



Enhanced Wasserstein Distributionally Robust OPF With Dependence Structure and Support Information

Arrigo, Adriano; Kazempour, Jalal; De Greve, Zacharie; Toubeau, Jean-François; Vallée, Francois

Published in:
Proceedings of 2021 IEEE PowerTech

Link to article, DOI:
[10.1109/PowerTech46648.2021.9494771](https://doi.org/10.1109/PowerTech46648.2021.9494771)

Publication date:
2021

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Arrigo, A., Kazempour, J., De Greve, Z., Toubeau, J-F., & Vallée, F. (2021). Enhanced Wasserstein Distributionally Robust OPF With Dependence Structure and Support Information. In *Proceedings of 2021 IEEE PowerTech* IEEE. <https://doi.org/10.1109/PowerTech46648.2021.9494771>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Enhanced Wasserstein Distributionally Robust OPF With Dependence Structure and Support Information

Adriano Arrigo¹, Jalal Kazempour², Zacharie De Grève¹, Jean-François Toubeau¹ and François Vallée¹

¹ Power Systems and Market Research Group, University of Mons, Mons, Belgium

{adriano.arrigo, zacharie.degreve, jean-francois.toubeau, francois.vallee}@umons.ac.be

² Department of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark, seykaz@elektro.dtu.dk

Abstract—This paper goes beyond the current state of the art related to Wasserstein distributionally robust optimal power flow problems, by adding dependence structure (correlation) and support information. In view of the space-time dependencies pertaining to the stochastic renewable power generation uncertainty, we apply a moment-metric-based distributionally robust optimization, which includes a constraint on the second-order moment of uncertainty. Aiming at further excluding unrealistic probability distributions from our proposed decision-making model, we enhance it by adding support information. We reformulate our proposed model, resulting in a semi-definite program, and show its satisfactory performance in terms of the operational results achieved and the computational time.

Index Terms—Distributionally robust optimization, space-time dependencies, optimal power flow, out-of-sample analysis.

I. INTRODUCTION

The power system operators across the world are still using a deterministic variant of an optimal power flow (OPF) problem to determine the energy and reserve schedule of generators in a day-ahead time stage, i.e., 24-36 hours ahead of the real-time operation. This implies that the system operators naively rely on a single-point forecast value of renewable power generation in the day-ahead stage to schedule the power system while resorting to thumb rules to set the required amount of reserve capacity. The recent literature suggests incorporating the probabilistic forecast information into the OPF problem in the form of, e.g., intervals, scenarios, uncertainty set, or probability distribution functions [1]. The key point is that the uncertainty stemming from renewable power generation exhibits space-time dependencies, and it is in general difficult to represent this uncertainty by a unique distribution function. For the sake of the most representative and realistic way to incorporate the uncertainty into the OPF problem, we apply a distributionally robust optimization framework which hedges against any misrepresentation of uncertainty and space-time dependencies.

Distributionally Robust Optimization (DRO) has recently emerged as a new paradigm for modeling uncertainty by considering a set of potential distributions, the so-called *ambiguity set* [2]. The rationale behind this paradigm is that the exact true distribution function describing the underlying uncertainty is almost never known [3]. Distributionally robust OPF problem minimizes the total operational cost of the power system in expectation under the worst-case distribution within the ambiguity set. It is worth noting that the worst-case distribution is not known a priori, and is endogenously determined. In the

current technical literature on DRO, two different methodologies have been mainly suggested to construct the ambiguity set, resulting in *metric-based* [4] and *moment-based* [5] DRO frameworks. In the metric-based DRO, the ambiguity set is composed of all distributions in the neighborhood (in the sense of a selected probability distance metric) of a central empirical distribution function, e.g., based on a number of historical observations [6]. In contrast, the moment-based DRO uses moment information, e.g. mean and covariance, to build the ambiguity set. The appealing benefit of the metric-based DRO is that it could perform in a similar way as a scenario-based stochastic program or a robust optimization by properly tuning the size of the ambiguity set, and even outperform them in some cases. However, it is known to be computationally more intensive than the moment-based counterpart, although its computational burden is still comparable to the scenario-based stochastic programming.

Both moment- and metric-based distributionally robust OPF problems may consider unrealistic distributions within the ambiguity set to describe the renewable power generation uncertainty, which would lead to making erroneous operational decisions. Therefore, it is crucial to exclude those unrealistic distributions from the ambiguity set. This challenge can be leveraged by incorporating additional information about the renewable power generation uncertainty, such as *support information* [7] as well as *dependence structure* [8], into the design of the ambiguity set. The support information imposes the lower and upper bounds for the renewable power generation. In addition, the potential spatio-temporal correlation of renewable power generation [9] can be added to the ambiguity set, making it aware of the dependence of random variables. There are only very few works in the literature of distributionally robust OPF problem that include additional information about renewable uncertainty, see Table I for a survey. In particular, [10] introduces a moment-metric-based ambiguity set for the distributionally robust OPF problem, which combines the benefits of the metric-based DRO with a constraint imposing the value of the second-order moment. The resulting ambiguity set therefore includes all distributions in the neighborhood of a central empirical distribution with a fixed covariance matrix. By doing so, the ambiguity set shrinks to distributions with a dependence structure defined by the given covariance matrix.

Our contribution in this paper is to develop an enhanced moment-metric-based distributionally robust OPF model by incorporating both support information and dependence

TABLE I: Relevant features of works reported in the literature and the model proposed in this paper

Reference	Type of ambiguity set	Dependence structure	Unimodality information	Support information	Resulting model
[11]	Moment	×	×	×	LP ¹
[12]	Moment	×	×	×	QCQP ²
[13]	Moment	×	✓	×	SOCp ³
[14]	Metric	×	✓	×	SOCp
[4], [15]	Metric	×	×	✓	LP
[8], [10]	Moment-metric	✓	×	×	SDP ⁴
This paper	Moment-metric	✓	×	✓	SDP

¹ LP: Linear program ² QCQP: Quadratically-constrained quadratic program
³ SOCP: Second-order cone program ⁴ SDP: Semi-definite program

structure of random variables. To the best of our knowledge, as highlighted in Table I, this is the first research work that develops this improved variant of the distributionally robust OPF problem. We use a Wasserstein function as our probability distance measure. Our resulting model is a semi-definite program (SDP). We provide a thorough numerical analysis to illustrate the benefits of our proposed enhanced distributionally robust OPF model.

The remainder of this paper is structured as follows. Section II presents the distributionally robust OPF model and provides three different types of ambiguity sets, including our proposed set which is aware of support information and dependence structure. Section III reformulates the model. Section IV provides a numerical study. Section V concludes the paper. Finally, a mathematical proof is provided in the appendix.

II. MODEL

Vectors and matrices are represented by bold lower cases and by bold upper cases, respectively. Vector $\mathbf{1}$ is a vector of ones. The set of real numbers is represented by \mathbb{R} . The operator $|\cdot|$ returns the cardinality of the underlying set, and $(\cdot)^\top$ is the transpose operator. Finally, $\mathbb{E}^\mathbb{Q}[\cdot]$ refers to the expected value with respect to the probability distribution \mathbb{Q} .

A. Distributionally Robust OPF

The distributionally robust OPF problem to be solved in the day-ahead stage writes as

$$\min_{\mathbf{p}, \bar{\mathbf{r}}, \underline{\mathbf{r}}, \mathbf{B}} \mathbf{c}^\top \mathbf{p} + \bar{\mathbf{c}}^\top \bar{\mathbf{r}} + \underline{\mathbf{c}}^\top \underline{\mathbf{r}} + \max_{\mathbb{Q} \in \mathcal{A}} \mathbb{E}^\mathbb{Q} \left[\mathbf{c}^\top \mathbf{B} \tilde{\boldsymbol{\xi}} \right] \quad (1a)$$

$$\text{s.t. } \mathbf{p} + \bar{\mathbf{r}} \leq \mathbf{p}^{\max} \quad (1b)$$

$$\mathbf{p} - \underline{\mathbf{r}} \geq \mathbf{0} \quad (1c)$$

$$\mathbf{0} \leq \underline{\mathbf{r}} \leq \mathbf{r}^{\max}; \mathbf{0} \leq \bar{\mathbf{r}} \leq \mathbf{r}^{\max} \quad (1d)$$

$$\mathbf{1}^\top \mathbf{p} + \mathbf{1}^\top \mathbf{W} \boldsymbol{\mu} - \mathbf{1}^\top \mathbf{d} = 0 \quad (1e)$$

$$\mathbf{1}^\top \mathbf{B} + \mathbf{1}^\top \mathbf{W} = \mathbf{0} \quad (1f)$$

$$\min_{\mathbb{Q} \in \mathcal{A}} \mathbb{Q} \left(-r_g \leq \mathbf{B}_g \tilde{\boldsymbol{\xi}} \right) \geq 1 - \epsilon_g \quad \forall g \in \mathcal{G} \quad (1g)$$

$$\min_{\mathbb{Q} \in \mathcal{A}} \mathbb{Q} \left(\mathbf{B}_g \tilde{\boldsymbol{\xi}} \leq \bar{r}_g \right) \geq 1 - \bar{\epsilon}_g \quad \forall g \in \mathcal{G} \quad (1h)$$

$$\min_{\mathbb{Q} \in \mathcal{A}} \mathbb{Q} \left(\mathbf{Z}_l^\mathcal{G} \left(\mathbf{p} + \mathbf{B} \tilde{\boldsymbol{\xi}} \right) + \mathbf{Z}_l^\mathcal{W} \mathbf{W} \left(\boldsymbol{\mu} + \tilde{\boldsymbol{\xi}} \right) - \mathbf{Z}_l^\mathcal{D} \mathbf{d} \leq f_l^{\max} \right) \geq 1 - \epsilon_l \quad \forall l \in \mathcal{L}. \quad (1i)$$

Objective function (1a) minimizes the day-ahead scheduling cost of generators (the first three terms) as well as the recourse operational cost (the last term). The first term refers to the total production cost of the system, where $\mathbf{c} \in \mathbb{R}^{|\mathcal{G}|}$ is the vector of production cost of conventional generators, and $\mathbf{p} \in \mathbb{R}^{|\mathcal{G}|}$ is the vector of their production schedules. The second and

third terms in (1a) correspond to the cost for booking the capacity of conventional generators for upward and downward reserve services, respectively. Vectors $\bar{\mathbf{c}} \in \mathbb{R}^{|\mathcal{G}|}$ and $\underline{\mathbf{c}} \in \mathbb{R}^{|\mathcal{G}|}$ give the upward and downward reserve procurement costs, whereas vectors $\bar{\mathbf{r}} \in \mathbb{R}^{|\mathcal{G}|}$ and $\underline{\mathbf{r}} \in \mathbb{R}^{|\mathcal{G}|}$ refer to the upward and downward reserve schedules. Finally, the last term in (1a) provides the worst-case expectation of the recourse operational cost of the system. Vector $\tilde{\boldsymbol{\xi}} \in \mathbb{R}^{|\mathcal{W}|}$ is the forecast error of renewable energy sources, which is the sole source of uncertainty in our OPF problem, and its true probability distribution is not necessarily known. In addition, $\mathbf{B} \in \mathbb{R}^{|\mathcal{G}| \times |\mathcal{W}|}$ is a matrix of participation factors, representing the contribution of conventional generators towards balancing the production deviation of renewable energy sources in the real-time operation with respect to their day-ahead forecast. The recourse cost is calculated in expectation with respect to the worst-case probability distribution \mathbb{Q} , which is endogenously selected within the ambiguity set \mathcal{A} . This set will be described later in Section II-B.

Constraints (1b) to (1d) enforce capacity limits of conventional generators. Vector $\mathbf{p}^{\max} \in \mathbb{R}^{|\mathcal{G}|}$ refers to the capacity of conventional generators, whereas vector $\mathbf{r}^{\max} \in \mathbb{R}^{|\mathcal{G}|}$ corresponds to their maximum capability for upward/downward reserve provision. Constraints (1e) and (1f) impose power balance conditions in the day-ahead and real-time stages, respectively. Note that $\mathbf{W} \in \mathbb{R}^{|\mathcal{W}| \times |\mathcal{W}|}$ is a diagonal matrix containing the installed capacity of renewable energy sources, and $\boldsymbol{\mu} \in \mathbb{R}^{|\mathcal{W}|}$ is the vector of day-ahead renewable power generation forecast in per-unit, such that $\mathbf{1}^\top \mathbf{W} \boldsymbol{\mu}$ represents the total renewable power generation forecast in the day-ahead stage. According to (1e), the total production of conventional generators $\mathbf{1}^\top \mathbf{p}$ plus the total production of renewables $\mathbf{1}^\top \mathbf{W} \boldsymbol{\mu}$ supply the total inelastic demand $\mathbf{1}^\top \mathbf{d}$, where $\mathbf{d} \in \mathbb{R}^{|\mathcal{D}|}$ is the vector of their consumption levels. The real-time power balance constraint (1f) sets the participation factors \mathbf{B} , such that the recourse actions will adjust the power deviations of renewable energy sources for all $\tilde{\boldsymbol{\xi}} \in \mathbb{R}^{|\mathcal{W}|}$.

Finally, (1g)-(1i) are probabilistic constraints including random variables $\tilde{\boldsymbol{\xi}}$, modeled as individual distributionally robust chance constraints [16]. Each probabilistic constraint is assigned with a pre-defined constraint violation probability $\epsilon \in [0, 1]$. With a probabilistic confidence of $1 - \epsilon_g$ under the worst-case distribution \mathbb{Q} within the ambiguity set \mathcal{A} , the distributionally robust chance constraint (1g) ensures that the downward recourse action of conventional generator $g \in \mathcal{G}$ in the real-time operation, i.e., $\mathbf{B}_g \tilde{\boldsymbol{\xi}}$, will not exceed the downward reserve capacity r_g booked in the day-ahead stage. Note that the subscript in \mathbf{B}_g selects the row corresponding to generator g , and r_g is the g^{th} element of vector $\underline{\mathbf{r}}$. A similar constraint for the upward reserve is enforced by (1h). Finally, (1i) imposes the capacity of each transmission line $l \in \mathcal{L}$. Matrices $\mathbf{Z}^\mathcal{G} \in \mathbb{R}^{|\mathcal{L}| \times |\mathcal{G}|}$, $\mathbf{Z}^\mathcal{W} \in \mathbb{R}^{|\mathcal{L}| \times |\mathcal{W}|}$ and $\mathbf{Z}^\mathcal{D} \in \mathbb{R}^{|\mathcal{L}| \times |\mathcal{D}|}$ represent the power transmission distribution factors for conventional generators, renewable energy sources and demands, respectively. Subscript $\mathbf{Z}_l^{(\cdot)}$ selects the row corresponding to line l , and f_l^{\max} gives its capacity.

B. Uncertainty Modeling

We define three different types of ambiguity sets that will be used throughout this paper. For the ease of reading, we denote them as *metric*-based, *moment-metric*-based and *enhanced moment-metric*-based ambiguity sets, and define them as follow:

1) *Metric-Based Ambiguity Set*: It collects all distributions in the neighborhood of an empirical distribution function, say $\hat{\mathbb{Q}}_N = \frac{1}{N} \sum_{i=1}^N \delta_{\hat{\xi}_i}$, where $\delta_{\hat{\xi}_i}$ is the Dirac distribution centered on $\hat{\xi}_i$. The latter is therefore constructed based on N historical observations $\hat{\xi}_i$ which are assigned with an equal probability of $\frac{1}{N}$. To measure the probabilistic distance between distributions, we use the Wasserstein metric [17], for which we provide a mathematical definition in the online companion of this paper [18]. The Wasserstein distance corresponds to the optimal transportation cost of a transportation problem which aims at optimally transferring the probability from one to another distribution function. We define the metric-based ambiguity set \mathcal{A}_1 as

$$\mathcal{A}_1 = \left\{ \mathbb{Q} \in \mathcal{M} \mid d_W(\mathbb{Q}, \hat{\mathbb{Q}}_N) \leq \rho \right\}, \quad (2)$$

where \mathcal{M} denotes the set of all distributions on $\mathbb{R}^{|\mathcal{W}|}$. In addition, $\rho \in \mathbb{R}$ corresponds to the maximum transportation budget, the so-called *radius*, which limits the Wasserstein distance $d_W(\mathbb{Q}, \hat{\mathbb{Q}}_N)$ between each distribution \mathbb{Q} within set \mathcal{A}_1 and the empirical distribution $\hat{\mathbb{Q}}_N$.

2) *Moment-Metric-Based Ambiguity Set*: This type of ambiguity set, as defined in [8], is described by

$$\mathcal{A}_2 = \left\{ \mathbb{Q} \in \mathcal{M} \cap \mathcal{A}_1 \mid \mathbb{E}^{\mathbb{Q}} \left[(\tilde{\xi} - \mu_0) (\tilde{\xi} - \mu_0)^\top \right] \preceq \Sigma \right\}, \quad (3)$$

which contains only those distributions within the previously defined set \mathcal{A}_1 that comply with the covariance matrix of the empirical distribution. The newly added constraint in (3) ensures that all distributions within \mathcal{A}_2 follow the empirical covariance matrix Σ . Note that μ_0 represents the empirical mean. Thereby, set \mathcal{A}_2 contains the distributions that are close to the empirical distribution $\hat{\mathbb{Q}}_N$ and satisfy the linear dependence structure expressed in Σ . This additional feature enables eliminating the unrealistic distributions from the ambiguity set \mathcal{A}_1 when the underlying uncertainty (here, the renewable power generation uncertainty) has a correlated structure.

3) *Enhanced Moment-Metric-Based Ambiguity Set*: As our methodological contribution, we propose an enhanced moment-metric-based ambiguity set described by

$$\mathcal{A}_3 = \left\{ \mathbb{Q} \in \mathcal{M} \cap \mathcal{A}_2 \mid \mathbb{Q}(\tilde{\xi} \in \mathcal{U}) = 1 \right\}, \quad (4)$$

which contains only those distributions within the previously defined set \mathcal{A}_2 that satisfy an additional constraint coming from the support information. This constraint restricts the probability distribution \mathbb{Q} to take values strictly and only within the support \mathcal{U} . The support of uncertainty may correspond to physical bounds of the uncertainty, e.g., minimum and maximum renewable power generation, or to probabilistic bounds,

e.g., quantile regression achieved by forecasting methods. By adding this support information, we restrict the distributions to take realistic values, thereby further eliminating unrealistic distributions from the ambiguity set. For instance, the distributions in sets \mathcal{A}_1 and \mathcal{A}_2 may take non-realistic values, e.g., negative renewable power generation, which will lead to sub-optimality or infeasibility of the resulting program for large values of ρ . This intuition is confirmed by our numerical analysis to be provided later in Section IV. To add the support information, we consider ellipsoidal support \mathcal{U} , where the random vector has to satisfy a constraint as

$$(\tilde{\xi} \quad 1) \mathbf{M} \begin{pmatrix} \tilde{\xi} \\ 1 \end{pmatrix} \leq 0, \text{ where } \mathbf{M} = \begin{pmatrix} \Sigma_0 & -\Sigma_0 \xi_0 \\ -\xi_0^\top \Sigma_0 & \xi_0^\top \Sigma_0 \xi_0 - 1 \end{pmatrix}. \quad (5)$$

Note that $\xi_0 \in \mathbb{R}^{|\mathcal{W}|}$ corresponds to the center of the ellipsoid, and $\Sigma_0 \in \mathbb{R}^{|\mathcal{W}| \times |\mathcal{W}|}$ parametrizes the span of the ellipsoid.

III. MODEL REFORMULATIONS

We provide the tractable reformulations of OPF problem (1), such that it can be solved by existing off-the-shelf commercial solvers. We need to get rid of the non-convex probability operator $\mathbb{Q}(\cdot)$ in (1g)-(1i). For this purpose, we use conditional-value-at-risk (CVaR) approximation to cast the distributionally robust chance constraints in the form of $\min_{\mathbb{Q} \in \mathcal{A}} \mathbb{Q}(\cdot \leq 0) \geq 1 - \epsilon$ into CVaR constraints $\max_{\mathbb{Q} \in \mathcal{A}} \mathbb{Q}\text{-CVaR}_\epsilon(\cdot) \leq 0$ [19]. The CVaR operator is defined as

$$\max_{\mathbb{Q} \in \mathcal{A}} \mathbb{Q}\text{-CVaR}_\epsilon(\cdot) = \min_{\tau \in \mathbb{R}} \tau + \frac{1}{\epsilon} \max_{\mathbb{Q} \in \mathcal{A}} \mathbb{E}^{\mathbb{Q}} [[\cdot - \tau]^+], \quad (6)$$

where $\tau \in \mathbb{R}$ is an auxiliary variable, and operator $[\cdot]^+ = \max(\cdot, 0)$. Using this definition, problem (1) now includes the worst-case expectation not only in the objective function but also in the constraints. Consider a generic worst-case expectation problem $\max_{\mathbb{Q} \in \mathcal{A}} \mathbb{E}^{\mathbb{Q}} [\mathbf{a}(\mathbf{x})^\top \xi + b(\mathbf{x})]$, where $\mathbf{x} \in \mathcal{X}$ gives the vector of decision variables and $\mathbf{a}^\top(\mathbf{x}) \in \mathbb{R}^{|\mathcal{W}|}$ and $b(\mathbf{x}) \in \mathbb{R}$ represent the decision-dependent coefficients. For this generic worst-case expectation problem under ambiguity sets \mathcal{A}_1 , \mathcal{A}_2 and \mathcal{A}_3 , we provide the equivalent tractable reformulations in the online companion [18], Section III-A and Section III-B of this paper, respectively.

A. Moment-Metric-Based Ambiguity Set

The worst-case expectation problem under the moment-metric-based ambiguity set \mathcal{A}_2 , as proposed in [8], writes as

$$\max_{\mathbb{Q} \in \mathcal{A}_2} \mathbb{E}^{\mathbb{Q}} [\mathbf{a}(\mathbf{x})^\top \xi + b(\mathbf{x})] = \quad (7a)$$

$$\begin{cases} \min_{\lambda, \sigma_i, \Lambda \succeq 0, \zeta_i} \lambda \rho + \langle \Lambda, \Sigma \rangle_{\mathbb{F}} + \frac{1}{N} \sum_{i=1}^N \sigma_i \\ \text{s.t. } \mathcal{F}_i \succeq 0 \quad \forall i \in \{1, \dots, N\} \\ \|\zeta_i\|_* \leq \lambda \quad \forall i \in \{1, \dots, N\}, \end{cases} \quad (7b)$$

where $\lambda \in \mathbb{R}$, $\sigma \in \mathbb{R}^N$, $\zeta_i \in \mathbb{R}^{|\mathcal{W}|}$ and $\Lambda \in \mathbb{R}^{|\mathcal{W}| \times |\mathcal{W}|}$ are auxiliary variables. Operator $\langle \Lambda, \Sigma \rangle_{\mathbb{F}} = \sum_{m=1}^{|\mathcal{W}|} \sum_{n=1}^{|\mathcal{W}|} \Lambda_{mn} \cdot \Sigma_{mn}$

refers to the Frobenius inner product between matrices $\mathbf{\Lambda}$ and $\mathbf{\Sigma}$. In addition, operator $\|\cdot\|_*$ refers to the dual norm, which is defined as $\|v\|_* = \max_{\|b\| \leq 1} v^\top b$. Problem (7) is an SDP, where a constraint involving the \succeq operator imposes that the matrix on the left-hand side is semi-definite positive. We define matrix $\mathcal{F}_i \in \mathbb{R}^{|\mathcal{W}|+1 \times |\mathcal{W}|+1}$ as

$$\mathcal{F}_i = \begin{bmatrix} \mathbf{\Lambda} & -\frac{1}{2}\mathbf{a}(\mathbf{x}) + \frac{1}{2}\boldsymbol{\zeta}_i - \mathbf{\Lambda}\boldsymbol{\mu}_0 \\ (-\frac{1}{2}\mathbf{a}(\mathbf{x}) + \frac{1}{2}\boldsymbol{\zeta}_i - \mathbf{\Lambda}\boldsymbol{\mu}_0)^\top & \sigma_i - \boldsymbol{\zeta}_i^\top \hat{\boldsymbol{\zeta}}_i \\ & -b(\mathbf{x}) + \boldsymbol{\mu}_0^\top \mathbf{\Lambda}\boldsymbol{\mu}_0 \end{bmatrix}. \quad (8)$$

The complete reformulation of the distributionally robust OPF problem (1) under ambiguity set \mathcal{A}_2 is available in our online companion [18].

B. Enhanced Moment-Metric-Based Ambiguity Set

As our main methodological contribution, this section provides the reformulation of the worst-case expectation problem under the enhanced moment-metric-based ambiguity set \mathcal{A}_3 .

Theorem 1. The worst-case expectation problem under ambiguity set \mathcal{A}_3 described by (4) with an ellipsoidal support of uncertainty \mathcal{U} given by (5) is equivalent to

$$\max_{\mathbb{Q} \in \mathcal{A}_2} \mathbb{E}^{\mathbb{Q}} \left[\mathbf{a}(\mathbf{x})^\top \boldsymbol{\xi} + b(\mathbf{x}) \right] = \quad (9a)$$

$$\begin{cases} \min_{\lambda, \sigma_i, \mathbf{\Lambda} \succeq 0, \boldsymbol{\zeta}_i, \beta_i \geq 0} \lambda \rho + \langle \mathbf{\Lambda}, \mathbf{\Sigma} \rangle_{\text{F}} + \frac{1}{N} \sum_{i=1}^N \sigma_i \\ \text{s.t. } \mathcal{F}_i \succeq -\beta_i \mathbf{M} \quad \forall i \in \{1, \dots, N\} \\ \|\boldsymbol{\zeta}_i\|_* \leq \lambda \quad \forall i \in \{1, \dots, N\}. \end{cases} \quad (9b)$$

Proof. See the appendix. \square

Using Theorem 1, we reformulate the distributionally robust OPF problem (1) under ambiguity set \mathcal{A}_3 . The resulting SDP is available in the online companion [18].

IV. NUMERICAL STUDY

In this section we provide a numerical study to illustrate the benefits of the distributionally robust OPF problem with the proposed ambiguity set \mathcal{A}_3 over existing sets \mathcal{A}_1 and \mathcal{A}_2 in the literature. First, we present our case study and input data in Section IV-A. Then, we provide a computational analysis in Section IV-B. Finally, we discuss the superiority of ambiguity set \mathcal{A}_3 in terms of the expected operational cost of the system in Section IV-C.

A. Case study and Input Data

We use an adapted version of the IEEE 24-node reliability test system [20], which is composed of 12 conventional generators, 2 wind farms, and 17 loads, all connected through a network with 34 transmission lines. We consider a single-hour time period, as inter-temporal constraints are excluded. The total conventional generating capacity is 2,362.5 MW, whereas the total installed capacity of wind farms is 1,600 MW. In addition, the total demand is 2,207 MW. Full details of network parameters are available in our online companion [18].

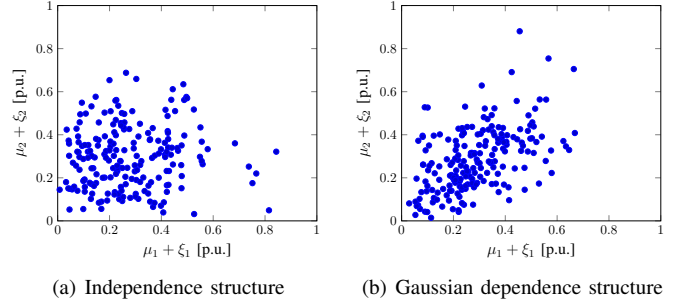


Fig. 1: Examples of wind power generation datasets for two wind farms with (a) independence and (b) dependence structures

Ambiguity sets \mathcal{A}_1 , \mathcal{A}_2 and \mathcal{A}_3 require the knowledge of empirical distribution of wind power deviations in the real-time operation with respect to the day-ahead forecast. In this paper, aiming at showing the benefits of the proposed uncertainty modeling technique under dependent uncertainty, we generate two different and synthetic datasets, namely *independent* and *dependent* sets. For this purpose, we use package *DatagenCopulaBased* v1.3.0 coded with Julia programming language v1.4.2. This package allows us to randomly draw samples from a copula, i.e., a distribution function with uniformly distributed marginals that embeds dependence information without any marginal information¹. We draw samples from the independent copula and the Gaussian copula parametrized with covariance matrix $\mathbf{\Sigma}_{\text{Gauss.}} = [1 \ 0.5; 0.5 \ 1]$. The uniformly distributed marginal samples are passed through the Weibull inverse cumulative distribution (scale parameter $\lambda = 1$ and shape parameter $k = 2$), resulting in two final sets of samples following a multivariate distribution function with a known copula (i.e., an independent or a dependent one) and known marginals. These two datasets are illustrated in Fig. 1.

We generate 1,050 samples from each distribution, 50 of which are in-sample data (depicted in Fig. 1) used for solving the proposed distributionally robust OPF problem under different ambiguity sets \mathcal{A}_1 , \mathcal{A}_2 and \mathcal{A}_3 . These 50 samples define the empirical distribution function. The remaining 1,000 samples are used later for an ex-post out-of-sample simulation. In addition, as the input data for the enhanced moment-metric-based ambiguity set \mathcal{A}_3 , we design an ellipsoidal support of uncertainty with parameters²

$$\mathbf{\Sigma}_0 = \begin{pmatrix} 2.20 & -0.25 \\ -0.25 & 2.20 \end{pmatrix} \text{ and } \boldsymbol{\xi}_0 = \begin{pmatrix} 0.2046 \\ 0.2046 \end{pmatrix}. \quad (10)$$

We consider an identical violation probability of $\epsilon = 0.05$ for all constraints (1g)-(1i). Finally, we take into account different values for radius ρ , ranging from 10^{-4} to 10^{-1} .

1. A copula allows to split the information on dependence structure and marginal distributions of a multivariate distribution function, i.e., any multivariate distribution function with k dimensions may be completely described by one copula and k marginals.

2. These parameters have been selected by trials and errors in the scope of this paper, with limited impact on the final results of our proposed model. A more accurate design of the support is left for the future work.

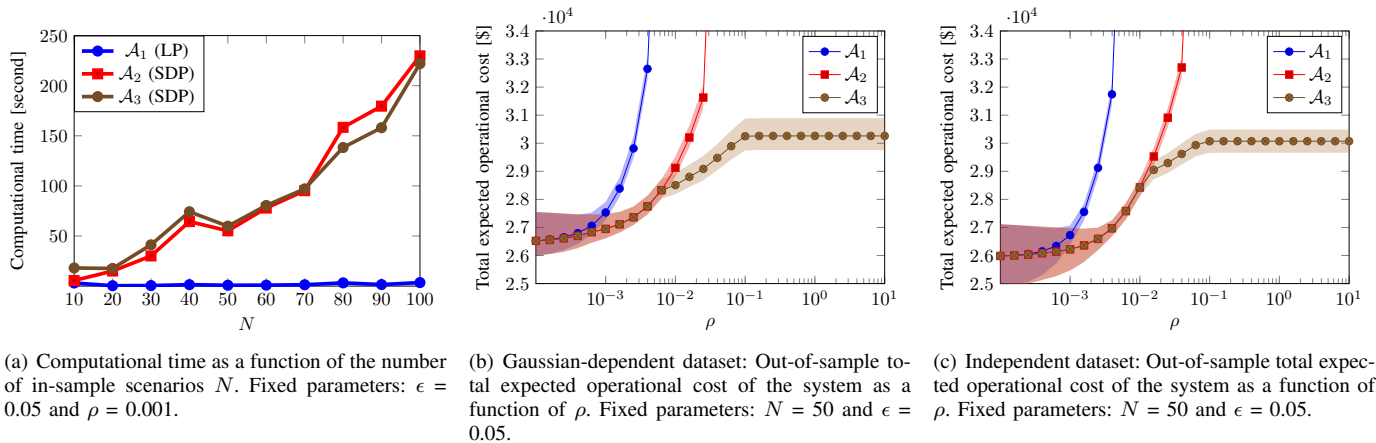


Fig. 2: Numerical study. The shaded area around each curve represents the corresponding standard deviation.

B. Computational Study

Recall that our resulting OPF model under ambiguity set \mathcal{A}_1 is a linear program (LP), while it is an SDP under sets \mathcal{A}_2 and \mathcal{A}_3 . We solve these programs using the solver Mosek 9.2 with the modeling language JuMP 0.21.3 and the programming language JuliaPro 1.4.2, on a computer clocking at 2.2 GHz with 16 GB installed RAM capacity. All source codes are publicly available in the online companion of this paper [18].

We retrieve the computational time corresponding to each model with different number of in-sample scenarios N , ranging from 10 to 100. The computational times are reported in Fig. 2(a). As expected, we observe that the computational time for the model with the metric-based ambiguity \mathcal{A}_1 is the lowest due to its linearity. In addition, these results show that the computational time related to the OPF model with the moment-metric-based and the enhanced moment-metric-based ambiguity sets are comparable. The reason for this is that the scale of problems (7) and (9) under ambiguity sets \mathcal{A}_2 and \mathcal{A}_3 are almost equivalent — only one supplementary scalar variable β_i appears in (9) compared to (7). These results further suggest that the computational time in both SDP models grows linearly when N increases.

C. Expected Operational Cost Based on Unseen Realizations

We use the remaining 1,000 samples of each dataset (independent and dependent) as unseen realizations of uncertainty that mimic the wind power generation in real-time. We first carry out the in-sample simulation by solving problem (1) with 50 samples. Given the optimal day-ahead schedule of conventional generators obtained, i.e., \mathbf{p} , $\bar{\mathbf{r}}$, and $\underline{\mathbf{r}}$, we run a real-time optimization problem 1,000 times, one per each unseen realization. This real-time optimization problem is given in the online appendix [18]. By solving this problem, we compute the cost in the real-time operation for the recourse actions (including involuntarily load shedding and wind spillage). We calculate the expected operational cost of the system and its standard deviation over 1,000 simulations, and report the results in Fig. 2(b) and 2(c), respectively, for the independent and Gaussian-dependent cases.

From Fig. 2(b), we observe that the system operator may achieve a lower expected operational cost when using the moment-metric-based ambiguity set \mathcal{A}_2 , compared to \mathcal{A}_1 . This interesting result stems from the smart design of \mathcal{A}_2 that imposes the dependence structure of probability distributions within the ambiguity set. Nevertheless, we observe that after a threshold value of radius ρ , i.e., 0.001 for \mathcal{A}_1 and 0.02 for \mathcal{A}_2 , the expected operational cost drastically increases. The program even fails to be feasible after a threshold value of 0.005 for \mathcal{A}_1 and 0.05 for \mathcal{A}_2 . The reason for this is that the worst-case distribution within the ambiguity set is not bounded by the support of uncertainty. For example, it may give a non-zero probability to a negative value of wind power generation. This may lead to infeasibility when the amount of available reserve is not sufficient for the system under the worst-case distribution. Our proposed technique relying on the enhanced moment-metric-based ambiguity set \mathcal{A}_3 that embeds the support information allows to recover feasibility when ρ takes a comparatively high value. By bounding distributions within the ambiguity set to lie within the support, the values taken by the worst-case distribution remain realistic. Eventually, our proposed technique achieves a lower expected operational cost when ρ is greater than 0.01, compared to a moment-metric-based approach. In each of the three explored models, the standard deviation decreases when the radius ρ increases.

It is worth noting that we observe similar results in Fig. 2(c) for the independent case, although with different threshold values. This counter-intuitive result shows that irrespective of the dependence structure of the underlying uncertainty, the ambiguity set \mathcal{A}_3 outperforms others. Eventually, our proposed model allows to reduce the expected operational cost when ρ increases. This happens by eliminating the unrealistic distributions from the ambiguity set while keeping a smart insight of the dependence structure of the underlying distributions, resulting in the minimum expected operational cost achieved for all values of radius ρ .

V. CONCLUSION

This paper explores distributionally robust optimization frameworks for the day-ahead optimal power flow problem. We

propose an enhanced moment-metric-based ambiguity set that accounts for the dependence among the uncertain variable as well as the support of uncertainty to make informed scheduling decisions. We provide the mathematical reformulations related to this new type of ambiguity set and apply it within a numerical study. We show the superiority of our proposed technique in terms of the out-of-sample expected operational cost of the system compared to the existing ambiguity sets, i.e., metric-based and moment-metric-based ones.

As a potential future research path, it is of interest to explore the implementation of ambiguity sets accounting for nonlinear correlations, e.g., based on copula information. Furthermore, the design of the ellipsoidal support of uncertainty could be fine-tuned by machine learning techniques with the aim of including more realistic distributions within the ambiguity set.

APPENDIX: PROOF OF THEOREM 1

We depart from Theorem 1 in [8], which is equivalently stated for discrete empirical distributions as

$$\max_{\mathbb{Q} \in \mathcal{A}_3} \mathbb{E}^{\mathbb{Q}} [\mathbf{a}(\mathbf{x})^\top \boldsymbol{\xi} + b(\mathbf{x})] = \quad (11a)$$

$$\begin{cases} \min_{\lambda, \sigma_i, \Lambda \succeq 0} \lambda \rho + \langle \Lambda, \Sigma \rangle_F + \frac{1}{N} \sum_{i=1}^N \sigma_i \\ \text{s.t. } \max_{\boldsymbol{\xi} \in \mathcal{U}} \mathbf{a}(\mathbf{x})^\top \boldsymbol{\xi} + b(\mathbf{x}) - \lambda \|\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}_i\| \\ \quad - (\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\top \Lambda (\boldsymbol{\xi} - \boldsymbol{\mu}_0) \leq \sigma_i \quad \forall i \in \{1, \dots, N\}. \end{cases} \quad (11b)$$

Constraint in (11b) requires further reformulations. First, we use the dual norm to recast the norm $\|\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}_i\|$ as

$$\begin{aligned} \max_{\boldsymbol{\xi} \in \mathcal{U}} \mathbf{a}(\mathbf{x})^\top \boldsymbol{\xi} + b(\mathbf{x}) - \lambda \max_{\|\zeta_i\|_* \leq 1} \zeta_i^\top (\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}_i) \\ - (\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\top \Lambda (\boldsymbol{\xi} - \boldsymbol{\mu}_0) \leq \sigma_i \quad \forall i \in \{1, \dots, N\}. \end{aligned} \quad (12)$$

The inner maximization operator is brought back to the left with a change in the optimization sense (i.e., $-\max = \min -$). Next, we permute the order of the optimization operators to result in a min-max structure; the permutation is allowed since the optimization sets are convex and independent. For the sake of simplicity, we realize a change of variable $\lambda \zeta_i \mapsto \zeta_i$ and drop the min operator inside the “less or equal to” constraint, adding the variable ζ_i to the decision variable, which allows us to equivalently reformulate the constraint (12) as

$$\max_{\boldsymbol{\xi} \in \mathcal{U}} \mathbf{a}(\mathbf{x})^\top \boldsymbol{\xi} + b(\mathbf{x}) - \zeta_i^\top (\boldsymbol{\xi} - \hat{\boldsymbol{\xi}}_i) \\ - (\boldsymbol{\xi} - \boldsymbol{\mu}_0)^\top \Lambda (\boldsymbol{\xi} - \boldsymbol{\mu}_0) \leq \sigma_i \quad \forall i \in \{1, \dots, N\} \quad (13a)$$

$$\|\zeta_i\|_* \leq \lambda. \quad (13b)$$

Constraint (13a) corresponds to a quadratically-constrained quadratic program (QCQP) when the uncertainty set \mathcal{U} is ellipsoidal and described by quadratic equations (5). By means of S-Lemma [21], this QCQP can be equivalently reformulated as SDP

$$\mathcal{F}_i \succeq -\beta_i \mathbf{M} \quad \forall i \in \{1, \dots, N\} \quad (14a)$$

$$\|\zeta_i\|_* \leq \lambda \quad \forall i \in \{1, \dots, N\} \quad (14b)$$

$$\beta_i \geq 0 \quad \forall i \in \{1, \dots, N\}, \quad (14c)$$

where $\beta_i \in \mathbb{R}$ is an auxiliary variable and \mathcal{F}_i is defined by (8). Constraint (11b) can therefore be equivalently recast as (14). This completes the proof. \square

REFERENCES

- [1] D. Bienstock, M. Chertkov, and S. Harnett, “Chance-constrained optimal power flow: Risk-aware network control under uncertainty,” *SIAM Review*, vol. 56, no. 3, pp. 461–495, 2014.
- [2] H. Rahimian and S. Mehrotra, “Distributionally robust optimization: A review,” 2019. [Online]. Available: www.optimization-online.org/DB_HTML/2019/08/7332.html
- [3] J. Botticau, L. Hubert, Z. D. Grève, F. Vallée, and J.-F. Toubeau, “Very short-term probabilistic forecasting for a risk-aware participation in the single price imbalance settlement,” *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1218–1230, Mar. 2020.
- [4] P. Mohajerin Esfahani and D. Kuhn, “Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations,” *Math. Prog.*, vol. 171, no. 1-2, pp. 115–166, 2018.
- [5] E. Delage and Y. Ye, “Distributionally robust optimization under moment uncertainty with application to data-driven problems,” *Oper. Res.*, vol. 58, no. 3, pp. 595–612, 2010.
- [6] Y. Guo, K. Baker, E. Dall’Anese, Z. Hu, and T. H. Summers, “Data-based distributionally robust stochastic optimal power flow,” *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1483–1492, Mar. 2019.
- [7] A. Arrigo, C. Ordoúdis, J. Kazempour, Z. D. Grève, J.-F. Toubeau, and F. Vallée, “Wasserstein distributionally robust chance-constrained optimization for energy and reserve dispatch: An exact and physically-bounded formulation,” 2020, Under review.
- [8] R. Gao and A. J. Kleywegt, “Distributionally robust stochastic optimization with dependence structure,” 2017, arXiv:1701.04200.
- [9] F. Vallée, J. Lobry, and O. Deblecker, “Impact of the wind geographical correlation level for reliability studies,” *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2232–2239, Nov. 2007.
- [10] C. Wang, R. Gao, F. Qiu, J. Wang, and L. Xin, “Risk-based distributionally robust optimal power flow with dynamic line rating,” *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6074–6086, Nov. 2018.
- [11] C. Zhao and R. Jiang, “Distributionally robust contingency-constrained unit commitment,” *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 94–102, Jan. 2018.
- [12] R. Mieth and Y. Dvorkin, “Data-driven distributionally robust optimal power flow for distribution systems,” *IEEE Control Syst. Lett.*, vol. 2, no. 3, pp. 363–368, Jul. 2018.
- [13] Y. Zhang, S. Shen, and J. L. Mathieu, “Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads,” *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1378–1388, Mar. 2017.
- [14] A. Esteban-Pérez and J. M. Morales, “Data-driven distributionally robust optimization via optimal transport with order cone constraints,” 2019. [Online]. Available: <https://arxiv.org/abs/1903.01769>
- [15] R. Gao and A. Kleywegt, “Distributionally robust stochastic optimization with Wasserstein distance,” 2016, arXiv:1604.02199.
- [16] Z. Chen, D. Kuhn, and W. Wiesemann, “Data-driven chance constrained programs over Wasserstein balls,” 2018. [Online]. Available: http://www.optimization-online.org/DB_FILE/2018/06/6671.pdf
- [17] L. V. Kantorovich and G. S. Rubinshtein, “On a space of totally additive functions,” *Vestnik Leningradskogo Universiteta*, vol. 13, pp. 52–59, 1958.
- [18] A. Arrigo, J. Kazempour, Z. De Grève, J.-F. Toubeau, and F. Vallée, “Online companion: Enhanced Wasserstein distributionally robust OPF with dependence structure and support information,” 2020. [Online]. Available: <https://doi.org/10.5281/zenodo.4293635>
- [19] S. Zymler, D. Kuhn, and B. Rustem, “Distributionally robust joint chance constraints with second-order moment information,” *Math. Prog.*, vol. 137, no. 1-2, pp. 167–198, 2013.
- [20] Reliability System Task Force, “The IEEE reliability test system 1996. A report prepared by the reliability test system task force of the application of probability methods subcommittee,” *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1010–1020, Aug. 1999.
- [21] I. Pólik and T. Terlaky, “A survey of the S-Lemma,” *SIAM Review*, vol. 49, no. 3, pp. 371–418, 2007.