsymbol{\lambda}_{3,l,i} - \boldsymbol{\lambda}_{4,l,i} \quad \forall l \in \mathcal{L} \tag{B22}$$

$$0 \le z_{l,i} \le p_l \quad \forall l \in \mathcal{L} \tag{B23}$$

$$\begin{cases} y_{l,i} \ge 0 \\ \boldsymbol{\lambda}_{1,l,i}, \boldsymbol{\lambda}_{2,l,i}, \boldsymbol{\lambda}_{3,l,i}, \boldsymbol{\lambda}_{4,l,i} \ge \boldsymbol{0} \end{cases} \quad \forall l \in \mathcal{L} \tag{B24}$$

REFERENCES

[1] G. A. Hanasusanto, V. Roitch, D. Kuhn *et al*., "A distributionally robust perspective on uncertainty quantification and chance constrained programming," *Mathematical Programming*, vol. 151, pp. 35-62, Mar. 2015.

[2] F. Pourahmadi, J. Kazempour, C. Ordoudis *et al*., "Distributionally robust chance-constrained generation expansion planning," *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 2888-2903, Jul. 2020.

[3] C. Duan, W. Fang, L. Jiang *et al*., "Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4924-4936, Sept. 2018.

[4] Y. Zhang, S. Shen, and J. L. Mathieu, "Distributionally robust chance constrained optimal power flow with uncertain renewables and uncertain reserve provided by loads," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp, 1378-1388, Mar. 2017.

[5] W. Xie and S. Ahmed, "Distributionally robust chance constrained optimal power flow with renewables: a conic reformulation," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp, 1860-1867, Mar. 2018.

[6] L. Yang, Y. Xu, W. Gu *et al*., "Distributionally robust chance-constrained optimal power-gas flow under bidirectional interactions considering uncertain wind power," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1722-1735, Mar. 2021.

[7] T. Ding, B. Rui, F. Li *et al.*, "Interval power flow analysis using linear relaxation and optimality-based bounds tightening (OBBT) methods," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 177-188, Jan. 2015.

[8] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12-19, Feb. 2011.

Lun Yang received the B.S. degree in electrical engineering from Shihezi University, Shihezi, China, in 2016, and the M.S. degree in electrical engineering from Chongqing University, Chongqing, China, in 2019. Currently, he is pursuing the Ph.D. degree in Tsinghua-Berkeley Shenzhen Institute, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, China. His research interests include convex optimization, distributionally robust optimization and their applications in power and energy systems.

Yinliang Xu received the B.S. and M.S. degrees in control science and engineering from Harbin Institute of Technology, Harbin, China, in 2007 and 2009, respectively, and the Ph.D. degree in electrical and computer engineering from New Mexico State University, Las Cruces, NM, USA, in 2013. He is now an Associate Professor with Tsinghua-Berkeley Shenzhen Institute, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, China. His research interests include distributed control and optimization of power systems, and renewable energy integration.

Zheng Xu received the B.S. degree in electrical engineering from Southeast University, Nanjing, China, in 2017. She is currently pursuing her Ph.D. degree with Tsinghua-Berkeley Shenzhen Institute, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, China. Her research interests include modeling and strategy analysis of virtual power plants in electricity market and renewable energy integration.

Hongbin Sun received his double B.S. degrees from Tsinghua University, Beijing, China, in 1992, the Ph.D. from Department of Electrical Engineering, Tsinghua University, in 1996. From September 2007 to September 2008, he was a Visiting Professor with School of EECS at the Washington State University, Pullman, USA. He is now ChangJiang Scholar Chair Professor in Department of Electrical Engineering and the Director of Energy Management and Control Research Center, Tsinghua University. He is currently an IEEE Fellow and IET Fellow. He also serves as the editor of the IEEE Transactions on Smart Grid, associate editor of IET Renewable Power Generation, and member of the Editorial Board of four international journals and several Chinese journals. His research interests include electric power system operation and control with specific interests on energy management system, system-wide automatic voltage control, and energy system integration.