TOP (2022) 30:594–617
https://doi.org/10.1007/s11750-022-00634-4

ORIGINAL PAPER

Integrating unimodality into distributionally robust optimal power flow

Bowen Li[1] · **Ruiwei Jiang**[2] · **Johanna L. Mathieu**[3]

Received: 12 October 2021 / Accepted: 25 May 2022 / Published online: 18 June 2022

Abstract

To manage renewable generation and load consumption uncertainty, chance-constrained optimal power flow (OPF) formulations have been proposed. However, conventional solution approaches often rely on accurate estimates of uncertainty distributions, which are rarely available in reality. When the distributions are not known but can be limited to a set of plausible candidates, termed an ambiguity set, distributionally robust (DR) optimization can reduce out-of-sample violation of chance constraints. Nevertheless, a DR model may yield conservative solutions if the ambiguity set is too large. In view that most practical uncertainty distributions for renewable generation are unimodal, in this paper, we integrate unimodality into a moment-based ambiguity set to reduce the conservatism of a DR-OPF model. We review exact reformulations, approximations, and an online algorithm for solving this model. We extend these results to derive a new, offline solution algorithm. Specifically, this algorithm uses a parameter selection approach that searches for an optimal approximation of the DR-OPF model before solving it. This significantly improves the computational efficiency and solution quality. We evaluate the performance of the offline algorithm against existing solution approaches for DR-OPF using modified IEEE 118-bus and 300-bus systems with high penetrations of renewable generation. Results show that including unimodality reduces solution conservatism and cost without degrading reliability significantly.

Keywords Optimal power flow · Chance constraints · Distributionally robust optimization · α-Unimodality

Mathematics Subject Classification 90C15 · 90C22 · 90C34

✉ Johanna L. Mathieu
jlmath@umich.edu

Extended author information available on the last page of the article

1 Introduction

Previous research has developed models to ensure power system reliability under uncertainties (such as renewable generation forecast error) by chance-constrained optimal power flow (OPF), in which physical constraints are required to be satisfied with high probability, e.g., (Zhang and Li 2011; Jabr 2013; Vrakopoulou et al. 2013; Bienstock et al. 2014; Roald et al. 2013; Vrakopoulou et al. 2019; Li et al. 2019a). Conventional solution approaches to solving chance-constrained OPF include scenario approximation (Campi et al. 2009; Margellos et al. 2014), analytical reformulations for specific distributions (e.g., Gaussian) (Bienstock et al. 2014; Roald et al. 2013; Li et al. 2019a; Roald et al. 2015), and sample average approximation (SAA) (Pagnoncelli et al. 2009; Ahmed and Shapiro 2008). Scenario approximation approaches rely on a large number of scenarios and often provide overly conservative solutions. Analytical reformulations usually require less computational effort; however, it is often difficult to accurately estimate the *joint* probability distribution of the uncertain parameters and so solutions can be unreliable. SAA performs better as the number of samples increases, but that also increases its computational burden as more binary variables and constraints are needed when recasting the SAA formulation as a mixed-integer program.

In contrast, distributionally robust (DR) optimization ensures that a chance constraint holds with regard to all probability distributions within an ambiguity set (Ghaoui et al. 2003; Delage and Ye 2010; Stellato 2014; Jiang and Guan 2016). This approach is closely related to both robust and stochastic optimization because (1) it reduces to robust optimization if the ambiguity set is characterized by the support information only (and, as a result, includes all distributions on the support) and (2) it reduces to a (nominal) chance constraint if the ambiguity set includes only a single distribution. By incorporating distributional information of the uncertainty (such as moments) into the ambiguity set, DR optimization can achieve a better trade-off between cost and reliability than existing approaches. The conservatism of the DR approach is related to the ambiguity set: if it includes unrealistic distributions, then the solution may be more costly than necessary. A recent thrust of research in DR optimization is to incorporate structural information, e.g., unimodality, into the ambiguity set so that some unrealistic distributions can be eliminated. However, incorporating additional information often comes with additional computational burden.

The objective of this work is to assess the value of using both moment and structural information, specifically, unimodality, in DR-OPF. We first review an existing DR-OPF model using an ambiguity set based on moments and unimodality, as well as its reformulation and approximations. Then, we analyze approaches to solve this model, comparing an existing online algorithm to a new offline algorithm developed here. The newly developed offline algorithm aims to reduce the computational time while producing high-quality solutions (i.e., low optimality gap). Lastly, we investigate the trade-off between solution quality (i.e., cost and reliability) and computational burden with a variety of chance-constrained OPF models. We also demonstrate the efficiency of the offline algorithm.

Prior work on DR-OPF has derived tractable reformulations for moment ambiguity sets (Lubin et al. 2016; Summers et al. 2015; Mieth and Dvorkin 2018; Zhang et al. 2017; Xie and Ahmed 2018; Lu et al. 2019; Tong et al. 2018; Mieth and Dvorkin 2018), discrepancy ambiguity sets (Guo et al. 2018; Duan et al. 2018; Wang et al. 2018; Huang et al. 2021; Guo et al. 2019, 2018; Arrigo et al. 2021, 2022; Esteban-Pérez and Morales 2021; Jabr 2020; Ordoudis et al. 2021), and structural information such as symmetry (Roald et al. 2015), unimodality (Roald et al. 2015; Li et al. 2016, 2019b; Summers et al. 2015; Pourahmadi and Kazempour 2021), and log-concavity (Li et al. 2018). Xie and Ahmed (2018) considered two-sided chance constraints for generator capacities and transmission line limits and Huang et al. 2021, Guo et al. 2019, 2018, 2018, Wang et al. 2018, Duan et al. 2018, Arrigo et al. 2022 and Jabr 2020 constructed ambiguity sets based on discrepancy measures between the real distribution and the empirical distribution. Further extensions have included embedding dependencies among uncertainties (Arrigo et al. 2021), considering contextual information (Esteban-Pérez and Morales 2021), using joint chance constraints (Ordoudis et al. 2021), and solving applications such as generation planning (Pourahmadi and Kazempour 2021). Here, we consider an ambiguity set that incorporates the first two moments and a generalized α-unimodality (Dharmadhikari and Joag-Dev 1988), which is usually satisfied by the uncertainties included in OPF models, such as the wind power forecast error. Our prior work (Li et al. 2019b, 2016) developed exact reformulations, approximations, and an online solution algorithm that we leverage here.

As new contribution, in addition to applying the existing online algorithm (Li et al. 2019b) to solve DR-OPF, we derive a new, offline algorithm that uses an optimal parameter selection (OPS) approach to help construct a high-quality sandwich approximation of the DR chance constraints. This approach significantly improves upon the approximations in Li et al. (2019b) and Li et al. (2016)). The main step of this approach finds an optimal piece-wise linear (PWL) sandwich approximation of a concave function. The approximation is independent of the values of the decision variables and leads to a provably smallest approximation error. We analyze the construction and theoretical properties of this OPS approach. Specifically, we give an optimality condition and a construction algorithm. Then, we compare the online and offline algorithms in case studies based on modified 118-bus and 300-bus systems with high wind power penetration.

The remainder of the paper is organized as follows. In Sect. 2, we review fundamental concepts and generalize the DR formulations and theoretical results in Li et al. (2019b) with unimodality information. In Sect. 3, we derive the DR-OPF reformulation using the generalized DR results. In Sect. 4, we introduce the OPS approach and give the online and offline DR-OPF solution algorithms. In Sect. 5, we compare the performance of the offline algorithm to existing alternatives and discuss the value of including unimodality information in DR-OPF. Section 6 concludes the paper.

2 Distributionally robust chance constraints

In this section, we review DR chance constraints and generalize the results from Li et al. (2019b). We consider random inequalities in the form

$$a(x)^\top \xi \le b(x), \tag{1}$$

where $x \in \mathbb{R}^n$ represents decision variables and $a(x) : \mathbb{R}^n \to \mathbb{R}^l$ and $b(x) : \mathbb{R}^n \to \mathbb{R}$ represent two affine functions of x. Uncertainty $\xi \in \mathbb{R}^l$ is defined on probability space $(\mathbb{R}^l, \mathcal{B}^l, \mathbb{P}_\xi)$ with Borel σ-algebra $\mathcal{B}^l$ and probability distribution $\mathbb{P}_\xi$. To ensure (1) is satisfied with at least a probability threshold $1-\epsilon$, we define the chance constraint (Charnes et al. 1958; Miller and Wagner 1965)

$$\mathbb{P}_\xi\big(a(x)^\top \xi \le b(x)\big) \ge 1-\epsilon, \tag{2}$$

where $1-\epsilon$ often takes a large value (e.g., 0.99).

In this paper, we consider two types of ambiguity sets. The first includes moment information only

$$\mathcal{D}_\xi := \Big\{ \mathbb{P}_\xi \in \mathcal{P}^l : \mathbb{E}_{\mathbb{P}_\xi}[\xi] = \mu,\ \mathbb{E}_{\mathbb{P}_\xi}[\xi\xi^\top] = \Sigma \Big\}, \tag{3}$$

and the second includes moment and unimodality information

$$\mathcal{U}_\xi := \big\{ \mathbb{P}_\xi \in \mathcal{P}^l_\alpha \cap \mathcal{D}_\xi :\ \mathcal{M}(\xi) = m \big\}, \tag{4}$$

where $\mathcal{P}^l_\alpha$ and $\mathcal{P}^l$ denote all probability distributions on $\mathbb{R}^l$ with and without the requirement of α-unimodality, respectively; μ and Σ denote the first and second moments of ξ; and $\mathcal{M}(\xi) = m$ specifies that the true mode of ξ is m. The value of α determines the shape of the unimodal distribution (Dharmadhikari and Joag-Dev 1988). When $\alpha = 1$, all the marginal distributions are univariate unimodal (i.e., the density function has a single peak called the mode and decaying tails). When $\alpha = l$, the density function of ξ has a single peak at m and is non-increasing along any rays emanating from m. As $\alpha \to \infty$, the requirement of unimodality gradually relaxes until it disappears. In reality, most uncertainties such as wind power forecast error follow a "bell-shaped" unimodal distribution, which is empirically justified with historical data in Li et al. (2018, 2019c).

The DR chance constraint with ambiguity set $\mathcal{D}_\xi$ is

$$\inf_{\mathbb{P}_\xi \in \mathcal{D}_\xi} \mathbb{P}_\xi\big(a(x)^\top \xi \le b(x)\big) \ge 1-\epsilon \tag{5}$$

and admits the following reformulation.

Theorem 1 (Theorem 2.2 in (Wagner 2008)) *The DR chance constraint* (5) *is equivalent to*

$$\sqrt{\left(\frac{1-\epsilon}{\epsilon}\right)a(x)^\top(\Sigma-\mu\mu^\top)a(x)} \le b(x)-a(x)^\top\mu. \tag{6}$$

Li et al. (2019b) derived an exact reformulation for the DR chance constraint with $\mathcal{U}_\xi$, i.e.,

$$\inf_{\mathbb{P}_\xi\in\mathcal{U}_\xi} \mathbb{P}_\xi\big(a(x)^\top\xi \le b(x)\big) \ge 1-\epsilon; \tag{7}$$

however, those results were derived assuming $m=0$, i.e., the mode is at the origin. Without loss of generality, we can rewrite (1) as $a(x)^\top(\xi-m) \le b(x)-a(x)^\top m$ with $\xi-m$ being our random vector, whose mode is at the origin, and generalize all of the corresponding results in Li et al. (2019b) to allow any non-zero mode value m.

Theorem 2 (A generalization of Theorem 1 in (Li et al. 2019b)) *The DR chance constraint* (7) *is equivalent to*

$$\begin{aligned}\sqrt{\frac{1-\epsilon-\tau^{-\alpha}}{\epsilon}}\|\Lambda a(x)\| &\le \tau\big(b(x)-a(x)^\top m\big)\\ &\quad-\left(\frac{\alpha+1}{\alpha}\right)(\mu-m)^\top a(x), \quad \forall\tau\ge\left(\frac{1}{1-\epsilon}\right)^{1/\alpha},\end{aligned} \tag{8}$$

where $\Lambda := \left(\left(\frac{\alpha+2}{\alpha}\right)(\Sigma-\mu\mu^\top)-\frac{1}{\alpha^2}(\mu-m)(\mu-m)^\top\right)^{1/2}$.

Reformulation (8) consists of an infinite number of second-order conic (SOC) constraints, parameterized by τ. To this end, Li et al. (2019b) developed a sandwich approximation of (8), which is asymptotic in the sense that it recovers (8) as more choices of τ are included. The sandwich approximation consists of a relaxed and a conservative approximation, which provide a superset and a subset of the original feasible region, respectively. In a minimization problem, solving the relaxed approximation results in a lower bound on the true optimal cost. An upper bound can be achieved by solving the conservative approximation. In this paper, to control the tightness of the approximations, we use parameters $\{n_k : k=1,\dots,K\}$ in the following two propositions. The approximations converge to the exact reformulations in Theorem 2 as the integer K increases. With the same K, different selections of n_k in the approximations result in different levels of tightness and performance.

Proposition 3 (Relaxed Approximation, a Generalization of Proposition 4 in (Li et al. 2019b)) *For integer* $K\ge 1$ *and real numbers* $\tau_0 \le n_1 < n_2 < \cdots < n_K \le \infty$, (7) *implies the SOC constraints*

$$\begin{aligned}\sqrt{\frac{1-\epsilon-n_k^{-\alpha}}{\epsilon}}\|\Lambda a(x)\| &\le n_k\big(b(x)-a(x)^\top m\big)\\ &\quad-\left(\frac{\alpha+1}{\alpha}\right)(\mu-m)^\top a(x), \quad \forall k=1,\dots,K.\end{aligned} \tag{9}$$

Proposition 4 (*Conservative Approximation, a Generalization of Proposition* 5 *in* (*Li et al.* 2019*b*)) *For integer* $K \geq 2$ *and real numbers* $\tau_0 = n_1 < \cdots < n_K = \infty$, *define a PWL function containing* $(K-1)$ *pieces*:

$$h(\tau) = \min_{k=2,\ldots,K} \left\{ \sqrt{\frac{1}{\epsilon(1-\epsilon-n_k^{-\alpha})}} \left[\left(\frac{\alpha n_k^{-\alpha-1}}{2} \right) \tau + 1 - \epsilon - \left(1 + \frac{\alpha}{2} \right) n_k^{-\alpha} \right] \right\}. \tag{10}$$

Set $q_1 = \tau_0$ *and denote* $q_2 < \cdots < q_{K-1}$ *as the* $(K-2)$ *break points of function* $h(\tau)$. *Then*, (7) *is implied by the SOC constraints*

$$h(q_k)\|\Lambda a(x)\| \leq q_k \big(b(x) - a(x)^\top m\big) - \left(\frac{\alpha+1}{\alpha} \right)(\mu - m)^\top a(x), \quad \forall k = 1, \ldots, K-1. \tag{11}$$

In this work, we consider a formulation that uses individual chance constraints, as opposed to joint chance constraints, mainly to take advantage of the high computational tractability granted by the exact reformulation and the sandwich approximation. Although the formulation does not provide probabilistic guarantees that all constraints are satisfied jointly, we may still achieve high performance due to the constraint correlations and the DR formulation (Li et al. 2019a, b). Furthermore, joint chance constraint techniques (Margellos et al. 2014; Hou and Roald 2020) might generate overly conservative solutions or require high sample complexity. Even when probability guarantees on the joint constraint satisfaction are required, we can apply individual chance constraints (e.g., using the Bonferroni approximation (Baker and Bernstein 2019; Xie et al. 2019)) to obtain a good and conservative approximation.

3 Distributionally robust optimal power flow

We apply the models and reformulations of Sect. 2 to a chance-constrained OPF that uses the DC power flow approximation (Vrakopoulou et al. 2013). We assume that the system has N_W wind power plants with forecast error $\tilde{w} \in \mathbb{R}^{N_W}$ (each element is represented by $\tilde{w}_i$), N_G generators with index set $\mathcal{G}$, N_B buses, N_D loads, and N_L transmission lines with index set $\mathcal{L}$. The forecast errors are calculated as the difference between actual wind power realizations and their corresponding forecasts and are compensated by reserves. Design variables include generation output $P_G \in \mathbb{R}^{N_G}$, up and down reserve capacities $R_G^{up} \in \mathbb{R}^{N_G}, R_G^{dn} \in \mathbb{R}^{N_G}$, and a distribution vector $d_G \in \mathbb{R}^{N_G}$, which determines how much reserve each generator provides to balance the total forecast error. Without loss of generality, we employ $\mathcal{U}_\xi$ as the ambiguity set. Then the DR-OPF formulation for a given confidence level ϵ is

$$\min\ P_G^{\top}[C_1]P_G + C_2^{\top}P_G + C_R^{\top}(R_G^{up} + R_G^{dn})$$

$$\text{s.t.}\ \inf_{\mathbb{P}_\xi \in \mathcal{U}_\xi} \mathbb{P}_\xi \Big(A_i(C_W - C_G d_G \mathbf{1}^{\top})\tilde{w} + A_i(C_L P_L - C_W P_W^f - C_G P_G) \le P_{l,i} \Big) \ge 1-\epsilon,\ \forall i \in \mathcal{L}, \tag{12a}$$

$$\inf_{\mathbb{P}_\xi \in \mathcal{U}_\xi} \mathbb{P}_\xi \Big(A_i(C_W - C_G d_G \mathbf{1}^{\top})\tilde{w} + A_i(C_L P_L - C_W P_W^f - C_G P_G) \ge -P_{l,i} \Big) \ge 1-\epsilon,\ \forall i \in \mathcal{L}, \tag{12b}$$

$$R_G = -d_G\Big(\sum_{i=1}^{N_W} \tilde{w}_i\Big), \tag{12c}$$

$$P_{\text{inj}} = C_G(P_G + R_G) + C_W(P_W^f + \tilde{w}) - C_L P_L, \tag{12d}$$

$$\inf_{\mathbb{P}_\xi \in \mathcal{U}_\xi} \mathbb{P}_\xi \Big(-d_{G,i}\mathbf{1}^{\top}\tilde{w} \le \overline{P}_{G,i} - P_{G,i} \Big) \ge 1-\epsilon,\ \forall i \in \mathcal{G}, \tag{12e}$$

$$\inf_{\mathbb{P}_\xi \in \mathcal{U}_\xi} \mathbb{P}_\xi \Big(-d_{G,i}\mathbf{1}^{\top}\tilde{w} \ge \underline{P}_{G,i} - P_{G,i} \Big) \ge 1-\epsilon,\ \forall i \in \mathcal{G}, \tag{12f}$$

$$\inf_{\mathbb{P}_\xi \in \mathcal{U}_\xi} \mathbb{P}_\xi \Big(-d_{G,i}\mathbf{1}^{\top}\tilde{w} \le R_{G,i}^{up} \Big) \ge 1-\epsilon,\ \forall i \in \mathcal{G}, \tag{12g}$$

$$\inf_{\mathbb{P}_\xi \in \mathcal{U}_\xi} \mathbb{P}_\xi \Big(-d_{G,i}\mathbf{1}^{\top}\tilde{w} \ge -R_{G,i}^{dn} \Big) \ge 1-\epsilon,\ \forall i \in \mathcal{G}, \tag{12h}$$

$$\mathbf{1}_{1\times N_G} d_G = 1, \tag{12i}$$

$$\mathbf{1}_{1\times N_B}(C_G P_G + C_W P_W^f - C_L P_L) = 0, \tag{12j}$$

$$P_G \ge \mathbf{0}_{N_G\times 1},\ d_G \ge \mathbf{0}_{N_G\times 1}, \tag{12k}$$

$$R_G^{up} \ge \mathbf{0}_{N_G\times 1},\ R_G^{dn} \ge \mathbf{0}_{N_G\times 1}, \tag{12l}$$

where $[C_1] \in \mathbb{R}^{N_G\times N_G}$, $C_2 \in \mathbb{R}^{N_G}$, and $C_R \in \mathbb{R}^{N_G}$ are cost parameters. Constraints (12a) and (12b) bound the power flows by the line limits $P_l \in \mathbb{R}^{N_L}$. The power flow

is calculated from the power injections $P_{\text{inj}} \in \mathbb{R}^{N_B}$ in (12d) and the constant matrix $A \in \mathbb{R}^{N_L \times N_B}$, which is calculated using the admittance matrix and network connections (Vrakopoulou et al. 2013). Constraint (12c) computes the real-time reserve response R_G that is bounded by the reserve capacities $R_G^{dn} \in \mathbb{R}^{N_G}$ and $R_G^{up} \in \mathbb{R}^{N_G}$ in (12g) and (12h). In (12d), $P_W^f \in \mathbb{R}^{N_W}$ is the wind power forecast, $P_L \in \mathbb{R}^{N_D}$ is the load (which is assumed to be known, though the formulation can be easily extended to handle uncertain loads[1]), and $C_G \in \mathbb{R}^{N_B \times N_G}$, $C_W \in \mathbb{R}^{N_B \times N_W}$, and $C_L \in \mathbb{R}^{N_B \times N_D}$ are constant matrices that map generators, wind power plants, and loads to buses. Constraints (12e) and (12f) bound generator outputs within their limits $[\underline{P}_G, \overline{P}_G]$; (12i), (12j) enforce power balance with and without wind power forecast error; and (12k), (12l) ensure all decision variables are non-negative.

4 Solution algorithm

The quality of the sandwich approximation in Propositions 3 and 4 is a function of the choice of parameters n_k. To solve the DR-OPF model, Li et al. (2019b) propose the following online algorithm that selects n_k values on-the-fly and solves the corresponding sandwich approximation iteratively.

Algorithm 1 Online Algorithm [37]

Initialization: $k = 1$, $\tau_0 = \left(\frac{1}{1-\epsilon}\right)^{1/\alpha}$, gap tolerance η;

Iteration k:

Step 1: Solve the conservative approximation (11) using τ_j for all $j = 0, \ldots, k-1$ and $\tau^k = \infty$ and obtain optimal solution x_k^u and objective f_k^u;

Step 2: Solve the relaxed approximation (9) using τ_j for all $j = 0, \ldots, k-1$ and obtain optimal solution x_k^l and objective f_k^l;

Step 3: **IF** $(f_k^u - f_k^l)/f_k^l \leq \eta$, **STOP** and **RETURN** x_k^u as optimal solution; **ELSE GOTO** Step 4;

Step 4 (**Separation**): Find worst case τ^* that results in the largest violation of (8) under x_k^l

$$\tau^* = \text{argmax}_{\tau \geq \tau_0} \left\{ \sqrt{\frac{1-\epsilon-\tau^{-\alpha}}{\epsilon}} \|\Lambda a(x_k^l)\| - \tau \left(b(x_k^l) - a(x_k^l)^\top m \right) \right\}; \quad (13)$$

Step 5: Set $\tau_k = \tau^*$ and $k = k+1$, **GOTO** Step 1.

[1] To handle uncertain loads, we denote P_L^f and $\tilde{d}$ as the load forecast and the uncertain load forecast error, respectively. We substitute the vector of loads P_L with $P_L^f + \tilde{d}$ and the system uncertainty $\tilde{w}$ with $\tilde{w} + \tilde{d}$ in formulation (12). Then, we can use the same reformulation method and solution algorithm as described in Section 4 to solve this extended model.

In each iteration k, Algorithm 1 chooses a value of n_k that yields the largest violation of constraint (8) given the incumbent solution x_k^I (see Step 4). This choice is critical to improve the relaxed approximation. However, it provides no guarantee on the quality of the conservative approximation, whose solution is what Algorithm 1 returns. Next, we propose a new, offline algorithm selects the n_k values in a provably optimal way and admits guarantees on the quality of both the relaxed and conservative approximations. It also reduces the computational effort in each iteration.

4.1 Optimal parameter selection (OPS)

The OPS approach seeks a concave and $|\mathcal{S}|$-piece PWL function $g(\tau)$, whose break points are denoted by q_s for $s \in \mathcal{S}$, that outer approximates the nonlinear function $v(\tau)$ on the interval $[\tau_0, \infty)$. Existing approaches to finding PWL approximations for convex/concave functions when they are defined on a bounded domain include (Cox 1971; Imamoto and Tang 2008; Vandewalle 1975; Gavrilović 1975). We extend the prior work to an unbounded domain by identifying optimality conditions for an optimal $g(\tau)$ and design an algorithm to construct such a $g(\tau)$.

4.2 Optimality conditions

We denote the $|\mathcal{S}|$-piece PWL function $g(\tau) = \max_{s\in\mathcal{S}} g_s(\tau)$, where $g_s(\tau) := d_s\tau + f_s$ represents a piece of $g(\tau)$ for all $s \in \mathcal{S}$ and d_s represents the slope of $g_s(\tau)$ and is non-increasing in s. If we let $\mathcal{H}_s$ denote the part of domain $g_s(\tau) = g(\tau)$, then the error $e^{\max}$ of the PWL outer approximation $g(\tau)$ for $v(\tau)$ can be evaluated by the largest discrepancy between these functions,

$$e^{\max} := \max_{s\in\mathcal{S}} \max_{\tau\in\mathcal{H}_s} \{d_s\tau + f_s - v(\tau)\}. \tag{13}$$

Next, we identify conditions for $g(\tau)$ to be an optimal PWL approximation of $v(\tau)$, i.e., one achieving the smallest $e^{\max}$ for fixed $|\mathcal{S}|$. We omit the proof because it is a straightforward extension of (Gavrilović 1975).

Theorem 5 (Optimality Conditions) *For $|\mathcal{S}| \geq 1$, an $|\mathcal{S}|$-piece PWL function $g(\tau)$ is an optimal PWL outer approximation of $v(\tau)$ if the following three conditions hold.*

1. $g_{|\mathcal{S}|}(\tau) = \sqrt{\frac{1-\epsilon}{\epsilon}}$.
2. *$g_s(\tau)$ is tangent to $v(\tau)$ for all $s \in \mathcal{S}$.*
3. $g(B_{s_1}) - v(B_{s_1}) = g(B_{s_2}) - v(B_{s_2})$, $\forall s_1, s_2 \in \mathcal{S}$, *where $\{B_s,\ s \in \mathcal{S}\}$ consists of all break points of $g(\tau)$ and τ_0.*

where ϵ is the violation probability in (2).

The intuition of the optimality condition is that, at optimality, all of the local approximation errors, evaluated at the break points of the optimal $g(\tau)$, are all equal (i.e., $E_1^i = E_2^i = E_3^i = e_T^i$ in Fig. 1). If the local approximation errors are not all equal, we can always move the break points to reduce the largest error $e^{\max}$. Defining $e^* := g(B_s) - v(B_s)$ for any $s \in \mathcal{S}$, we note that $g(\tau) - e^*$ is an inner approximation of $v(\tau)$ with a (maximum discrepancy) error e^*.

4.3 Construction algorithm

We provide a constructive algorithm to search for an optimal $g(\tau)$ that satisfies the optimality conditions identified in Theorem 5. This algorithm is adapted from the recursive descent algorithm in Imamoto and Tang (2008). We first define the following notation.

- $\Delta^i \in \mathbb{R}^{\mathcal{I}}$: the step size in iteration i;
- δ: the threshold for the termination criterion;
- $\mathcal{I}$: the maximum number of iterations;
- $g^i(\tau)$: estimation of $g(\tau)$ in iteration i ($g_s^i(\tau)$ represents the sth piece of $g^i(\tau)$);
- $B^i \in \mathbb{R}^{|\mathcal{S}|}$: the break points of $g^i(\tau)$ in iteration i (B_s^i represents the sth break point and $B_1^i = \tau_0$);
- $T^i \in \mathbb{R}^{|\mathcal{S}|-1}$: the points at which $g^i(\tau)$ is tangent to $v(\tau)$ in iteration i;
- $E^i \in \mathbb{R}^{|\mathcal{S}|}$: the distance between $g(\tau)$ and $v(\tau)$ at all B_s^i in iteration i.
- $e_T^i := \sqrt{(1-\epsilon)/\epsilon} - v(B_{|\mathcal{S}|}^i)$ represents the distance between the last constant piece (i.e., $g_{|\mathcal{S}|}(\tau)$) and $v(\tau)$ at $B_{|\mathcal{S}|}^i$.

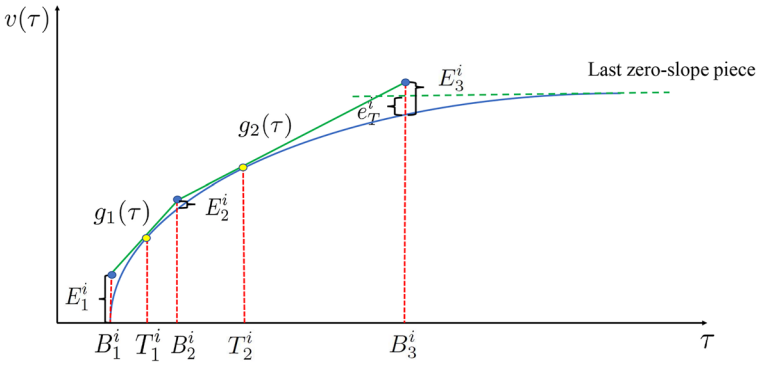

Fig. 1 Illustrative example of Algorithm 2 for $|\mathcal{S}| = 3$, at iteration i. Function $g_i(\tau)$ (green solid lines) is a PWL outer approximation of $v(\tau)$ (blue curve) with tangent points T^i (yellow dots) and break points B^i (blue dots). At convergence, $E_1^i = E_2^i = E_3^i = e_T^i$ (color figure online)

Algorithm 2 Parameter searching algorithm

Initialization: $i = 1$, $\Delta^1 = 1$, $\delta = 0.01$, $\mathcal{I} = 50$, $B^1_{|\mathcal{S}|} = 10$, $T^1_s = \tau_0 + \frac{s(B^1_{|\mathcal{S}|} - \tau_0)}{|\mathcal{S}|}$ for $s = 1, \ldots, |\mathcal{S}| - 1$;

Iteration i:

Step 1: **IF** $i \leq \mathcal{I}$, calculate B^i_s for $s = 2, \ldots, |\mathcal{S}| - 1$ by solving $g^i_{s-1}(B^i_s) = g^i_s(B^i_s)$, where $g^i_s(\tau) = v'(T^i_s)(\tau - T^i_s) + v(T^i_s)$; **ELSE STOP** and **RETURN** no convergence under current initialization;

Step 2: Calculate E^i and e^i_T; **IF** $E^i_{|\mathcal{S}|}/e^i_T > 1 + \delta$ or $E^i_{|\mathcal{S}|}/e^i_T < 1/(1 + \delta)$, set $B^i_{|\mathcal{S}|} = 0.5(\hat{\tau} + B^i_{|\mathcal{S}|})$, where $\hat{\tau}$ is the solution of $g^i_{|\mathcal{S}|-1}(\hat{\tau}) = \sqrt{(1 - \epsilon)/\epsilon}$, and **GOTO** Step 1; **ELSE GOTO** Step 3;

Step 3: **IF** $\max(E^i) \leq (1 + \delta)\min(E^i)$, **STOP** and **RETURN** $g^i(\tau)$ as an optimal solution; **ELSE GOTO** Step 4;

Step 4: **IF** $i = 1$, **GOTO** Step 5; **ELSEIF** $\max(E^i) > \max(E^{i-1})$, set $i = i-1$, $\Delta^i = \Delta^i/2$, and **GOTO** Step 5; **ELSE GOTO** Step 5;

Step 5: For $s = 1, \ldots, |\mathcal{S}| - 1$, calculate

$$T^{i+1}_s = T^i_s + \frac{\Delta^i(E^i_{s+1} - E^i_s)}{\frac{E^i_{s+1}}{B^i_{s+1} - T^i_s} + \frac{E^i_s}{T^i_s - B^i_s}}$$

and set $i = i + 1$ and **GOTO** Step 1.

The algorithm is described in Algorithm 2 and Fig. 1 displays an example for $|\mathcal{S}| = 3$. First, Δ^1, δ, $\mathcal{I}$, and $B^1_{|\mathcal{S}|}$ are initialized (for different $v(\tau)$, they may take other reasonable values) and T^1 is computed by evenly dividing $[\tau_0, B^1_{|\mathcal{S}|}]$ into $|\mathcal{S}|$ segments. In each iteration i, Step 1 calculates B^i using T^i and $B^i_{|\mathcal{S}|}$. Step 2 calculates E^i and e^i_T and adjusts $B^i_{|\mathcal{S}|}$ accordingly. Step 3 checks for convergence. Step 4 repeats the iteration with a smaller step size if the previous step did not produce an improvement. Step 5 adjusts T^i to further reduce the differences among E^i_s. The optimal break points B^* obtained from Algorithm 2 can be used in Proposition 4 to establish the corresponding conservative approximation.

Remark: The optimal parameters obtained from Algorithm 2 are unique given a choice of $|\mathcal{S}|$. That is, they are independent of the decision variables but dependent on parameters ϵ and α. Hence, Algorithm 2 can be conducted offline to find an optimal PWL approximation.

4.4 Performance

Figure 2 shows the convergence of Algorithm 2 under various values of $|\mathcal{S}|$. We observe that as $|\mathcal{S}|$ increases, the optimal approximation error $e^{\max}$ decreases and the total number of iterations grows almost linearly.

4.5 Algorithm

Since OPS can be conducted offline via Algorithm 2, we can compute the optimal PWL approximations prior to solving the DR-OPF model (e.g., as a look-up table). The offline algorithm is summarized in Algorithm 3, which takes the optimal PWL approximations as input in each iteration.

Algorithm 3 Offline Algorithm

Initialization: $k = 1$, $\tau_0 = \left(\frac{1}{1-\epsilon}\right)^{1/\alpha}$, gap tolerance η;

Iteration K:

Step 1 (**Optimal Parameter Selection**): find an optimal K-piece PWL approximation $g(\tau)$ using Algorithm 2 and set $\{q_k\}_{k=1}^K$ as the optimal break points;

Step 2: Solve the conservative approximation (11) using $\{q_k\}_{k=1}^K$ to obtain an optimal solution x_K^u and an optimal value f_K^u;

Step 3: Solve the relaxed approximation (9) using $\{q_k\}_{k=1}^K$ to obtain an optimal solution x_K^l and an optimal value f_K^l;

Step 4: **IF** $(f_K^u - f_K^l)/f_K^l \leq \eta$, **STOP** and **RETURN** x_K^u as an optimal solution; **ELSE** set $K = K + 1$ and **GOTO** Step 1.

5 Case studies

We test our approaches on the IEEE 118-bus and 300-bus systems, which are modified to include a large number of wind power plants with a total of 400 MW ($\sim$ 9.4% of total load) and 2000 MW ($\sim$ 8.5% of total load) of forecasted wind power, respectively. We use the network and cost parameters from Coffrin et al. (2016) and set $C_R = 10C_2$. We add wind power to all buses with generators and allocate the forecast wind power to these buses in proportion to their generation limit.

We also test our approaches using two forecast error data sets with different characteristics. We define the forecast error ratio as the ratio between the forecast error and the corresponding forecast. *Data Set 1* **(DS1)** was used in Vrakopoulou et al. (2019). The data set is generated using the Markov–Chain Monte Carlo mechanism (Papaefthymiou and Klockl 2008) on real wind power forecasts and realizations from Germany. The wind power is well-forecasted with small forecast error ratios (-30 to 60%). For each wind bus, we randomly select the forecast errors from the same data pool without considering spatial correlation. *Data Set 2* **(DS2)** is constructed from the RE-Europe data set (Jensen and Pinson 2017), which contains hourly wind power forecasts and realizations based on the European energy system. The data set includes strong spatiotemporal correlation. However, the data set also contains poor forecasts with extreme forecast error ratios, up to 5300% (Li 2019). Therefore, we scale down the forecast errors by 60% and then filter outliers with forecast error ratios larger than 100%.

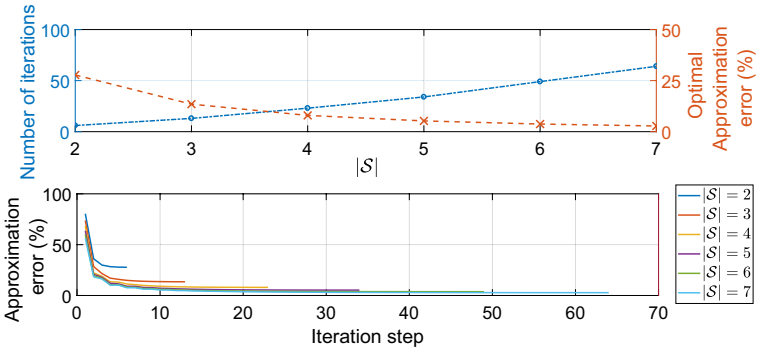

Fig. 2 Number of iterations and optimal approximation error as a function of $|\mathcal{S}|$ (top); Convergence of the approximation error for different $|\mathcal{S}|$ (bottom)

We use 5000 randomly selected data points for the 118-bus system and 8000 for the 300-bus system to construct $\mathcal{D}_\xi$ and $\mathcal{U}_\xi$. More data is needed for the 300-bus system since the uncertainty dimension is larger. In addition, we use histograms with 15 and 20 bins to justify our assumption of unimodality and determine the locations of mode m for DS1 (see Fig. 3) and DS2 (see Fig. 4) by identifying the bin with the most points (Li et al. 2019c).

Further, to evaluate reliability, we randomly select 5000 and 8000 data points to conduct out-of-sample tests for the 118-bus and 300-bus systems, respectively. We define the reliability as the percentage of wind power forecast errors for which all chance constraints are satisfied. We perform three parallel tests by randomly reselecting the data used to construct the ambiguity sets.

We benchmark our approaches against two conventional approaches. *Analytical reformulation assuming multivariate Gaussian distributions* **(AR)** is used in (Bienstock et al. 2014; Li et al. 2019a; Roald et al. 2013), which uses moments determined from the data. Then all chance constraints can be exactly reformulated as SOC constraints. The *scenario-based method* **(SC)** developed in Margellos et al. (2014) requires that the constraints affected by uncertainties are satisfied for any realization within a probabilistically robust set constructed using a sufficient number of randomly selected uncertainty realizations.

We solve all optimization problems using CVX with the Mosek solver (Grant and Boyd 2014, 2008). We set $\epsilon = 5\%$, $\alpha = 1$, and optimality gap $\eta = 1\%$. The choice of α is valid because, in general, wind power forecast error is marginally unimodal (Li 2019).

The results are organized as follows. In Sect. 5.1, we compare the performance of DR approaches to the benchmark approaches in terms of cost, reliability, and overall computational performance. We also show the advantage of including unimodality information in the DR-OPF ambiguity set. In Sect. 5.2, we analyze the computational performance of the online Algorithm 1 to demonstrate the computational burden for large systems. In Sect. 5.3, we compare the proposed offline Algorithm 3 to the online Algorithm 1 to demonstrate the computational improvement.

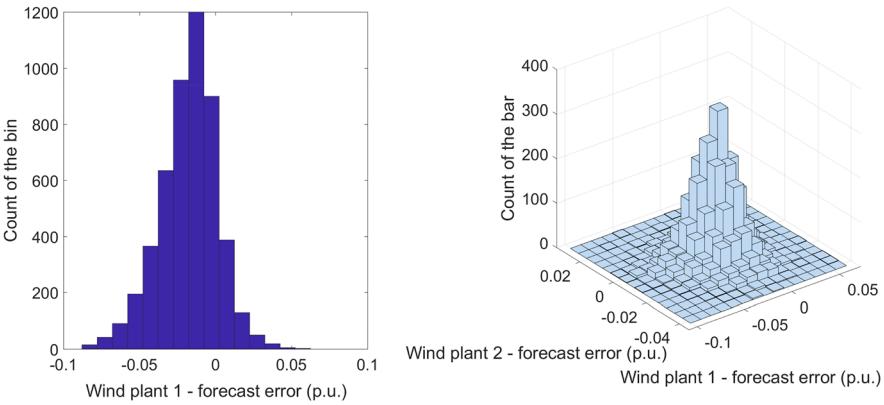

Fig. 3 Histograms of univariate and bivariate wind forecast errors of DS1 (15 bins)

5.1 Cost and reliability performance

We first compare the DR approaches to the benchmark approaches in terms of objective cost and reliability. The results are summarized in Table 1. **(DR-M)** is the DR approach with ambiguity set $\mathcal{D}_\xi$, which does not include the unimodality assumption. **(DR-U)** is the DR approach with ambiguity set $\mathcal{U}_\xi$, solved using the online Algorithm 1. To facilitate comparisons, we define a percentage difference on cost (C/Diff) and reliability (R/Diff) against the benchmarks, where AR generally produces low-cost solutions that are not sufficiently reliable and SC generally produces high-cost solutions with higher reliability than necessary. Specifically, we calculate the C/Diff of a DR approach as the difference in cost compared to that of the AR approach divided by the difference in cost between the AR and SC approaches. The R/Diff is defined similarly. Small C/Diffs are desirable, i.e., low costs approaching that of the AR approach. Large R/Diffs are desirable, i.e., high reliability approaching that of the SC approach. We define the improvement (Improv) of a DR approach to be its R/Diff divided by its C/Diff. Large Improvs are desirable, indicating a better trade-off between cost and reliability.

From Table 1, we observe that SC provides overly conservative results with the highest costs and 100% reliability in all instances, because it provides a joint probabilistic guarantee and uses a robust optimization method. In contrast, AR provides the least conservative results with the lowest costs and the lowest reliability (below 95% in all instances), because it assumes a Gaussian distribution, which may not hold in reality. Meanwhile, DR approaches provide intermediate costs and reliability, with all reliability values lying above the target (95%). Of the two DR approaches, DR-U provides higher costs and higher reliability than DR-M since DR-M does not include the unimodality assumption. DR-U incorporates the unimodality information and achieves a better cost performance, while maintaining satisfactory reliability. If we compare the Diffs and Improvs of DR-U and DR-M, we see that DR-U provides a better trade-off between cost and reliability. Solutions using DS1 are more

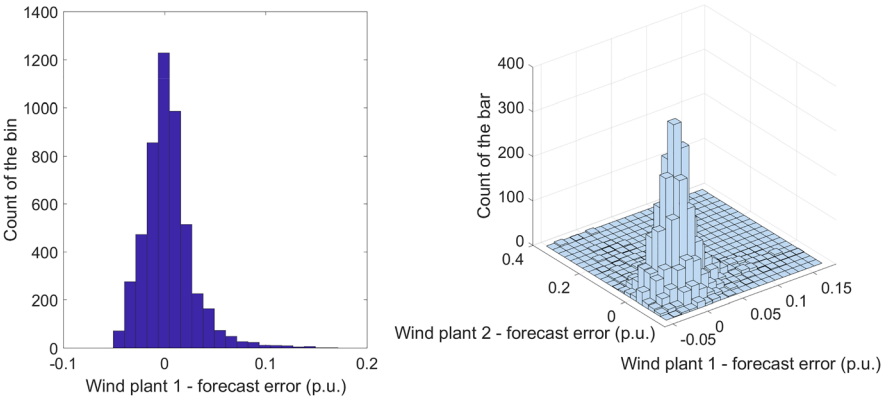

Fig. 4 Histograms of univariate and bivariate wind forecast errors of DS2 (20 bins)

stable with less variability across parallel tests than those using DS2. Additionally, solutions using DS1 have higher Improvs than those using DS2.

5.2 Computational performance

As shown in Table 1, DR-U requires significantly more computational time than the other approaches. This is because DR-U employs an iterative solution algorithm, while the other approaches do not. For large system dimensions, the computational burden becomes severe pointing to the need for a computationally efficient alternative.

Table 2 summarizes the percent of the total computational time to complete the **separation** step of Algorithm 1 and the required number of iterations. DS2 (poorer forecasts) requires a larger number of iterations than DS1, leading to overall computational times that are longer for DS2 than DS1, as seen in Table 1. The 118-bus system requires a higher percentage of computational time to complete the **separation** step than the 300-bus system. The total computation time of each iteration slightly increases over the iterations, while the time needed for the **separation** step is approximately constant.

Recall that in Algorithm 1, we output the solutions of the conservative approximation (i.e., x_k^u from Step (1) at termination. However, solutions from the relaxed approximation (i.e., x_k^l from Step (2) can also have low optimality gaps. Here, we check if the intermediate solutions from relaxed approximation of Algorithm 1 are good approximates of the optimal solution. Figure 5 shows the optimality gap and reliability of all the intermediate solutions x_k^l for the 118-bus system using DS2.[2] We find that the intermediate solutions are not good approximates because even solutions with small absolute optimality gaps (< 1%) can have low reliability (< 70%). We also observe that higher objective cost does not always guarantee higher reliability in out-of-sample tests.

[2] Note that each intermediate solution comes from a relaxed approximation and so the objective cost is lower than the true objective cost. Here we use a negative optimality gap to illustrate this relation.

Table 1 Objective costs, reliability (%), and computational times (seconds) for each approach

Bus/data set		AR			SC			DR-M						DR-U (online)					
		Cost	Reliab	Time	Cost	Reliab	Time	Cost	C/Diff	Reliab	R/Diff	Improv	Time	Cost	C/Diff	Reliab	R/Diff	Improv	Time
118/DS1	Min	3309	81.7	11.0	4935	100.0	11.2	3466	9.6	99.7	98.3	10.1	11.1	3340	1.9	97.0	83.6	40.4	475.1
	Avg	3310	81.8	11.4	4937	100.0	14.4	3467	9.6	99.7	98.3	10.2	11.3	3343	2.0	97.1	84.2	41.6	478.4
	Max	3310	81.9	11.8	4942	100.0	18.7	3468	9.7	99.7	98.4	10.2	11.4	3344	2.1	97.2	84.5	44.0	483.4
118/DS2	Min	3491	79.5	11.0	5902	100.0	10.5	4064	23.5	98.9	94.3	3.3	11.1	3703	8.7	95.0	74.7	8.1	2490.2
	Avg	3520	81.8	11.7	5926	100.0	11.6	4141	25.9	99.2	95.8	3.7	11.6	3736	9.0	95.6	75.7	8.4	3160.7
	Max	3564	85.4	12.4	5942	100.0	13.3	4261	29.3	99.7	97.9	4.1	12.3	3780	9.2	96.6	76.7	8.7	3815.9
300/DS1	Min	14408	72.9	125.4	18032	100.0	134.0	14579	4.7	99.6	98.5	20.9	124.5	14479	1.9	96.4	86.7	44.4	1948.5
	Avg	14409	73.6	127.0	18038	100.0	136.9	14580	4.7	99.6	98.6	21.0	125.3	14479	1.9	96.6	87.0	44.9	2158.5
	Max	14410	74.1	130.2	18046	100.0	140.2	14581	4.7	99.7	98.9	21.0	125.9	14480	2.0	96.7	87.4	45.4	2263.6
300/DS2	Min	15125	85.7	234.8	21249	100.0	141.7	16956	29.0	99.7	97.7	3.1	236.5	15724	8.9	96.7	76.9	6.8	6547.7
	Avg	15161	86.2	236.1	21373	100.0	143.1	17052	30.5	99.7	97.8	3.2	237.7	15777	9.9	97.1	78.8	8.0	10880.4
	Max	15191	87.1	237.1	21436	100.0	144.9	17128	32.0	99.7	97.9	3.4	239.4	15880	11.4	97.4	81.8	8.7	15720.6

Table 2 Algorithm 1, percent of time used for solving **separation** and number of iterations

	118/DS1			118/DS2			300/DS1			300/DS2		
	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max
Percent	86.2	86.8	87.4	85.4	85.6	85.9	59.1	59.1	59.2	43.1	43.6	43.9
Iterations	5	5	5	26	32	36	6	7	7	14	23	33

5.3 Online vs offline algorithm

In this section, we compare the online Algorithm 1 to the offline Algorithm 3 using OPS. For each network and data set we show the optimality gap, computational time (scaled to the total time in Table 1), and the reliability of the solutions obtained in each iteration. We focus on the first few iterations to highlight the progression of the algorithms.

IEEE 118-bus network with DS1

Figure 6 compares the online and offline algorithms using the IEEE 118-bus network and DS1. We find that the online algorithm fails to achieve a less than 1% optimality gap within the first four iterations while the offline algorithm achieves this goal by the second iteration and with much less computational time (6.02% of the time needed by the the online algorithm). In terms of reliability, we see that the offline algorithm's solutions are less reliable than those of the online algorithm but they still satisfy our requirement of 95%.

IEEE 300-bus network with DS1

Figure 7 compares the online and offline algorithms using the IEEE 300-bus network and DS1. Here, we observe similar trends to those of the IEEE 118-bus network. Specifically, the offline algorithm converges to below 1% optimality gap using only 19.62% of the total computational time needed by the online algorithm. However, the offline algorithm takes a longer time to solve these early iterations than the online algorithm.

IEEE 118-bus network with DS2

Figure 8 compares the online and offline algorithms using the IEEE 118-bus network and DS2. Here, we observe an even larger relative advantage from the offline algorithm. When the offline algorithm reaches less than 1% optimality gap, the online algorithm has a 20% optimality gap. The offline algorithm converges to below 1% optimality gap with only 5.51% of the total computational time needed by the online algorithm.

IEEE 300-bus network with DS2

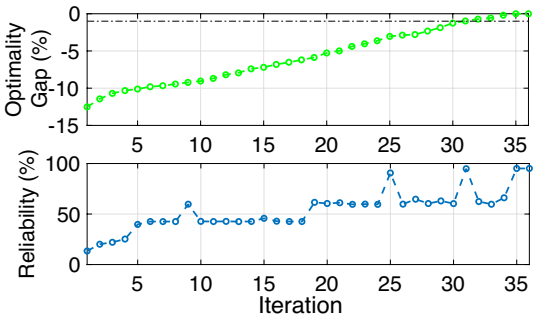

Fig. 5 Optimality gap and reliability of intermediate solutions from the relaxed approximation of Algorithm 1. Black dashed line marks 1% optimality gap (color figure online)

Figure 9 compares the online and offline algorithms using the IEEE 300-bus network and DS2. Again, the online algorithm has a large relative advantage. When the offline algorithm reaches less than 1% optimality gap, the online algorithm has a 11.39% optimality gap. Further, the offline algorithm converges to below 1% optimality gap with only 21.7% of the total computational time needed by the online algorithm. However, like with the IEEE 300-bus network and DS1, the offline algorithm takes a longer time to solve these early iterations than the online algorithm.

Cross comparisons

In all cases, the offline algorithm takes much less computational time than the online algorithm. The offline algorithm takes a similar amount of time as AR, SC, and DR-M, while the online algorithm exhibits an approximately linear relationship between computational time and iterations. All intermediate solutions of both algorithms satisfy the 95% constraint satisfaction level. Note that this does not contradict the conclusions from Fig. 5, in which solutions x_k^l are generated from a collection of relaxed approximations. Solutions x_k^l are not guaranteed to be feasible in the original DR-U. In contrast, the offline Algorithm 3 takes advantage of the conservative approximation and outputs solutions x_k^u, which guarantee feasibility in the original DR-U, hence high reliability. For both the online and offline algorithms, solutions become less conservative as iteration continues.

Figures 6 and 8 show that the intermediate solutions for DS2 generally have larger optimality gaps than those for DS1. In Figs. 7 and 9, we observe that the offline algorithm can take more computational time per iteration than the online algorithm in the early iterations (see Fig. 7). In general, the offline algorithm is less computationally advantageous for the IEEE-300 bus system than the IEEE-118 bus system.

In summary, the offline algorithm with OPS produces good quality solutions of DR-U with small optimality gaps and high reliability. In addition, it shows a much better convergence rates than the online algorithm.

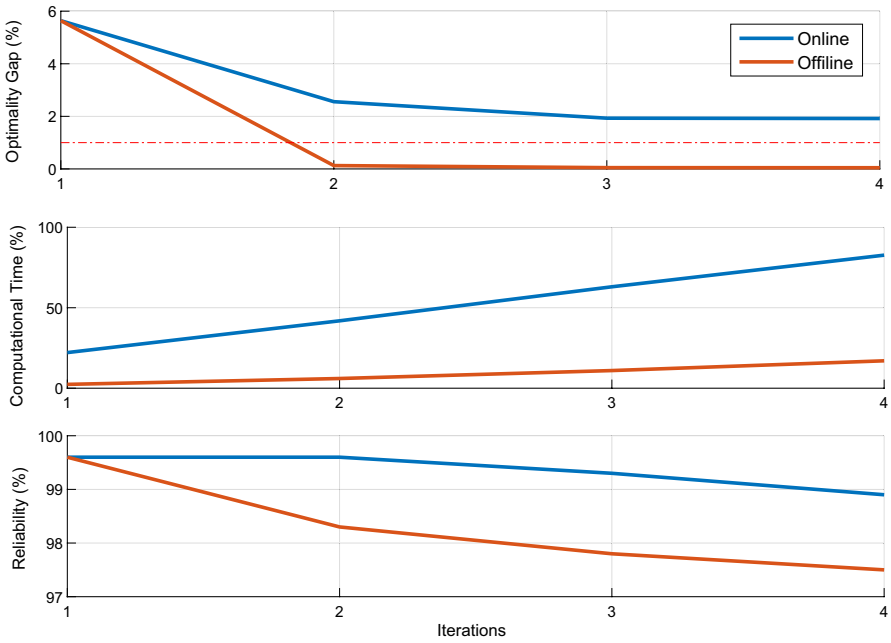

Fig. 6 Online vs Offline (118/DS1): optimality gap, computational time, and reliability of the solutions from the early iterations. The red dashed line marks the 1% optimality gap (color figure online)

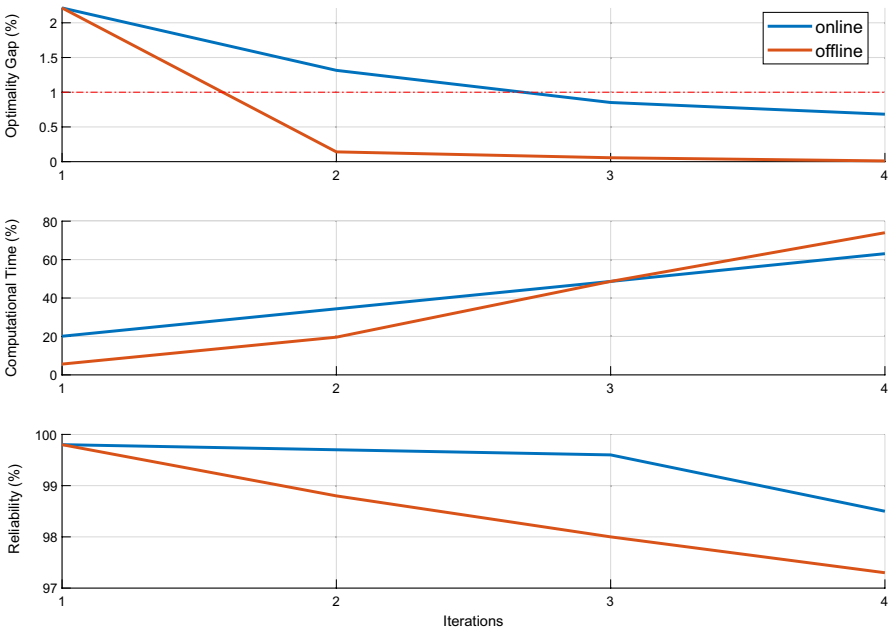

Fig. 7 Online vs Offline (300/DS1): optimality gap, computational time, and reliability of the solutions from the early iterations. The red dashed line marks the 1% optimality gap (color figure online)

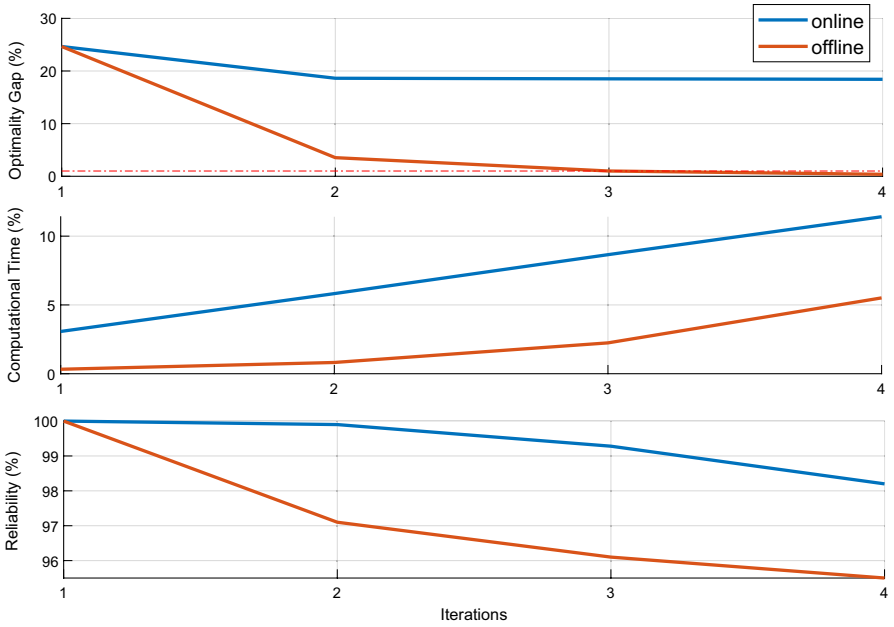

Fig. 8 Online vs Offline (118/DS2): optimality gap, computational time, and reliability of the solutions from the early iterations. The red dashed line marks the 1% optimality gap (color figure online)

6 Conclusions

In this paper, we analyzed the value of including unimodality information in DR-OPF. Specifically, we first showed how to integrate unimodality into the formulation and obtain an exact reformulation/sandwich approximation with SOC constraints. Second, we described an existing online approach and developed a new offline approach to solve DR-OPF with unimodality. For the offline approach, we use an OPS method to achieve a better sandwich approximation. Third, we demonstrated the benefit of including unimodality information in DR-OPF by evaluating our proposed approaches against the current state of the art. Through case studies on modified IEEE 118-bus and 300-bus systems, we demonstrated that including unimodality information within a DR-OPF problem with wind power uncertainty leads to a better cost/reliability trade-off than benchmark approaches or a DR-OPF that includes only moment information. However, the online solution approach suffers from large computational burden. We showed that our offline solution approach significantly reduces the computational time, providing fast convergence to low optimality gaps while also satisfying desired reliability levels. We also showed how the results vary across two forecast error data sets. We found that both the data set and choice of test system have a significant impact on the value of including unimodality information in DR-OPF, indicating that, in practice, the value is highly system-dependent. Moreover, the relative performance of the online versus offline

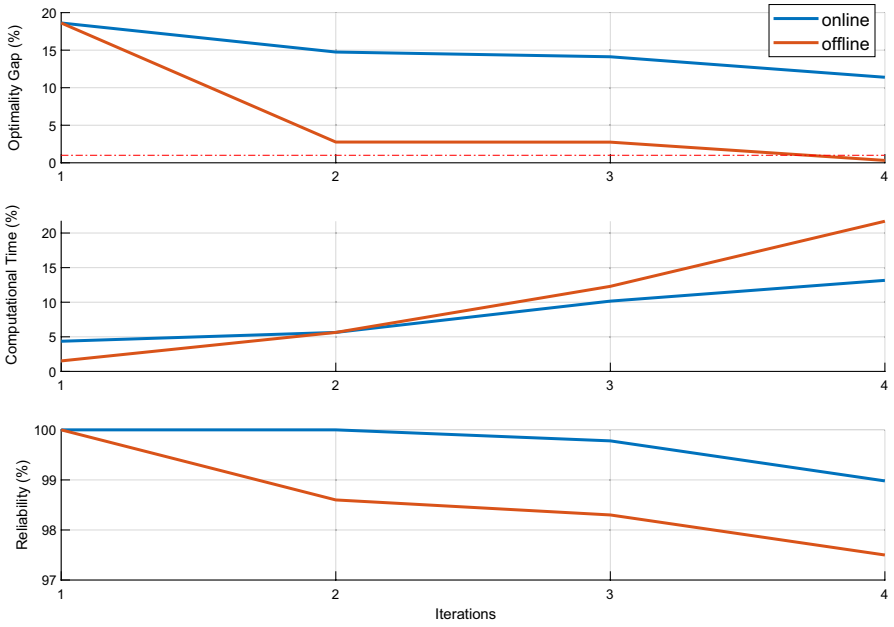

Fig. 9 Online vs Offline (300/DS2): optimality gap, computational time, and reliability of the solutions from the early iterations. The red dashed line marks the 1% optimality gap (color figure online)

algorithms, in terms of optimality gap, computational time, and solution reliability, is also system-dependent.

Acknowledgements This research was supported by the U.S. National Science Foundation Awards CCF-1442495 and CMMI-1662774.

References

Ahmed S, Shapiro A (2008) Solving chance-constrained stochastic programs via sampling and integer programming. In: State-of-the-art decision-making tools in the information-intensive age. INFORMS TutORials Oper Res:261–269

Arrigo A, Kazempour J, De Grève Z, Toubeau J.-F, Vallée F (2021) Embedding dependencies within distributionally robust optimization of modern power systems. arXiv preprint arXiv:2104.08101

Arrigo A, Ordoudis C, Kazempour J, De Grève Z, Toubeau J-F, Vallée F (2022) Wasserstein distributionally robust chance-constrained optimization for energy and reserve dispatch: an exact and physically-bounded formulation. Eur J Oper Res 296(1):304–322

Baker K, Bernstein A (2019) Joint chance constraints in ac optimal power flow: improving bounds through learning. IEEE Trans Smart Grid 10(6):6376–6385

Bienstock D, Chertkov M, Harnett S (2014) Chance-constrained optimal power flow: risk-aware network control under uncertainty. SIAM Rev 56(3):461–495

Campi M, Calafiore G, Prandini M (2009) The scenario approach for systems and control design. Annu Rev Control 33(2):149–157

Charnes A, Cooper W, Symonds G (1958) Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. Manage Sci 4(3):235–263

Coffrin C, Gordon D, Scott P (2016) NESTA, the NICTA energy system test case archive. arXiv preprint arXiv:1411.0359v5
Cox MG (1971) An algorithm for approximating convex functions by means by first degree splines. Comput J 14(3):272–275
Delage E, Ye Y (2010) Distributionally robust optimization under moment uncertainty with application to data-driven problems. Oper Res 58(3):595–612
Dharmadhikari SW, Joag-Dev K (1988) Unimodality, convexity, and applications. Academic Press, Cambridge
Duan C, Fang W, Jiang L, Yao L, Liu J (2018) Distributionally robust chance-constrained approximate AC-OPF with Wasserstein metric. IEEE Trans Power Syst 33(5):4924–4936
Esteban-Pérez A, Morales J.M (2021) Distributionally robust optimal power flow with contextual information. arXiv preprint arXiv:2109.07896
Gavrilović MM (1975) Optimal approximation of convex curves by functions which are piecewise linear. J Math Anal Appl 52(2):260–282
Ghaoui LE, Oks M, Oustry F (2003) Worst-case value-at-risk and robust portfolio optimization: a conic programming approach. Oper Res 51(4):543–556
Grant M, Boyd S (2008) Graph implementations for nonsmooth convex programs. In: Recent Advances in Learning and Control. Lecture Notes in Control and Information Sciences
Grant M, Boyd S (2014) CVX: matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx
Guo Y, Baker K, Dall'Anese E, Hu Z, Summers TH (2018) Data-based distributionally robust stochastic optimal power flow - Part II: case studies. IEEE Trans Power Syst 34(2):1493–1503
Guo Y, Baker K, Dall'Anese E, Hu Z, Summers T (2018) Stochastic optimal power flow based on data-driven distributionally robust optimization. In: IEEE Annual American Control Conference, Milwaukee, WI
Guo Y, Baker K, Dall'Anese E, Hu Z, Summers TH (2019) Data-based distributionally robust stochastic optimal power flow - Part I: methodologies. IEEE Trans Power Syst 34(2):1483–1492
Hou A.M, Roald L.A (2020) Chance constraint tuning for optimal power flow. In: 2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), pp. 1–6. IEEE
Huang W, Zheng W, Hill DJ (2021) Distributionally robust optimal power flow in multi-microgrids with decomposition and guaranteed convergence. IEEE Trans Smart Grid 12(1):43–55
Imamoto A, Tang B (2008) A recursive descent algorithm for finding the optimal minimax piecewise linear approximation of convex functions. In: World Congress on Engineering and Computer Science, San Francisco, CA
Jabr RA (2013) Adjustable robust OPF with renewable energy sources. IEEE Trans Power Syst 28(4):4742–4751
Jabr RA (2020) Distributionally robust cvar constraints for power flow optimization. IEEE Trans Power Syst 35(5):3764–3773
Jensen TV, Pinson P (2017) Re-Europe, a large-scale dataset for modeling a highly renewable European electricity system. Sci data 4:170–175
Jiang R, Guan Y (2016) Data-driven chance constrained stochastic program. Math Program 158(1):291–327
Li B (2019) Distributionally robust optimal power flow with strengthened ambiguity sets. PhD thesis, University of Michigan Ann Arbor
Li B, Jiang R, Mathieu J.L (2016) Distributionally robust risk-constrained optimal power flow using moment and unimodality information. In: IEEE Conference on Decision and Control, Las Vegas, NV
Li B, Mathieu J.L, Jiang R (2018) Distributionally robust chance constrained optimal power flow assuming log-concave distributions. In: Power Systems Computation Conference, Dublin, Ireland
Li B, Vrakopoulou M, Mathieu JL (2019a) Chance constrained reserve scheduling using uncertain controllable loads Part II: analytical reformulation. IEEE Trans Smart Grid 10(2):1618–1625
Li B, Jiang R, Mathieu JL (2019b) Ambiguous risk constraints with moment and unimodality information. Math Program 173(1–2):151–192
Li B, Jiang R, Mathieu JL (2019c) Distributionally robust chance-constrained optimal power flow assuming unimodal distributions with misspecified modes. IEEE Trans Control Netw Syst 6(3):1223–1234
Lu X, Chan KW, Xia S, Zhou B, Luo X (2019) Security-constrained multi-period economic dispatch with renewable energy utilizing distributionally robust optimization. IEEE Trans Sustain Energy 10(2):768–779

Lubin M, Dvorkin Y, Backhaus S (2016) A robust approach to chance constrained optimal power flow with renewable generation. IEEE Trans Power Syst 31(5):3840–3849
Margellos K, Goulart P, Lygeros J (2014) On the road between robust optimization and the scenario approach for chance constrained optimization problems. IEEE Trans Autom Control 59(8):2258–2263
Mieth R, Dvorkin Y (2018) Data-driven distributionally robust optimal power flow for distribution systems. IEEE Control Syst Lett 2(3):363–368
Miller B, Wagner H (1965) Chance constrained programming with joint constraints. Oper Res 13(6):930–945
Ordoudis C, Nguyen VA, Kuhn D, Pinson P (2021) Energy and reserve dispatch with distributionally robust joint chance constraints. Oper Res Lett 49(3):291–299
Pagnoncelli BK, Ahmed S, Shapiro A (2009) Sample average approximation method for chance constrained programming: theory and applications. J Optim Theory Appl 142(2):399–416
Papaefthymiou G, Klockl B (2008) MCMC for wind power simulation. IEEE Trans Energy Conv 23(1):234–240
Pourahmadi F, Kazempour J (2021) Distributionally robust generation expansion planning with unimodality and risk constraints. IEEE Trans Power Syst 36(5):4281–4295
Roald L, Oldewurtel F, Krause T, Andersson G (2013) Analytical reformulation of security constrained optimal power flow with probabilistic constraints. In: IEEE PowerTech Conference, Grenoble, France
Roald L, Oldewurtel F, Parys B.V, Andersson G (2015) Security constrained optimal power flow with distributionally robust chance constraints. arXiv preprint arXiv:1508.06061
Stellato B (2014) Data-driven Chance Constrained Optimization. Master thesis, ETH Zurich
Summers T, Warrington J, Morari M, Lygeros J (2015) Stochastic optimal power flow based on conditional value at risk and distributional robustness. Int J Elect Power Energy Syst 72:116–125
Tong X, Sun H, Luo X, Zheng Q (2018) Distributionally robust chance constrained optimization for economic dispatch in renewable energy integrated systems. J Global Optim 70(1):131–158
Vandewalle J (1975) On the calculation of the piecewise linear approximation to a discrete function. IEEE Trans Comput 24:843–846
Vrakopoulou M, Margellos K, Lygeros J, Andersson G (2013) 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
Vrakopoulou M, Li B, Mathieu JL (2019) Chance constrained reserve scheduling using uncertain controllable loads Part I: formulation and scenario-based analysis. IEEE Trans Smart Grid 10(2):1608–1617
Wagner M (2008) Stochastic 0–1 linear programming under limited distributional information. Oper Res Lett 36(2):150–156
Wang C, Gao R, Qiu F, Wang J, Xin L (2018) Risk-based distributionally robust optimal power flow with dynamic line rating. IEEE Trans Power Syst 33(6):6074-6086
Xie W, Ahmed S (2018) Distributionally robust chance constrained optimal power flow with renewables: a conic reformulation. IEEE Trans Power Syst 33(2):1860–1867
Xie W, Ahmed S, Jiang R (2019) Optimized bonferroni approximations of distributionally robust joint chance constraints. Math. Program: 1–34
Zhang H, Li P (2011) Chance constrained programming for optimal power flow under uncertainty. IEEE Trans Power Syst 26(4):2417–2424
Zhang Y, Shen S, Mathieu JL (2017) Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads. IEEE Trans Power Syst 32(2):1378–1388

Authors and Affiliations

Bowen Li[1] · Ruiwei Jiang[2] · Johanna L. Mathieu[3]

Bowen Li
bowen.li@anl.gov

Ruiwei Jiang
ruiwei@umich.edu

[1] Argonne National Laboratory, Lemont, IL, USA

[2] Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI, USA

[3] Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA