Embedding Dependencies Between Wind Farms in Distributionally Robust Optimal Power Flow

Adriano Arrigo, Jalal Kazempour, *Senior Member, IEEE,* Zacharie De Grève, *Member, IEEE,*
Jean-François Toubeau, *Member, IEEE,* and François Vallée, *Member, IEEE*

Abstract—**The increasing share of renewables in the electricity generation mix comes along with an increasing uncertainty in power supply. In the recent years, distributionally robust optimization has gained significant interest due to its ability to make informed decisions under uncertainty, which are robust to misrepresentations of the distributional information (e.g., from probabilistic forecasts). This is achieved by introducing an ambiguity set that describes the potential deviations from an empirical distribution of all uncertain parameters. However, this set typically overlooks the inherent dependencies of uncertainty, e.g., spatial dependencies of weather-dependent energy sources. This paper goes beyond the state-of-the-art models by embedding such dependencies within the definition of ambiguity set. In particular, we propose a new copula-based ambiguity set which is tailored to capture any type of dependencies. The resulting problem is reformulated as a conic program which is kept generic such that it can be applied to any decision-making problem under uncertainty in power systems. Given the Optimal Power Flow (OPF) problem as one of the main potential applications, we illustrate the performance of our proposed distributionally robust model applied to *i)* a DC-OPF problem for a meshed transmission system and *ii)* an AC-OPF problem using *LinDistFlow* approximation for a radial distribution system.**

Index Terms—**Distributionally robust optimization, copula, dependencies, optimal power flow.**

I. INTRODUCTION

THE continuous balance between power injections and offtakes is paramount to ensure the safe operation of power systems, however it is being challenged by the uncertainty introduced by the weather-dependent renewable power generation [1]. To cope with such an uncertainty, it is crucial to develop uncertainty-aware decision-making tools for operational and planning purposes. While the power system operators traditionally use a deterministic optimization approach by considering a single-point forecast of renewable power generation, stochastic approaches account for a probabilistic forecast, allowing to make more informed operational and planning decisions (depending on the forecast accuracy) [2].

The power system research community has put a lot of efforts in the last decades for developing various types of stochastic decision-making tools, ranging from scenario-based stochastic programming [3] to robust optimization [4] and

A. Arrigo, Z. De Grève and F. Vallée are with the Power Systems and Markets Research group at University of Mons, 31, Boulevard Dolez, 7000, Mons, Belgium. E-mails : {adriano.arrigo, zacharie.degreve, francois.vallee}@umons.ac.be.

J.-F. Toubeau is with the TME Branch, Energy Institute, University of Leuven (KU Leuven), 13, Oude Markt, 3000, Leuven, Belgium. E-mail: jean-francois.toubeau@kuleuven.be.

J. Kazempour is with the Department of Wind and Energy Systems at Technical University of Denmark, 325, Elektrovej, 2800 Kgs. Lyngby, Denmark. E-mail: jalal@dtu.dk.

chance-constrained programming [5]. However, these techniques disregard the inherent misrepresentation of the input probabilistic information, which may lead to high discrepancies between the expected results and the actual (ex-post) realizations [6]. To address this issue, we focus on Distributionally Robust Optimization (DRO) [7], which has shown the capability to outperform all previous techniques in presence of high uncertainties.

DRO allows to hedge against the inherent errors arising from modeling the probabilistic distribution that is typically derived from the probabilistic forecast. These errors may propagate to the resulting decisions and compromise the reliability and cost-efficiency of the power system. The exposure to uncertainty in the probability distribution itself is called *ambiguity* [8]. By design, DRO considers a family of potential distributions, the so-called ambiguity set, to hedge against any inexactness or biasedness of the distribution function. Owing to its appealing properties, one popular approach to define an ambiguity set is based on the Wasserstein probability metric, which calculates the distance between two distribution functions [9]. In the Wasserstein DRO approach, all distributions in the neighborhood of a central empirical distribution, based on historically observed samples of uncertainty, are collected and fed to the decision-making problem. Although this framework accounts for potential forecast errors, there is no guarantee that the distributions within the ambiguity set remain realistic (e.g., from a correlation perspective). As our main contribution, we go beyond the state of the art by developing a DRO approach for power systems that is aware of all potential *dependencies* among uncertain parameters, especially the spatial dependencies of weather-dependent renewable energy sources.

In the current literature, there is a willingness to enhance the Wasserstein ambiguity set by removing the unrealistic distributions of renewable power generation uncertainty. To do so, supplementary constraints should be incorporated into the design of the ambiguity set. References [9] and [10] include the support information in the definition of Wasserstein ambiguity set, excluding distributions with unrealistic realizations of uncertainty. An example of such unrealistic distributions is those with a negative renewable power generation. References [11] and [12] include modality information, i.e., the number of spikes in a probability distribution, within the ambiguity set to get rid of potentially unrealistic distributions, e.g., those with two or more spikes. References [13] and [14] embed information on dependencies, which is based on imposing the value of a covariance matrix in the ambiguity set definition. However, the covariance matrix is only able to capture *linear* relations and fails to capture more complex non-linear dependencies (e.g., stemming from wind power generation).

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2022.3221348

2

Moreover, the models derived in these works usually impose the value of covariance via a positive semi-definite matrix restriction which is not flexible, i.e., the distributions within the set must follow the dependence structure contained in an arbitrarily fixed semi-definite cone centered on the empirically observed covariance matrix. Overall, all these works allow to improve the representation of multi-dimensional uncertainties, however all have their own limitations and cannot represent the whole dependence structure.

Several additional works have contributed towards embedding the full non-linear dependence information into the distributionally robust optimization. However, it is worth emphasizing that there are currently very few contributions in this research strand, especially focusing on power system applications. References [15]-[16] envisage the use of *copula* as the mathematical object to describe the full dependence structure among random variables. A copula is a distribution function with uniformly distributed marginals that contains solely information about the dependence structure among uncertain parameters [17]-[18]. To incorporate the dependence structure into the optimization problem, [15]-[16] introduce a constraint binding the Wasserstein distance between the copula of the distributions inside the ambiguity set and the one of the empirical distribution. In all these works, the marginals are fixed and assumed to be known. In other words, the only source of uncertainty stems from the dependence structure, and the ambiguity around the empirical distribution is no longer accounted for. This uncertainty scheme is usually suited for portfolio management problems in finance applications, but is expected to be insufficient for power system applications, where the marginal distributions of each renewable energy source are usually not predicted with high accuracy.

To address this issue, [19] represents the first effort in the operations research literature towards embedding the full dependence information along with the ambiguity around the empirical distribution function. To do so, the authors combine the findings in [15]-[16] with the classical approach to DRO [9]. They develop a Wasserstein ambiguity set that uses copula as the mathematical object to describe the full dependence structure among renewable energy sources.

In this paper, inspired by [19], we propose a copula-based ambiguity set as the framework that *generalizes (i)* the constraint on the second-order moment [13]-[14], and, *(ii)* the model with known marginals [15]-[16] by enabling the DRO problem to consider *any* potential type of correlation (not necessarily linear) among uncertain parameters. By this generalization, we are able to properly model the renewable power generation uncertainty that may exhibit any shape of dependencies [20]. To the best of our knowledge, this is the first effort in the literature to introduce copula-based ambiguity sets for power system applications. In particular, our contributions are threefold:

(i) We first develop an optimization program in Lemma 1 which computes the value of an empirical cumulative distribution function. This allows us to link any distribution within the ambiguity set to its corresponding copula.
(ii) Building up on Lemma 1, we develop a conic reformulation of a distributionally robust worst-case expectation problem[1] in Theorem 1, using the proposed copula-based ambiguity set. Theorem 1 and the solution approach in Appendix C are kept generic such that they could be applied to a broad range of problems in power systems.
(iii) We apply the proposed conic formulation to a day-ahead distributionally robust Optimal Power Flow (OPF) problem with dependent renewable power generation uncertainty. We explore both DC and AC optimal power flow formulations and numerically show the benefits of the proposed reformulation in both cases compared to a state-of-the-art distributionally robust model, while identifying potential paths for future research.

The remainder of this paper is structured as follows. Sections II and III provide a definition for the traditional metric-based and the proposed copula-based ambiguity sets, respectively. Section IV explains the reformulation process of a generic worst-case expectation problem under the copula-based ambiguity set. Section V introduces the day-ahead distributionally robust OPF problem and details the required reformulations for both DC and AC formulations. Section VI discusses the proposed model through an extensive numerical analysis. Main conclusion and prospects are given in Section VII. Finally, the mathematical proofs of Lemma 1 and Theorem 1 are provided in Appendixes A and B, respectively, while the proposed solution approach for Theorem 1 is discussed in Appendix C.

II. THE TRADITIONAL METRIC-BASED AMBIGUITY SET

We start with the definition of an empirical distribution function. Consider N number of equiprobable historical observations of the uncertainty, i.e., $\widehat{\xi}_i$, where $i = \{1, 2, ..., N\}$. The empirical distribution $\widehat{\mathbb{Q}}_N$ is defined as

$$\widehat{\mathbb{Q}}_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\widehat{\xi}_i}, \qquad (1)$$

where $\delta_{\widehat{\xi}_i}$ represents the Dirac distribution centered on $\widehat{\xi}_i$ and is assigned with a probability equal to $\frac{1}{N}$. Note that every symbol with a hat, e.g., $\widehat{\mathbb{Q}}_N$ and $\widehat{\xi}_i$, corresponds to historical observations. Next, we define the Wasserstein probability metric. The metric $d_W \left(\mathbb{Q}, \widehat{\mathbb{Q}}_N \right) : \mathcal{A} \times \mathcal{A} \to \mathbb{R}^+$ computes the distributional distance between two distribution functions $\mathbb{Q}$ and $\widehat{\mathbb{Q}}_N$ as the optimal value of a transportation problem between the probability mass of those two distributions [22]. The complete definition of this metric is given in the online appendix [23]. Note that $\mathcal{A}$ defines the space of all distribution functions.

Given the definition of the Wasserstein probability metric, the metric-based ambiguity set $\mathcal{M}_1$ is defined as

$$\mathcal{M}_1 = \left\{ \mathbb{Q} \in \mathcal{A} \ \middle| \ d_W \left(\mathbb{Q}, \widehat{\mathbb{Q}}_N \right) \le \theta_1 \right\}, \qquad (2)$$

which contains a family of distributions $\mathbb{Q}$ in the space of all distribution functions $\mathcal{A}$ that are in the neighbourhood of the empirical distribution function $\widehat{\mathbb{Q}}_N$. The distributional

---

[1]The worst-case expectation problem is a well-known problem in the field of DRO [21]. Its generic reformulation usually facilitates the application of DRO to any kind of problem where it naturally appears.

distance $d_W\left(\mathbb{Q}, \widehat{\mathbb{Q}}_N\right)$ is limited to be lower than or equal to $\theta_1 \in \mathbb{R}^+$. This value is tuned by the decision-maker, and is referred to as the radius of the ambiguity set. One can intuitively interpret that a larger value of θ_1 yields an ambiguity set that contains more distributions, implying that the decision-maker is less certain about the true distribution of the uncertainty. The reader is referred to [9] for more details about the traditional Wasserstein ambiguty set. In Section III, we innovatively enhance the traditional Wasserstein metric-based ambiguity set with a copula model to prevent unrealistic uncertainty distributions within the ambiguity set.

III. THE PROPOSED COPULA-BASED AMBIGUITY SET

We first mathematically define copula $\mathbb{C}$ of the distribution $\mathbb{Q}$. Let us consider $|\mathcal{W}|$ number of renewable power units, e.g., wind farms. Let $\widetilde{\xi} \in \mathbb{R}^{|\mathcal{W}|}$ be linked to distribution $\mathbb{Q}$. Unlike symbols with a hat that refer to historical observations, those with a tilde, e.g., $\widetilde{\xi}$, correspond to uncertain parameters. The symbol $\widetilde{\xi}_k$ refers to the uncertain generation of the renewable power unit k. Mathematically speaking, the copula $\mathbb{C}$ of distribution $\mathbb{Q}$ is defined as the cumulative distribution function of the uncertain parameter $\widetilde{U}$ [17]-[18], i.e.,

$$\left(\widetilde{U}_1, ..., \widetilde{U}_k, ..., \widetilde{U}_{|\mathcal{W}|}\right)$$
$$= \left(F_1\left(\widetilde{\xi}_1\right), ..., F_k\left(\widetilde{\xi}_k\right), ..., F_{|\mathcal{W}|}\left(\widetilde{\xi}_{|\mathcal{W}|}\right)\right), \quad (3)$$

where $\widetilde{U} \in \mathbb{R}^{|\mathcal{W}|}$ is linked to distribution $\mathbb{C}$. In addition, the function $F_k\left(.\right) = \mathbb{Q}_k\left(\widetilde{\xi}_k \leq .\right)$ represents the cumulative distribution function of element $\widetilde{\xi}_k$ of the $|\mathcal{W}|$-dimensional uncertain vector $\widetilde{\xi}$. Note that the probability operator $\mathbb{Q}_k\left(.\right)$ represents the marginal distribution function of the random variable $\widetilde{\xi}_k$. Therefore, $F_k\left(\widetilde{\xi}_k\right)$ defines a random variable which is uniformly distributed on the interval $[0,1]$.

By (3), the resulting distribution of $\widetilde{U}$ has marginals that are uniformly distributed on the interval $[0,1]$ and has the property to embody the dependence between the components of $\widetilde{\xi}$. In other words, the information contained in any multivariate distribution function can be splitted into *(i)* a collection of marginal distributions (containing the univariate information) and, *(ii)* a copula (containing the dependence information) from which the univariate information has been filtered.

We now elaborate on the need for a copula-based ambiguity set. Whereas the metric-based ambiguity set $\mathcal{M}_1$ defined in (2) allows the decision-maker to incorporate useful information as much as possible into the distributionally robust program, it may still contain erroneous distributions, i.e., distributions that do not reflect with fidelity the potential outcome of the uncertainty. For instance, distributions within set $\mathcal{M}_1$ may exhibit a completely different dependence structure than the one observed empirically. Hence, the decisions may unnecessarily be optimized for an over-conservative and/or non-representative insight of uncertainty, resulting in a higher expected total operational cost. In order to avoid such a situation, we aim to eliminate those distributions from the ambiguity set. In that direction, this paper introduces the copula-based ambiguity set

$\mathcal{M}_2$ which generally contains more representative distribution functions, such that

$$\mathcal{M}_2 = \left\{ \mathbb{Q} \in \mathcal{A} \left| \begin{array}{c} d_W\left(\mathbb{Q}, \widehat{\mathbb{Q}}_N\right) \leq \theta_1 \\ d_W\left(\mathbb{C}, \widehat{\mathbb{C}}_N\right) \leq \theta_2 \end{array} \right. \right\}. \quad (4)$$

The first constraint in the ambiguity set $\mathcal{M}_2$ is identical to the one in $\mathcal{M}_1$, yielding the desirable properties of the classical definition of the ambiguity set. The newly added second constraint limits the distributional distance $d_W\left(\mathbb{C}, \widehat{\mathbb{C}}_N\right)$ between the endogenously selected copula $\mathbb{C}$ of the endogenously selected distribution $\mathbb{Q}$ and the empirical copula $\widehat{\mathbb{C}}_N$ of the distribution $\widehat{\mathbb{Q}}_N$. This distance should not be greater than $\theta_2 \in \mathbb{R}^+$. Again, θ_2 is a parameter to be tuned by the decision-maker. By restricting the distributions $\mathbb{Q}$ inside the ambiguity set to have a copula $\mathbb{C}$ in the neighbourhood of the empirical one, the distributions inside the ambiguity set will follow a dependence structure which remains close to the historically observed one[2]. Similar to θ_1, a greater value of θ_2 implies that the decision-maker is less confident about the true dependence structure of the uncertainty, and includes distributions whose dependencies are less similar to those of the empirical one within the ambiguity set.

IV. WORST-CASE EXPECTED PROBLEM UNDER THE COPULA-BASED AMBIGUITY SET

By design, DRO aims to determine the *worst-case* distribution within the given ambiguity set, and makes decisions in *expectation* with respect to such a worst-case distribution. This section derives reformulation for a generic distributionally robust worst-case expectation problem, using the copula-based ambiguity set $\mathcal{M}_2$. This problem writes as

$$\min_{x \in \mathcal{X}} \overbrace{\max_{\mathbb{Q} \in \mathcal{M}_2} \mathbb{E}^{\mathbb{Q}}\left[a\left(x\right)^\top \widetilde{\xi} + b\left(x\right)\right]}^{\text{Worst-case expected cost}}, \quad (5)$$

where $x \in \mathcal{X}$ is the vector of decision variables. The inner maximization operator in (5) picks the worst-case distribution $\mathbb{Q}$ in the ambiguity set $\mathcal{M}_2$. The probability distribution of the uncertain parameter $\widetilde{\xi} \in \mathbb{R}^{|\mathcal{W}|}$ is $\mathbb{Q}$. The objective function (5) is linear[3], comprising of the decision-dependent vector $a(x) \in \mathbb{R}^{|\mathcal{W}|}$ and the decision-dependent scalar $b(x) \in \mathbb{R}$.

One key step in deriving the reformulation of (5) is to establish an analytical link between the copula $\mathbb{C}$ and its distribution function $\mathbb{Q}$. This can be achieved by using (3), in which $\mathbb{Q}$ and $\mathbb{C}$ are linked through the marginal cumulative distribution functions $F_k(.)$.

Remark 1. Theoretically, the link between variable copula $\mathbb{C}$ and variable distribution $\mathbb{Q}$ is established via the marginal cumulative distribution functions $F_k^{\mathbb{Q}} \forall k$ of $\mathbb{Q}$. However, in

---

[2]Another potential methodology relies on fixing the value of the second-order moment of the distributions within the ambiguity set [13]-[14]. However, only the *linear* dependence structure will be captured by such a constraint. On the contrary, the copula-based approach allows to capture *any* kind of dependence structure (not necessarily linear) and therefore offers a more general framework.

[3]For the sake of simplicity, we assume linearity of the objective function. This assumption is aligned with the current practice of electricity markets with linear bids. An extension to a non-linear objective function is straightforward, but requires additional reformulations that are out of the scope of this paper.
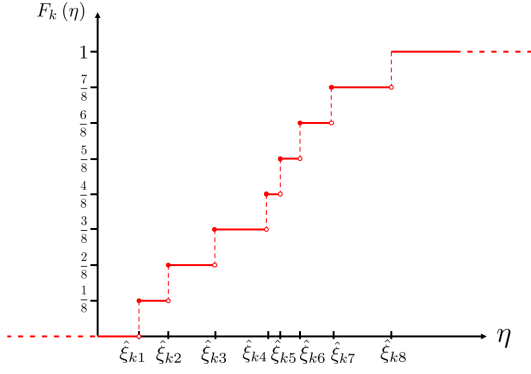
Fig. 1. Illustration of an empirical marginal cumulative distribution function for the renewable power unit k when the number of historical observations is $N = 8$ (an arbitrarily selected number).

practice, these functions are not straightforwardly accessible, due to their complex variable nature which impedes their endogenous reformulation. Consequently, considering these variable functions within the optimization framework would require research efforts that are beyond the scope of this paper. Therefore, we use the *empirical* marginal cumulative distribution functions $F_k^{\hat{\mathbb{Q}}_N} \forall k$ of $\hat{\mathbb{Q}}_N$. By doing so, we are able to derive a tractable reformulation of (5). Our hypothesis is that the approximation made by assessing the endogenous copula $\mathbb{C}$ via the functions $F_k^{\hat{\mathbb{Q}}_N}$ is required (to avoid non-linear formulation of functions $F_k^{\mathbb{Q}}$) and is valid in practice when $F_k^{\mathbb{Q}}$ and $F_k^{\hat{\mathbb{Q}}_N}$ are close to each other (e.g., when θ_1 is small enough). This setting is typically suited for a day-ahead probabilistic forecast embedded in OPF problems, as further demonstrated in our numerical analysis in Section VI. In the following, functions $F_k^{\hat{\mathbb{Q}}_N}$ are denoted F_k for the ease of notation.

From now on, we denote by $\eta \in \mathbb{R}$, the value of the argument of the function, i.e., $F_k(\eta)$. Let us further clarify this function by a schematic illustration. Fig. 1 shows the shape of the function $F_k(\eta)$ for the renewable power unit k, given the arbitrarily selected eight equiprobable historical observations $\hat{\xi}_{ki}, \ i \in \{1, ..., N = 8\}$ of $\tilde{\xi}_k$. In particular, this figure shows how the historical observations $\hat{\xi}_{ki}$, the variable η in the x-axis and the function $F_k(\eta)$ in the y-axis are linked. Accordingly, the empirical marginal cumulative distribution function for the renewable power unit k writes as

$$F_k(\eta) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\eta \geq \hat{\xi}_{ki}} \text{ with } \mathbb{1}_{\eta \geq \hat{\xi}_{ki}} = \begin{cases} 1 \text{ if } \eta \geq \hat{\xi}_{ki} \\ 0 \text{ otherwise.} \end{cases} \quad (6)$$

Lemma 1. The empirical marginal cumulative distribution function $F_k(\eta)$ for the renewable power unit k is equivalent to the following linear optimization program:

$$\max_{z_{ki}} \quad \frac{1}{N} \sum_{i=1}^{N} z_{ki} \quad (7a)$$

$$\text{s.t. } z_{ki} \left(\eta - \hat{\xi}_{ki} \right) \geq 0 \quad \forall i \in \{1, ..., N\}, \quad (7b)$$

$$0 \leq z_{ki} \leq 1 \qquad \forall i \in \{1, ..., N\}, \quad (7c)$$

where $\eta \in \mathbb{R}$ is the real number that corresponds to the

argument of the function and $z_{ki} \in \mathbb{R} \forall k \forall i$ is the decision variable.

Proof. See Appendix A. $\qquad \square$

Based on this analytical link between $\mathbb{C}$ and $\mathbb{Q}$, it is now possible to reformulate the worst-case expectation problem (5) with the ambiguity set $\mathcal{M}_2$.

Theorem 1. Given N historical observations $\hat{\xi}_i, \ i \in \{1, ..., N\}$ of the random variable $\tilde{\xi}$, the worst-case expectation problem $\max_{\mathbb{Q} \in \mathcal{M}_2} \mathbb{E}^{\mathbb{Q}} \left[a(x)^{\top} \tilde{\xi} + b(x) \right]$ is equivalent to the following conic reformulation:

$$\min_{\Pi} \quad \alpha \theta_1 + \beta \theta_2 + \frac{1}{N} \sum_{i=1}^{N} y_i \quad (8a)$$

$$\text{s.t. } y_i \geq \max_{\xi \in \Xi} \ a(x)^{\top} \xi + b(x) - \zeta_i^{(1)^{\top}} \left(\hat{\xi}_i - \xi \right)$$
$$- \zeta_i^{(2)^{\top}} \left(F\left(\hat{\xi}_i \right) - F(\xi) \right) \quad \forall i \quad (8b)$$

$$\left\| \zeta_i^{(1)} \right\|_* \leq \alpha \ \forall i \quad (8c)$$

$$\left\| \zeta_i^{(2)} \right\|_* \leq \beta \ \forall i, \quad (8d)$$

where $i \in \{1, ..., N\}$. The decision variables are collected in $\Pi = \{x, \ \alpha, \ \beta, \ y_i, \ \zeta_i^{(1)}, \ \zeta_i^{(2)}\}$. In particular, $\alpha, \ \beta \in \mathbb{R}^+$, $y_i \in \mathbb{R} \ \forall i \in \{1, ..., N\}$, and $\zeta_i^{(1)}, \ \zeta_i^{(2)} \in \mathbb{R}^{|\mathcal{W}|}$ are auxiliary variables. Recall that parameters θ_1 and θ_2 in (8a) are the Wasserstein radii defined for the ambiguity set $\mathcal{M}_2$. In addition, $F(\hat{\xi}_i) = \left(F_1(\hat{\xi}_{1i}), ..., F_k(\hat{\xi}_{ki}), ..., F_{|\mathcal{W}|}(\hat{\xi}_{|\mathcal{W}|i}) \right)^{\top}$ and the operator $||.||_*$ computes the dual norm of a vector. The vector $F(\xi) = \left(F_1(\xi_1), ..., F_{|\mathcal{W}|}(\xi_{|\mathcal{W}|}) \right)^{\top}$ is reformulated using Lemma 1, and is given by

$$F(\xi) = \left(\begin{array}{c} \left\{ \begin{array}{l} \max_{z_{jki}} \ \frac{1}{N} \sum_{j=1}^{N} z_{jki} \\ \text{s.t. } z_{jki} \left(\xi_k - \hat{\xi}_{kj} \right) \geq 0 \ \forall j \\ \quad 0 \leq z_{jki} \leq 1 \ \forall j. \end{array} \right\}, k = 1 \\ \vdots \\ \left\{ \begin{array}{l} \max_{z_{jki}} \ \frac{1}{N} \sum_{j=1}^{N} z_{jki} \\ \text{s.t. } z_{jki} \left(\xi_k - \hat{\xi}_{kj} \right) \geq 0 \ \forall j \\ \quad 0 \leq z_{jki} \leq 1 \ \forall j. \end{array} \right\}, k = |\mathcal{W}| \end{array} \right). \quad (9)$$

Proof. See Appendix B. $\qquad \square$

The outcome of Theorem 1 is an optimization problem with constraints involving optimization operators, which is not straightforward to be solved by using off-the-shelf solvers. In Appendix C, we aim to reformulate problem (8), especially constraint (8b), such that it can be incorporated into an optimization problem. To do so, we leverage duality theory and McCormick relaxation of bilinear terms. This will eventually enable us to reformulate the inner worst-case expectation in (5).

V. A POWER SYSTEM APPLICATION

Theorem 1 and its tractable solution approach in Appendix C provide a conic reformulation for a general decision-making problem under uncertainty, and therefore it can be applied to any decision-making optimization problem in power systems. As an example, we apply our proposed model to the day-ahead distributionally robust OPF problem, while capturing the dependencies among renewable energy sources. We derive the reformulation in two particular cases, i.e., *i)* meshed transmission systems using a DC power flow approximation[4], and *ii)* radial distribution systems using a *LinDistFlow* approximation of AC power flow equations[5,6].

A. The OPF Problem for Meshed Transmission Systems

Given a generic ambiguity set $\mathcal{M}$ (which is either $\mathcal{M}_1$ or $\mathcal{M}_2$ introduced in Section II and III), the distributionally robust day-ahead optimal power flow problem for meshed transmission systems reads as

$$\min_{g,\overline{r},\underline{r},V} \quad c^\top g + \overline{c}^\top \overline{r} + \underline{c}^\top \underline{r} + \max_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}^{\mathbb{Q}} \left[c^\top V \widetilde{\xi} \right] \tag{10a}$$

$$\text{s.t.} \quad g + \overline{r} \le g^{\max}, \quad g - \underline{r} \ge g^{\min}, \tag{10b}$$

$$0 \le \underline{r} \le r^{\max}, \quad 0 \le \overline{r} \le r^{\max}, \tag{10c}$$

$$\mathbb{1}^\top g + \mathbb{1}^\top W \mu - \mathbb{1}^\top d = 0, \tag{10d}$$

$$\sum_{p \in \mathcal{P}} V_{p,w} + W_{w,w} = 0 \ \forall w \in \mathcal{W}, \tag{10e}$$

$$\min_{\mathbb{Q} \in \mathcal{M}} \mathbb{Q} \left(-\underline{r}_p \le V_p \widetilde{\xi} \right) \ge 1 - \underline{\epsilon}_p \ \forall p \in \mathcal{P}, \tag{10f}$$

$$\min_{\mathbb{Q} \in \mathcal{M}} \mathbb{Q} \left(V_p \widetilde{\xi} \le \overline{r}_p \right) \ge 1 - \overline{\epsilon}_p \quad \forall p \in \mathcal{P}, \tag{10g}$$

$$\min_{\mathbb{Q} \in \mathcal{M}} \mathbb{Q} \left(T_f^{\mathcal{P}} \left(g + V \widetilde{\xi} \right) + T_f^{\mathcal{W}} W \left(\mu + \widetilde{\xi} \right) \right. $$
$$\left. - T_f^{\mathcal{D}} d \right) \le f_f^{\max} \right) \ge 1 - \epsilon_f \quad \forall f \in \mathcal{F}. \tag{10h}$$

[4]We highlight that the DC model of power flows [24] is a linear and convex approximation to the original complete set of non-linear and non-convex equations governing the AC power flow model. The resulting DC power flow model relies on assumptions that are usually deemed comparatively less strong for high-voltage transmission lines, i.e., lossless power lines, constant voltage magnitudes and small voltage angle differences along a line. However, these assumptions do not guarantee that the solution obtained is necessarily feasible with respect to the AC power flow equations [25].

[5]The solution of the DC-OPF problem may not be feasible with respect to the complete set of AC power flow equations [25]. As a prospect to the contribution in this work, we highlight the extension of our work to the most complete set of AC power flow equations. In partial fulfilment to this prospect, we consider the *LinDistFlow* approximation of AC power flow equations, which is valid for radial distribution systems, and show in the numerical results that the proposed copula-based approach for considering dependencies between wind farms outperforms the traditional techniques which disregard the correlations.

[6]To the best of our knowledge, this is the first time in the literature that a metric-based distributionally robust OPF problem is derived using *LinDistFlow* approximation of AC power flow equations (please see [26] for a moment-based counterpart). The reason for this is that the resulting distributionally robust models are usually highly complex to solve. The complexity arises from the additional modeling layer required by DRO, where the operating constraints such as transmission line capacity limits need to be formulated in an appropriate linear or convex form for introducing uncertain fluctuations. Any additional contribution towards using more extended AC modeling of power flows, e.g., those investigated in [27], within the proposed distributionally robust framework would require advanced convexification techniques dedicated to DRO, which is left for future work.

The problem (10) optimizes the day-ahead dispatch $g \in \mathbb{R}^{|\mathcal{P}|}$ of conventional generating units as well as their upward and downward reserve capacity $\overline{r} \in \mathbb{R}^{|\mathcal{P}|}$ and $\underline{r} \in \mathbb{R}^{|\mathcal{P}|}$ to be booked in the day-ahead stage. Objective function (10a) minimizes the total operational cost of the system. The first three terms in (10a) are linear and contains the energy production cost $c \in \mathbb{R}^{|\mathcal{P}|}$ and the upward/downward reserve procurement cost $\overline{c} \in \mathbb{R}^{|\mathcal{P}|}$ and $\underline{c} \in \mathbb{R}^{|\mathcal{P}|}$. The fourth term in (10a), i.e., $\max_{\mathbb{Q} \in \mathcal{M}} \mathbb{E}^{\mathbb{Q}} \left[c^\top V \widetilde{\xi} \right]$, refers to the operational cost of conventional units, incurred by the activation of their reserve capacity in the real-time operation. This cost is calculated in expectation with respect to the worst-case distribution $\mathbb{Q}$, which is endogenously selected within the ambiguity set $\mathcal{M}$. It is approximated by linear decision rules [28] using matrix $V \in \mathbb{R}^{|\mathcal{P}| \times |\mathcal{W}|}$, whose elements are decision variables. The uncertainty in the distributionally robust OPF problem arises from renewable wind power in-feed, when the real-time realization deviates from its day-ahead forecast. We model the renewable wind power in-feeds as

$$\omega = \mu + \tilde{\xi}, \tag{11}$$

where $\omega \in \mathbb{R}^{|\mathcal{K}|}$ describes the actual real-time wind power generation, $\mu \in \mathbb{R}^{|\mathcal{K}|}$ defines the day-ahead forecast for wind power generation and $\tilde{\xi} \in \mathbb{R}^{|\mathcal{K}|}$ is the deviation in real-time from day-ahead forecast[7,8]. In the following, we consider $\tilde{\xi}$ as the uncertain parameter in the distributionally robust OPF problem, and model it as a random variable described by the probability distribution $\mathbb{Q}$. Note that the renewable production cost is assumed to be zero.

Constraints (10b) ensure that the generation level of conventional units lies within their maximum and minimum limits, i.e., $g^{\max} \in \mathbb{R}^{|\mathcal{P}|}$ and $g^{\min} \in \mathbb{R}^{|\mathcal{P}|}$. Constraints (10c) enforce the maximum amount of reserve capacity $r^{\max} \in \mathbb{R}^{|\mathcal{P}|}$ that can be provided by conventional units. Constraints (10d) and (10e) ensure the day-ahead and real-time power balance, respectively. In particular, the day-ahead constraint (10d) enforces the sum of total production of conventional generating units $\mathbb{1}^\top g$ and total forecasted production of renewable units $\mathbb{1}^\top W \mu$ to be equal to total demand $\mathbb{1}^\top d$. Note that the demands $d \in \mathbb{R}^{|\mathcal{D}|}$ are assumed inelastic to prices. In addition, $W \in \mathbb{R}^{|\mathcal{W}| \times |\mathcal{W}|}$ is a diagonal matrix of the installed capacity of renewable power units, whereas $\mu \in \mathbb{R}^{|\mathcal{W}|}$ gives their per-unit power generation forecast. The real-time balance in (10e) is ensured via the elements of matrix V, the so-called participation factors, which can be interpreted as follows. The conventional unit p responds to any deviation in renewable power generation

[7]We assume a non-biased forecaster, such that the average forecast error $\mathbb{E} \left[\tilde{\xi} \right]$ equals 0. Furthermore, we assume that the wind power units are dispatched in day-ahead at a value equal to the single-point forecast μ.

[8]An alternative modeling approach would consider the renewable wind power units as dispatchable units which requires an additional decision variable modeling their day-ahead schedule. The main difference of this alternative compared to (11) pertains to the source of uncertainty (i.e., forecast errors in one case, and actual real-time wind power generation in the other one). We hypothesize that the innovation of incorporating the dependence information about the uncertain parameters into the definition of ambiguity set remains beneficial in both cases. The copula-based restriction results in an enhanced representation of dependencies whatsoever the underlying source of uncertainty is. Therefore, the choice of modeling approach will have a limited impact on the final conclusions of the paper.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2022.3221348

6

of farm w, i.e., $\widetilde{\xi}_w$ (with respect to the day-ahead forecast μ_w) based on its corresponding participation factors $V_{p,w}$. Therefore, for renewable power deviation $\widetilde{\xi}_w$, the recourse action of conventional unit p is $V_{p,w}\widetilde{\xi}_w$, such that the total recourse action $\sum_{p\in\mathcal{P}} V_{p,w}\widetilde{\xi}_w$ compensates the total deviation $W_{w,w}\widetilde{\xi}_w$. Note that (10e) is an equality constraint, so $\widetilde{\xi}_w$ can be dropped from both sides. In the real-time operation, the recourse action $V_p\widetilde{\xi} = \sum_{w\in\mathcal{W}} V_{p,w}\widetilde{\xi}_w$ of conventional units should be limited to the reserve capacities procured in the day-ahead stage, i.e., $\overline{r}_p$ and $\underline{r}_p$. Similarly, the real-time flow within each transmission line $f \in \mathcal{F}$ should respect the capacity $f_f^{\max}$. The real-time flows are expressed using the power transfer distribution factor matrices $\mathbf{T}_f^{\mathcal{P}} \in \mathbb{R}^{|\mathcal{F}|\times|\mathcal{P}|}$, $\mathbf{T}_f^{\mathcal{W}} \in \mathbb{R}^{|\mathcal{F}|\times|\mathcal{W}|}$ and $\mathbf{T}_f^{\mathcal{D}} \in \mathbb{R}^{|\mathcal{F}|\times|\mathcal{D}|}$ for conventional generating units, renewable energy sources, and demands, respectively. Note that these restrictions are enforced via via probabilistic constraints, namely, Distributionally Robust Chance Constraints (DRCCs), for which we provide further details later in Section V-C.

B. The OPF Problem for Radial Distribution Systems

In this section, we derive the distributionally robust day-ahead OPF problem for radial distribution systems using the *LinDistFlow* approximation of AC power flow equations [29]. Given the generic ambiguity set $\mathcal{M}$, the model reads as

$$\min_{g^\dagger, f^\dagger, V, u} \sum_{i\in\mathcal{N}} c_i g_i^P + \max_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}}\left[\sum_{i\in\mathcal{N}} c_i V_i \widetilde{\xi}^P\right] \quad (12a)$$

$$\text{s.t. } g_0^\dagger = \sum_{i\in\mathcal{D}_0} \left(d_i^\dagger - g_i^\dagger\right) - \mu_i^\dagger, \, \dagger \in \{P, Q\}, \quad (12b)$$

$$u_0 = 1, \quad (12c)$$

$$f_l^\dagger = \sum_{i\in\mathcal{D}_l} \left(d_i^\dagger - g_i^\dagger - \mu_i^\dagger\right), \, \dagger \in \{P, Q\}, \forall l \in \mathcal{L}, \quad (12d)$$

$$u_i = u_0 - 2\sum_{l\in\mathcal{R}_i} \left(f_l^P R_l + f_l^Q X_l\right), \forall i \in \mathcal{L}, \quad (12e)$$

$$\left(f_l^P\right)^2 + \left(f_l^Q\right)^2 \le \bar{f}_l^2, \, \forall l \in \mathcal{L}, \quad (12f)$$

$$\sum_{i\in\mathcal{N}} V_{i,w} + 1 = 0 \, \forall w \in \mathcal{W}, \quad (12g)$$

$$\min_{\mathbb{Q}\in\mathcal{M}} \mathbb{Q}\left(\widetilde{g}_i^\dagger \le \bar{g}_i^\dagger\right) \ge 1 - \underline{\epsilon}_i^\dagger, \, \dagger \in \{P, Q\}, \, \forall i \in \mathcal{N}, \quad (12h)$$

$$\min_{\mathbb{Q}\in\mathcal{M}} \mathbb{Q}\left(\underline{g}_i^\dagger \le \widetilde{g}_i^\dagger\right) \ge 1 - \underline{\epsilon}_i^\dagger, \, \dagger \in \{P, Q\}, \, \forall i \in \mathcal{N}, \quad (12i)$$

$$\min_{\mathbb{Q}\in\mathcal{M}} \mathbb{Q}\left(\widetilde{u}_i \le \bar{v}_i^2\right) \ge 1 - \epsilon_i, \, \forall i \in \mathcal{N}_{\backslash 0}, \quad (12j)$$

$$\min_{\mathbb{Q}\in\mathcal{M}} \mathbb{Q}\left(\underline{v}_i^2 \le \widetilde{u}_i\right) \ge 1 - \epsilon_i, \, \forall i \in \mathcal{N}_{\backslash 0}, \quad (12k)$$

where $\dagger = \{P, Q\}$ represents the superscripts related to, respectively, active and reactive power. The set $\mathcal{N}$ collects the nodes of the radial distribution system and the sets $\mathcal{D}_i$ and $\mathcal{R}_i$ respectively include the downstream and root nodes to node i. The set $l \in \mathcal{L}$ represents the set of distribution lines. Given that, the problem (12) seeks the optimal generation dispatch of controllable generators $g^P \in \mathbb{R}^{|\mathcal{N}|}$ at the distribution level.

The objective function (12a) minimizes the total operating cost, composed of the day-ahead energy production cost $c \in \mathbb{R}^{|\mathcal{N}|}$ and the real-time operating cost of recourse actions. Similarly to model (10), the real-time cost is calculated in expectation for the worst-case distribution $\mathbb{Q}$ of wind power generation forecast errors $\widetilde{\xi}^\dagger \in \mathbb{R}^{|\mathcal{W}|}$ in the ambiguity set $\mathcal{M}$, and the recourse actions are modeled via linear decision rules, using $V \in \mathbb{R}^{|\mathcal{N}|\times|\mathcal{W}|}$.

Constraint (12b) calculates the required active and reactive power injections $g_0^P \in \mathbb{R}$ and $g_0^Q \in \mathbb{R}$ at the interface between transmission and distribution systems (node 0) to balance the mismatch between consumption $d^\dagger \in \mathbb{R}^{|\mathcal{N}|}$ and generation $g^\dagger \in \mathbb{R}^{|\mathcal{N}|}$ in the downstream distribution system. Constraint (12c) imposes the voltage magnitude u_0 at the slack node, to be equal to 1 per-unit. Constraint (12d) calculates the active and reactive power flows $f_l^P \in \mathbb{R}$ and $f_l^Q \in \mathbb{R}$ in the distribution lines $l \in \mathcal{L}$. Constraint (12e) defines the per-unit square voltage $u_i \in \mathbb{R}$ at each node of the distribution system, given the per-unit resistances R_l and reactances X_l, and the active and reactive power flows f_l^P and f_l^Q in the root lines. The apparent power flow $f_l^2 \in \mathbb{R}$ is limited to $\bar{f}_l^2 \in \mathbb{R}$ via (12f) for all lines in the distribution system. Constraints (12g) to (12j) impose the real-time operational restrictions. Constraint (12g) enforces the real-time balancing between generation and consumption, using the linear decision rules, similarly to model (10). The minimum and maximum operating limits of active and reactive power generation, and voltage magnitude are enforced via the DRCCs (12h) to (12k). The rationale behind this is that the voltage magnitude and the active and reactive power in-feeds are real-time state variables that depend on the wind power uncertainty $\widetilde{\xi}^\dagger$. To appropriately model these real-time variations, we further define the real-time power generation of controllable generators as

$$\widetilde{g}_i^\dagger = g_i^\dagger + V_i \widetilde{\xi}^\dagger, \, \dagger \in \{P, Q\}, \quad (13a)$$

where the first term refers to the day-ahead dispatch $g_i^\dagger$ and the second term refers to the real-time response $V_i \widetilde{\xi}^\dagger$ to deviations $\widetilde{\xi}^\dagger$, using the matrix of participation factors $V \in \mathbb{R}^{|\mathcal{G}|\times|\mathcal{W}|}$. In addition, we model the real-time flows $\widetilde{f}_l^P$ and $\widetilde{f}_l^Q$ as follows:

$$\widetilde{f}_l^\dagger = f_l^\dagger + \sum_{i\in\mathcal{D}_l} \left(-V_i \widetilde{\xi}^\dagger - \widetilde{\xi}_l^\dagger\right), \, \dagger \in \{P, Q\}, \quad (14a)$$

where the power flow $f_l^\dagger$ scheduled in day-ahead is supplemented with the corresponding real-time deviation, composed of the deviation in renewable power generation and the compensation made by the controllable generators in the downstream nodes. Finally, the real-time square voltage $\widetilde{u}_i$ is expressed as follows:

$$\widetilde{u}_i = u_0 - 2\sum_{l\in\mathcal{R}_i} \left(\widetilde{f}_l^P R_l + \widetilde{f}_l^Q X_l\right), \quad (15a)$$

where the real-time expression of power flows are used. The combination of (13a), (14a) and (15a) with the model (12) results in the complete distributionally robust OPF formulation for radial distribution systems.

C. Reformulation of (10) and (12)

The inequalities (10f) to (10h) as well as (12h) to (12k) are enforced via DRCCs. These DRCCs state that the probabilistic constraints within parentheses should be respected under the worst-case distribution $\mathbb{Q}$ within the ambiguity set $\mathcal{M}$ with a probability not lower than $1 - \epsilon$. Note that the value of parameter $\epsilon \in \mathbb{R}$ lies between zero and one, fixed by the power system operator.

The procedure to reformulate these DRCCs is the same for all inequalities considered in both power system applications. For this purpose, we use a Conditional-Value-at-Risk (CVaR) approximation [30]. Consider a generic DRCC in the form of

$$\min_{\mathbb{Q} \in \mathcal{M}} \; \mathbb{Q}\left(. \leq 0\right) \geq 1 - \epsilon. \tag{16a}$$

In order to get rid of the probability operator $\mathbb{Q}(.)$, the DRCC (16a) can be approximated by the following CVaR constraint:

$$\max_{\mathbb{Q} \in \mathcal{M}} \; \mathbb{Q}\text{-CVaR}_\epsilon(.) \leq 0. \tag{16b}$$

The CVaR operator in the left-hand side of (16b) is defined as

$$\min_{\tau \in \mathbb{R}} \; \tau + \frac{1}{\epsilon} \; \max_{\mathbb{Q} \in \mathcal{M}} \; \mathbb{E}^{\mathbb{Q}}\left[\lceil . - \tau \rceil^+\right], \tag{16c}$$

where $\tau \in \mathbb{R}$ is an auxiliary variable and $\lceil . \rceil^+ = \max(., 0)$. By this approximation, the worst-case expectation problem appears not only in the objective function (10a), but also in every approximated DRCC in the form of (16c). Note that the inequalities (10f) to (10h) as well as (12h) to (12k) are linear. Therefore, the worst-case expectation problem appearing in (16c) complies with the setting of Theorem 1. In the following, we provide all reformulations required to solve the distributionally robust OPF problem under both ambiguity sets $\mathcal{M}_1$ and $\mathcal{M}_2$, in a generic way such that it can be applied to both DC and AC formulations.

1) Metric-Based Ambiguity Set $\mathcal{M}_1$: Reference [9] provides the reformulation of a generic worst-case expectation problem in the form of $\max_{\mathbb{Q} \in \mathcal{M}_1} \; \mathbb{E}^{\mathbb{Q}}\left[a(x)^\top \widetilde{\xi} + b(x)\right]$. Similarly to Theorem 1, this maximization over probability distributions $\mathbb{Q} \in \mathcal{M}_1$ can be recast into a minimization problem at the cost of introducing a set of additional auxiliary variables. With this reformulation, the objective function (10a) or (12a) would contain two min operators that can be merged, yielding a formulation that can be fed into an off-the-shelf solver. Next, we reformulate the worst-case expectation appeared in constraints (10g) to (10h) after using the CVaR approximation. Following [9], the minimization operators appeared in the constraints can be eventually dropped, but again at the cost of additional auxiliary variables. For the sake of conciseness and completeness, we provide the reformulation of a generic worst-case expectation problem, as well as the final reformulation of problems (10) and (12) under ambiguity set $\mathcal{M}_1$ in the online companion [23].

2) Copula-Based Ambiguity Set $\mathcal{M}_2$: The general procedure to reformulate (10) and (12) under ambiguity set $\mathcal{M}_2$ is similar to the one under set $\mathcal{M}_1$. However, instead of following the reformulations in [9], we apply the outcomes of Theorem 1 in Section IV. Recall that the solution approach for Theorem 1 proposed in Appendix C, allows us to recast the maximization problem over probability distributions $\mathbb{Q} \in \mathcal{M}_2$ as a minimization problem, but at the cost of additional auxiliary variables. We provide the final reformulation of problems (10) and (12) under ambiguity set $\mathcal{M}_2$ in the online companion [23].

VI. NUMERICAL EXPERIMENTS

We first perform numerical experiments for a meshed transmission system considering the distributionally robust OPF problem (10). We discuss the procedure of out-of-sample analysis, the computational performance and the out-of-sample performance. Next, we introduce a distribution system case study and derive similar experiments considering the distributionally robut OPF problem (12).

Our meshed transmission system case study is based on a slightly updated version of the 24-node IEEE reliability test system [31], whose input data are provided in the online companion [23]. This case study is composed of 12 conventional generating units with an aggregate capacity of $2,362.5$ MW, two wind farms with a total maximum installed capacity of $1,000$ MW, and eventually 17 loads with an aggregate demand of $2,207$ MW. These power suppliers and demands are connected through a network composed of 24 nodes and 34 transmission lines. The sole source of uncertainty is the deviation of renewable power generation in the real-time operation with respect to the day-ahead forecast values. To cope with such an uncertainty, the system operator reserves a fraction of capacities of conventional units in the day-ahead stage, and activates them in the real-time operation, if necessary. The total maximum reserve capacity that conventional units can provide is 798 MW.

To assess the impacts of dependence structure on operational decisions of the system operator, we generate a dataset of $1,000$ samples representing the historical wind power observations. To do so, we use the package *DatagenCopulaBased* v1.3.0 in Julia programming language v1.4.2, which allows to generate samples with a predefined dependence structure (e.g., following the Gaussian copula). The resulting dataset mimics historical wind power observations from which we retrieve the mean value μ corresponding to the day-ahead forecast. The final dataset of forecast errors and its corresponding copula will be illustrated later in Fig. 3(a).

A. Procedure of the Out-of-Sample Analysis

Aiming to conduct an ex-post out-of-sample analysis and therefore to compare different models on a fair basis, we split the dataset with $1,000$ samples into two different sets of samples. The first one contains 30 samples only (the in-sample data), which are used to characterize the uncertain wind power generation within the models. Therefore, $N = 30$ and indices i and j in (26) run from 1 to 30. The remaining 970 samples in the dataset are used as unseen wind power realizations to assess the quality of operational decisions made in the day-ahead stage.

The out-of-sample analysis is as follows. We first solve problem (10) under ambiguity sets $\mathcal{M}_1$ and $\mathcal{M}_2$, fed with in-

sample dataset[9], and obtain the optimal day-ahead decisions g, $\underline{r}$, and $\overline{r}$. Given fixed values of those day-ahead decisions, we then solve a deterministic optimization problem in the real-time operation for each of 970 unseen samples, whose outcomes are the recourse action of conventional generating units as well as the involuntarily load shedding and the wind curtailment as two extreme recourse actions. The involuntarily load shedding is required when there is a wind power deficit in the real time, and that the upward reserve capacities provided by conventional units are insufficient to compensate the entire deficit. Similarly, the wind curtailment occurs when there is a wind power excess in the real time, combined with the lack of downward reserve capacities provided by conventional units to absorb all available wind power. The formulation of this deterministic model is available in the online companion [23]. Once this deterministic optimization problem is solved 970 times for each model, we calculate the average real-time operational cost of the system. In the rest of this section, we report the out-of-sample cost, which is the sum of operational cost of the system in the day-ahead stage, i.e., the first three terms in (10a), and the average real-time operational cost calculated from the out-of-sample analysis.
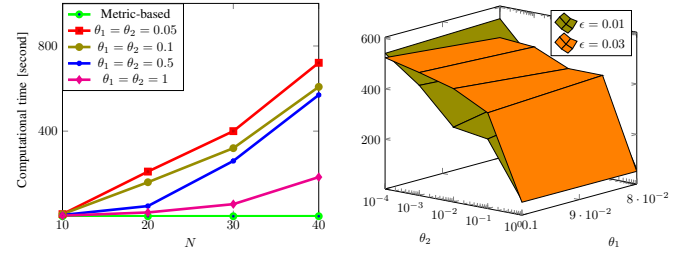
B. Computational Performance

All models are solved using Gurobi v8.0.1 in JuMP v0.21.3 under programming language JuliaPro v1.4.2 on a usual computer clocking at 2.2 GHz with 16 GB of RAM. All source codes are publicly available in the online companion [23].

Fig. 2(a) shows the computational time as a function of the number of historical in-sample observations N under different settings. We observe that the computational time increases with the number of historical observations in a non-linear manner. These results suggest that the computational time growth follows a quadratic trend, which is in line with the increase in the number of variables and constraints — see Theorem 1, where the number of certain variables and constraints increases quadratically[10] in N. In addition, Fig. 2(b) depicts the computational time for various values of parameters θ_1 and θ_2 in the proposed ambiguity set $\mathcal{M}_2$, and for $\epsilon = 0.01$ and 0.03. The results highlight that the computational time decreases when θ_2 increases. Recall that two wind farms only have been considered so far. We will investigate later in Section VI-F the scalability of the proposed model with respect to the dimensionality of the uncertainty

(a) Computational time as a function of the number of historical in-sample observations N. Fixed value: $\epsilon = 0.05$.

(b) Computational time as a function of parameters θ_1 and θ_2. Fixed values: $N = 25$. The metric-based ambiguity set approach finds a solution in less than 1 second for the values of θ_1 explored in this study.

Fig. 2. Computational study.

space by increasing the number of wind farms[11].

In the next two subsections, we investigate the out-of-sample performance of the DRO model with the ambiguity sets $\mathcal{M}_1$ and $\mathcal{M}_2$. Recall that the ambiguity set $\mathcal{M}_1$ contains one risk-tuning parameter only[12], i.e., θ_1, that restricts the distance of distributions within the ambiguity set to the empirical one. In contrast, the ambiguity set $\mathcal{M}_2$ comprises of two risk-tuning parameters, i.e., θ_1 and θ_2, where θ_2 restricts the similarity of distributions within the ambiguity set in terms of the dependence structure to that of the empirical one. For the ease of comparison, we alternately fix one of the two risk-tuning parameters θ_1 and θ_2, while varying the other one.

C. Out-of-Sample Performance: Operational Cost of the System as a Function of θ_1

We assign three different values to parameter θ_2, namely 0.01, 0.02, and 0.1. We vary the value of parameter θ_1 from 10^{-4} to 10^0, where the exponent increases linearly with a step of 0.2, and compare the out-of-sample performance of the DRO model with the ambiguity sets $\mathcal{M}_1$ and $\mathcal{M}_2$. We retrieve the average out-of-sample operational cost of the system. In addition, the standard deviation of the cost over 970 out-of-sample simulations is calculated. These results are illustrated in logarithmic Fig. 3(b).

One can observe that the results obtained from the DRO model with the ambiguity set $\mathcal{M}_1$ and those with the ambiguity set $\mathcal{M}_2$ when θ_2 is comparatively high (e.g., 0.1) are similar. This implies that the second constraint in $\mathcal{M}_2$ regarding the distance with respect to the empirical copula is not binding, and therefore the worst-case distributions for which the optimal decisions are made are identical in both DRO models. In other words, the ambiguity set $\mathcal{M}_2$ is identical to $\mathcal{M}_1$ when θ_2 takes comparatively high values. In contrast,

---

[9]Recall that the ambiguity set $\mathcal{M}_1$ defines the state-of-the-art Wasserstein metric-based ambiguity set, which serves as the benchmark solution in the subsequent numerical experiments, for assessing the performances of the proposed copula-based ambiguity set $\mathcal{M}_2$ under various values of θ_1 and θ_2.

[10]Compared to other uncertainty modeling techniques, e.g., scenario-based stochastic programming, the DRO approach generally provides more qualified decisions in terms of the out-of-sample performance, when the number of historical in-sample observations is relatively low. This may further motivate the use of DRO when there is limited historical data, or when the decision-maker aims to reduce the computational time by intentionally reducing the number of historical observations. As shown in Fig. 2(a), the computational time for the DRO approach is satisfactory when N is comparatively low.

[11]In this paper, we consider the scalability of our proposed framework with respect to the number of in-sample and the number of wind farms, as these indicators directly impact the description of uncertainty. The scales of networks used in these simulations are usually deemed acceptable for numerical experiment purposes, and allow us to bring valuable and informative discussions on the outcomes of the proposed method.

[12]The violation probability $\epsilon_{(.)}$ of each DRCC can also be seen as a risk-tuning parameter. However, we keep this value unchanged (fixed to 0.05) in our numerical study, and focus on the risk-tuning parameters related to the proposed ambiguity set.

(a) Historical wind power generation of two farms (red plot) and the corresponding empirical copula (blue plot)

(b) Out-of-sample cost as a function of θ_1 considering different values for θ_2

(c) Out-of-sample cost as a function of θ_2 considering different values for θ_1
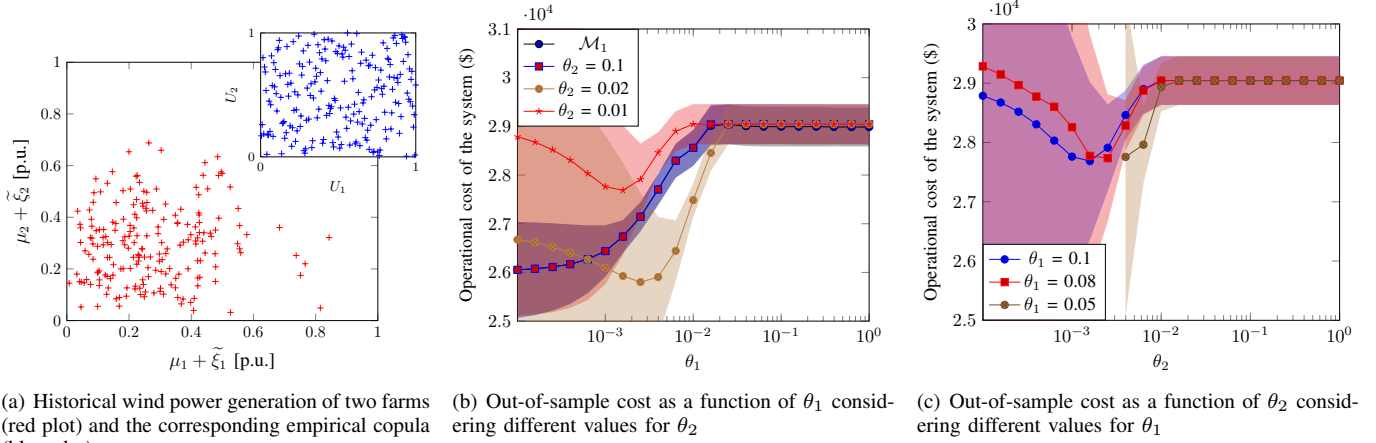
Fig. 3. Numerical study. The shaded area around each curve represents the corresponding standard deviation.

when the value of θ_2 decreases, e.g., to 0.02, the second constraint in $\mathcal{M}_2$ becomes binding, leading to improved results in terms of the out-of-sample cost. This improvement has led to the lowest out-of-sample cost, since the constraint on dependence structure eliminates the unrealistic distributions from the ambiguity set. This numerical finding suggests that the proposed ambiguity set $\mathcal{M}_2$ outperforms $\mathcal{M}_1$, provided that appropriate values for parameters θ_1 and θ_2 are selected (i.e., $\theta_1 = 0.0025$ and $\theta_2 = 0.02$). Given these appropriate values for θ_1 and θ_2, we observe a potential cost saving of 1% in expectation with respect to the classical ambiguity set $\mathcal{M}_1$ given the value for $\theta_1 = 10^{-4}$, which achieves the minimal expected total cost for $\mathcal{M}_1$.

It is worth mentioning that the DRO model becomes infeasible, when both parameters θ_1 and θ_2 take very low values, meaning that both constraints in $\mathcal{M}_2$ are highly restrictive. We have observed infeasibility for values of θ_1 and θ_2 lower than 0.01 in our simulations in Section V-C and for values lower than $\theta_1 = 0.05$ and $\theta_2 = 0.004$ in our simulations in Section V-D. One can intuitively interpret the observation on infeasibility as a case under which the two Wasserstein balls, defined by two constraints in $\mathcal{M}_2$, have no intersection. In other words, there is no distribution within the ambiguity set that satisfies both constraints at the same time.

D. Out-of-Sample Performance: Operational Cost of the System as a Function of θ_2

We fix the value of parameter θ_1 to either 0.05, 0.08, or 0.1, and report the out-of-sample operational cost of the system for different values of θ_2. In particular, this value ranges from 10^{-4} to 10^0, where the exponent increases linearly with a step equal to 0.2. Fig. 3(c) shows the out-of-sample results in a logarithmic scale.

Our first observation is that for comparatively high values of θ_2, the outcomes of the DRO model with the ambiguity set $\mathcal{M}_2$ are very similar to those with the ambiguity set $\mathcal{M}_1$. This numerically indicates that the copula constraint is not binding. Second, we observe that the cost curves reach a minimum at intermediate values for θ_2, suggesting that an optimal value for the parameter θ_2 exists. This difference between the outcomes

obtained by the DRO model with $\mathcal{M}_1$ and those with $\mathcal{M}_2$, given the optimal value for θ_2, represents the maximum cost earned due to the use of the more elegant $\mathcal{M}_2$ instead of $\mathcal{M}_1$, for fixed value of θ_1. For instance, when $\theta_1 = 0.1$, the optimum is obtained for $\theta_2 = 0.0016$, and it corresponds to a cost saving of 4.7% compared to $\mathcal{M}_1$ with the same value of $\theta_1 = 0.1$. Third, we observe that the program becomes infeasible for comparatively low values of θ_2, when θ_1 is also low (e.g., 0.01), confirming our intuition in the previous analysis.

Remark 2 (On the selection of θ_1 and θ_2 by the system operator). It is worth mentioning that when the decision-maker has less confidence in the available representation of uncertainty and dependence structure (e.g., when the number of historical in-sample observations is relatively low), the system operator should select θ_1 and θ_2 as large values, e.g., $\theta_1 = \theta_2 = 1$. This would result in similar outcomes to those obtained by the state-of-the-art metric-based ambiguity set. Next, by slightly decreasing the value of θ_1 and θ_2 from day-to-day, the system operator should observe a decrease in the average total operating cost, followed by an increase (corresponding to a displacement from right to left on the curves of Fig. 2(b) and 2(c)). The meaning behind this observation is that the optimal values for θ_1 and θ_2 have been found and should no longer be decreased, for the corresponding OPF problem to be solved.

E. Operational Decisions

We explore the impacts of the proposed ambiguity set $\mathcal{M}_2$ on the resulting operational decisions. Fig. 4(a) reports the total upward reserve capacity, i.e., $\mathbb{1}^\top \overline{r}$, procured from conventional units in the day-ahead stage. In particular, we solve the DRO model under three different settings: (*i*) with the ambiguity set $\mathcal{M}_1$, (*ii*) with the ambiguity set $\mathcal{M}_2$ where $\theta_2 = 0.001$, and (*iii*) with the ambiguity set $\mathcal{M}_2$ where $\theta_2 = 0.1$. In all three cases, we fix the value of θ_1 to be 0.1. We observe that when θ_2 takes a comparatively higher value (i.e., 0.1), the resulting operational decisions obtained from the DRO model with the ambiguity set $\mathcal{M}_2$ are identical

(a) Comparison of the total upward reserve capacity dispatched (i.e., $\mathbb{1}^{\top}\overline{r}$), obtained from the DRO model under different settings. Fixed values: $\epsilon = 0.05$, $\theta_1 = 0.1$, and $N = 30$.

(b) Computational time as a function of the number of wind farms. Fixed values: $\theta_1 = 0.1$, $\epsilon = 0.05$, and $N = 15$.

(c) Out-of-sample operational cost of the system as a function the number of wind farms. Legends are the same at those in plot (b). Fixed values: $\theta_1 = 0.1$, $\epsilon = 0.05$, and $N = 15$.
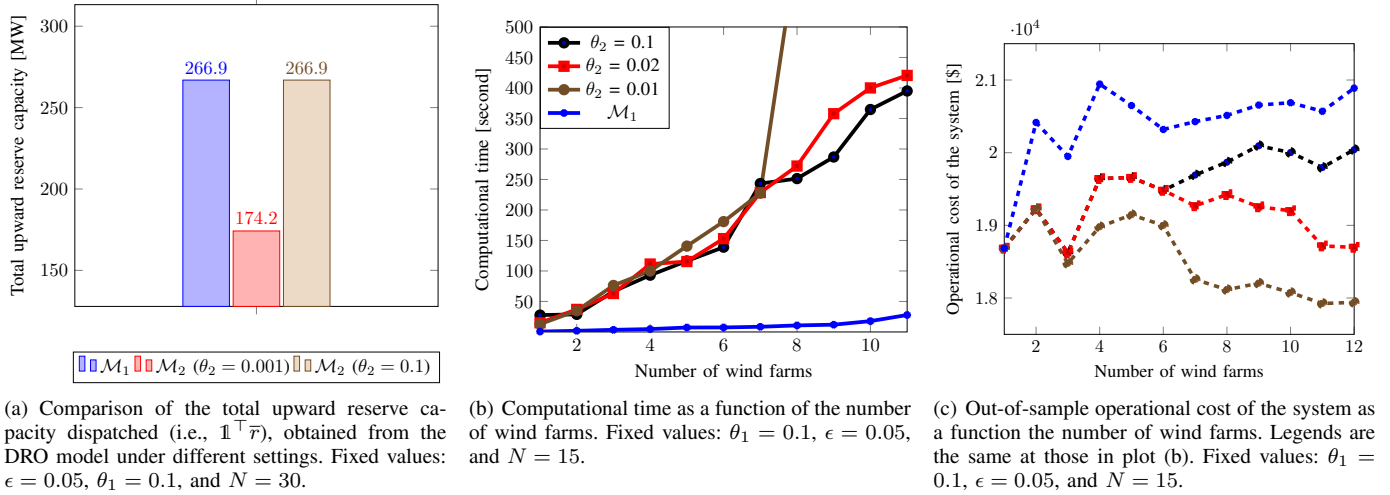
Fig. 4. Operational results with two (the first plot) and more (the next two plots) wind farms. The aggregate capacity of farms is always equal to $1,000$ MW.

to those with the ambiguity set $\mathcal{M}_1$. This confirms our earlier observation that the copula constraint in $\mathcal{M}_2$ is non-binding. On the contrary, when the parameter θ_2 is given an appropriate value (here, e.g., 0.001), not only the out-of-sample cost (see Fig. 3(c)) but also the total amount of reserve capacity (see Fig. 4(a)) decreases. This interesting observation implies that by elegantly incorporating the dependence structure into the ambiguity set definition, the DRO model will provide *less* conservative operational decisions.

F. Impacts of Increasing the Number of Wind Farms

We increase the number of wind farms connected to the system, enlarging the dimension of the uncertainty space. We consider up to 12 wind farms, which represent the number of wind uncertainty sources which is typically considered for studies related to the real-life Belgian electricity grid (i.e., one source per province and an additional one for off-shore wind). We re-scale the wind farm capacities, such that the aggregate capacity of wind farms is always equal to $1,000$ MW. We consider 15 historical observations.

Fig. 4(b) shows the computational time to solve the problem (10) with both ambiguity sets $\mathcal{M}_1$ (metric-based) and $\mathcal{M}_2$ (with different values of θ_2). The value of parameter θ_1 is fixed to 0.1 in all cases. In most cases, the computational time increases linearly with the number of wind farms. However, this increase is drastic when the value of θ_2 is comparatively low, indicating the case with a tight copula constraint.

Fig. 4(c) illustrates the out-of-sample operational cost of the system as a function of the number of wind farms. Legends are the same as those in Fig. 4(b). There are two important observations. First, we obtain lower values for the operational cost under the ambiguity set $\mathcal{M}_2$, in particular when the value of θ_2 is comparatively low, e.g., 0.01 or 0.02, meaning that the copula constraint is binding and that the dependence structure plays a role on the collection of distributions within $\mathcal{M}_2$. Second, the operational cost decreases by increasing the number of wind farms, provided that the value of θ_2 is comparatively low. Therefore, the benefits of using $\mathcal{M}_2$ instead of $\mathcal{M}_1$ increases when the number of wind farms is comparatively

high. Both observations highlight the importance of adding dependence structure in the definition of the ambiguity set, particularly when the number of wind farms is relatively high.

G. Performance in Case of Radial Distribution Systems

We consider a radial distribution system composed of 15 nodes, two of which host a controllable generator, each being able to produce electricity for up to 1 MW. Two wind turbines with a total capacity of 1 MW are also connected to the distribution system. The complete dataset of economical and technical parameters is provided in the online companion [23].

The renewable and dependent wind power generation introduces uncertainty in the operation of the system. To cope with this uncertainty, we solve the distributionally robust OPF (12) for radial distribution systems, given the metric-based ambiguity set $\mathcal{M}_1$ and the proposed copula-based ambiguity set $\mathcal{M}_2$. We leverage the per-unit wind power deviation dataset generated in Section VI-A for the meshed transmission case study and perform an out-of-sample analysis, with a procedure similar to the one showcased in Section VI-A, to fairly compare different models with respect to unseen realizations of uncertainty. The formulation of the underlying deterministic real-time optimization problem is available in the online companion [23].

We report the results in terms of expected total operating cost and its standard deviation, as well as the Expected Energy Not Served (EENS) in MWh, in Table I. The EENS is an indicator that is usually chosen by the system operator to evaluate the amount of load that is expected to be shed during the corresponding time period. In addition to providing insights of constraint violation probability, this indicator embeds the information of the severity of the constraint violation.

Similarly to the results in Sections VI.C and VI.D, we observe that for given appropriate values for θ_1 and θ_2, e.g., $\theta_1 = 0.01$ and $\theta_2 = 0.01$, the proposed copula-based approach achieves the lowest expected total operating cost (and standard deviation) compared to the traditional metric-based approach. For a fixed value of θ_1, the maximum cost saving observed in

		Expected cost [k€]	Standard deviation [k€]	EENS [MWh]
	Metric-based	120.76	11.7	0.068
$\theta_1 = 0.1$	$\theta_2 = 0.01$	120.9	11.7	0.068
	$\theta_2 = 0.001$	109.1	11.7	0.069
	Metric-based	120.8	11.7	0.068
$\theta_1 = 0.05$	$\theta_2 = 0.001$	109.1	11.7	0.069
	$\theta_2 = 0.0001$	107.5	11.6	0.070
	Metric-based	120.8	11.7	0.068
$\theta_1 = 0.01$	$\theta_2 = 0.1$	120.8	11.7	0.068
	$\theta_2 = 0.01$	106.2	11.6	0.072

our numerical experiments may reach up to 12 % compared to the traditional metric-based approach with the same fixed θ_1, e.g., when $\theta_1 = 0.01$. Regarding the EENS, we observe a decrease of curtailed load when the parameters θ_1 and θ_2 increase. Recall that the Wasserstein radii θ_1 and θ_2 relate to the distributional robustness. This means that, when θ_1 or θ_2 increase, the number of distributions within the set increases, ensuing a potentially more conservative worst-case distribution. This, in turn, impacts the empirical violation probabilities (and therefore, the EENS), because the violation probability ϵ has not the same implication when evaluated for different distributions.

VII. CONCLUSION

This paper introduces an ambiguity set that includes an additional Wasserstein constraint on copula, therefore capturing the whole dependence structure among all uncertain parameters. We develop a generic distributionally robust model that can be applied to any kind of decision-making optimization problems in power systems under uncertainty. In particular, we apply the proposed model to a distributionally robust day-ahead OPF problem. The results show the potential for a significant operational cost saving for the whole system, achieved by taking into account the dependence structure of the uncertain renewable energy sources.

As a potential path for future research, the development of a tool, e.g., using machine learning techniques, that determines *a priori* the optimal values of parameters θ_1 and θ_2 may help the decision-maker efficiently use the proposed model. Finally, it is interesting to apply the outcome of Theorem 1 to various short-term operational and long-term planning decision-making problems under uncertainty in power systems to further illustrate the potential benefits of the proposed model. In particular, we highlight the extension to AC-OPF formulation for the meshed transmission networks.

APPENDIX A
PROOF OF LEMMA 1

Proof. For a given argument value η, the program counts the number of historical observations $\widehat{\xi}_{ki}$, $i \in \{1, ..., N\}$ corresponding to the renewable power unit k under which the renewable power generation is lower than η. The constraint (7b) imposes that the variable z_{ki} takes a non-negative value

whenever η is higher than the underlying observation $\widehat{\xi}_{ki}$, and vice versa. However, the value of z_{ki} is restricted to lie within 0 and 1 in (7c). Therefore, the optimal value of z_{ki} will mimic the function $\mathbb{1}_{\eta \geq \widehat{\xi}_{ki}}$. This enables the objective function (7a) to compute the value of the empirical marginal cumulative distribution function. □

APPENDIX B
PROOF OF THEOREM 1

Proof. We depart from the equivalent reformulation of the worst-case expectation problem in (5), given by

$$\min_{\alpha, \beta \geq 0, y_i} \alpha\theta_1 + \beta\theta_2 + \frac{1}{N}\sum_{i=1}^{N} y_i \quad (17a)$$

$$\text{s.t. } y_i \geq \max_{\xi \in \Xi} \ a(x)^\top \xi + b(x) - \alpha\,\mathrm{d}\left(\widehat{\xi}_i, \xi\right) - \beta\,\mathrm{d}_F\left(\widehat{\xi}_i, \xi\right) \ \forall i. \quad (17b)$$

Note that (17) is equivalent to the findings in [19, Theorem 2]. For the sake of completeness, we also provide the proof of assertion (17) in the online companion [23]. Complicating constraint (17b) requires reformulations, as it contains a maximization operator over variable ξ and the distance functions, which are defined in the following. From now on, the proof will focus on the reformulation of (17b). The distance functions are defined using norms, such that

$$\mathrm{d}\left(\widehat{\xi}_i, \widetilde{\xi}\right) = \left\|\widehat{\xi}_i - \widetilde{\xi}\right\| \text{ and } \mathrm{d}_F\left(\widehat{\xi}_i, \widetilde{\xi}\right) = \left\|F\left(\widehat{\xi}_i\right) - F\left(\widetilde{\xi}\right)\right\| \quad (18a)$$

where $F(\widehat{\xi}_i) = \left(F_1\left(\widehat{\xi}_{1i}\right), ..., F_k\left(\widehat{\xi}_{ki}\right), ..., F_{|\mathcal{W}|}\left(\widehat{\xi}_{|\mathcal{W}|i}\right)\right)^\top$ and $F(\widetilde{\xi}) = \left(F_1\left(\widetilde{\xi}_1\right), ..., F_k\left(\widetilde{\xi}_k\right), ..., F_{|\mathcal{W}|}\left(\widetilde{\xi}_{|\mathcal{W}|}\right)\right)^\top$ are vectors in $\mathbb{R}^{|\mathcal{W}|}$. In other words, each component of $\widehat{\xi}_i$ or $\widetilde{\xi}$ is given as argument of its corresponding marginal cumulative distribution function. Note that the resulting vector $F\left(\widehat{\xi}_i\right)$ is a sample of the copula, which can be evaluated *a priori* using (3). Note also that $F\left(\widetilde{\xi}\right)$ is a decision variable related to the variations within the variable copula, which requires further reformulations. Using such distance definitions, (17b) can be recast into

$$y_i \geq \max_{\widetilde{\xi} \in \Xi} \ a(x)^\top \widetilde{\xi} + b(x) - \alpha\left\|\widehat{\xi}_i - \widetilde{\xi}\right\| - \beta\left\|F\left(\widehat{\xi}_i\right) - F\left(\widetilde{\xi}\right)\right\|. \quad (19)$$

To get rid of the norms inside the objective function of the inner maximization problem, we use dual norms ($\|x\|_* = \max_{\|v\| \leq 1} v^\top x$) as

$$y_i \geq \max_{\widetilde{\xi} \in \Xi} a(x)^\top \widetilde{\xi} + b(x) - \alpha \max_{\left\|\zeta_i^{(1)}\right\|_* \leq 1} \zeta_i^{(1)\top}\left(\widehat{\xi}_i - \widetilde{\xi}\right)$$

$$- \beta \max_{\left\|\zeta_i^{(2)}\right\|_* \leq 1} \zeta_i^{(2)\top}\left(F\left(\widehat{\xi}_i\right) - F\left(\widetilde{\xi}\right)\right) \ \forall i. \quad (20)$$

Next, we eliminate the maximization operators on variables $\zeta_i^{(1)}$ and $\zeta_i^{(2)}$ by (*i*) switching the $-\max$ to a $\min-$, (*ii*) moving the resulting minimization operators to the left, (*iii*) merging the min operators over variables $\zeta_i^{(1)}$ and $\zeta_i^{(2)}$, (*iv*) permuting with the max operator over $\widetilde{\xi}$ (which is allowed using the reformulation given by Lemma 1, because the

objective function is linear and the feasible sets are convex and independent), and finally (*v*) dropping the min operator from the right-hand side of the $\geq$ constraint. The variables $\zeta_i^{(1)}$ and $\zeta_i^{(2)}$ are added to the overall set of decision variables and the constraints are added to the overall set of constraints, such that

$$y_i \geq \max_{\widetilde{\xi} \in \Xi} \ a(x)^\top \widetilde{\xi} + b(x) - \alpha \zeta_i^{(1)\top} \left(\widehat{\xi}_i - \widetilde{\xi} \right) \tag{21a}$$
$$- \beta \zeta_i^{(2)\top} \left(F\left(\widehat{\xi}_i\right) - F\left(\widetilde{\xi}\right) \right) \ \forall i$$

$$\left\| \zeta_i^{(1)} \right\|_* \leq 1 \ \forall i \tag{21b}$$

$$\left\| \zeta_i^{(2)} \right\|_* \leq 1 \ \forall i, \tag{21c}$$

is equivalent to (17b). With the changes of variables $\alpha \zeta_i^{(1)} \to \zeta_i^{(1)}$ and $\beta \zeta_i^{(2)} \to \zeta_i^{(2)}$, the constraints become

$$y_i \geq \max_{\xi \in \Xi} \ a(x)^\top \xi + b(x) - \zeta_i^{(1)\top} (\xi_i - \xi)$$
$$- \zeta_i^{(2)\top} (F(\xi_i) - F(\xi)) \ \forall i \tag{22a}$$

$$\left\| \zeta_i^{(1)} \right\|_* \leq \alpha \ \forall i \tag{22b}$$

$$\left\| \zeta_i^{(2)} \right\|_* \leq \beta \ \forall i. \tag{22c}$$

We use Lemma 1 to reformulate the functions $F(\xi)$. This key step allows us to make the link between the variable distribution and the variable copula, by using the empirical marginal cumulative distributions (see Remark 1). The vector $F(\xi)$ now becomes (9) which completes the proof. □

APPENDIX C
SOLUTION APPROACH

We aim to reformulate problem (8), especially constraint (8b). The first step of the reformulation consists in getting rid of the maximization operators within $F(\xi)$. In that direction, we derive the optimality conditions related to the optimization problems in (9) and add them into the constraints of the outer maximization problem. We observe that the strong duality theorem holds for the underlying optimization problems. Therefore, the necessary and sufficient optimality conditions for the k-th element of vector $F(\xi)$ in (9) are composed of *(i)* the primal constraints, *(ii)* the dual constraints, and *(iii)* the strong duality equality. This mathematically translates to the following set of constraints:

$$z_{jki} \left(\xi_k - \widehat{\xi}_{kj} \right) \geq 0 \ \forall j \tag{23a}$$

$$0 \leq z_{jki} \leq 1 \ \forall j \tag{23b}$$

$$\frac{1}{N} + \sigma_{kji} \left(\xi_k - \widehat{\xi}_{kj} \right) - \pi_{kji} \leq 0 \ \forall j \tag{23c}$$

$$\sigma_{kji}, \pi_{kji} \geq 0 \ \forall j \tag{23d}$$

$$\frac{1}{N} \sum_{j=1}^{N} z_{kji} = \sum_{j=1}^{N} \pi_{kji}. \tag{23e}$$

By doing so, the constraint (8b) becomes

$$y_i \geq \left\{ \begin{array}{l} \max\limits_{\xi, z_{kji}, \sigma_{kji}, \pi_{kji}} \ a(x)^\top \xi + b(x) - \zeta_i^{(1)\top} \left(\widehat{\xi}_i - \xi \right) \\[2mm] - \zeta_i^{(2)\top} \left(F\left(\widehat{\xi}_i\right) - \frac{1}{N} \left(\begin{array}{c} \sum\limits_{j=1}^{N} z_{j1i} \\ \vdots \\ \sum\limits_{j=1}^{N} z_{j|\mathcal{W}|i} \end{array} \right) \right) \\[4mm] \text{s.t. } C\xi \leq D \\ z_{jki} \left(\xi_k - \widehat{\xi}_{kj} \right) \geq 0 \ \forall k, j \\ 0 \leq z_{jki} \leq 1 \ \forall k, j \\ \frac{1}{N} + \sigma_{kji} \left(\xi_k - \widehat{\xi}_{kj} \right) - \pi_{kji} \leq 0 \ \forall k, j \\ \sigma_{kji}, \pi_{kji} \geq 0 \ \forall k, j \\ \frac{1}{N} \sum\limits_{j=1}^{N} z_{kji} = \sum\limits_{j=1}^{N} \pi_{kji} \ \forall k \end{array} \right\} \ \forall i, \tag{24a}$$

where the inner maximization operator has been dropped and the optimality conditions have been added to the set of constraints of the outer maximization problem. We write the support as $\left\{ \widetilde{\xi} \in \mathbb{R}^{|\mathcal{W}|} \ \middle| \ C\widetilde{\xi} \leq D \right\}$, where $C \in \mathbb{R}^{2|\mathcal{W}| \times |\mathcal{W}|}$ and $D \in \mathbb{R}^{|\mathcal{W}|}$. We notice that the resulting outer maximization problem contains bilinear terms in the form of $z_{jki}\xi_k$ and $\sigma_{kji}\xi_k$. We use the McCormick relaxation of bilinear terms to restore linearity, such that the constraint can be cast into

$$y_i \geq \left\{ \begin{array}{l} \max\limits_{\xi, z_{kji}, \sigma_{kji}, \pi_{kji}} \ a(x)^\top \xi + b(x) - \zeta_i^{(1)\top} \left(\widehat{\xi}_i - \xi \right) \\[2mm] - \zeta_i^{(2)\top} \left(F\left(\widehat{\xi}_i\right) - \frac{1}{N} \left(\begin{array}{c} \sum\limits_{j=1}^{N} z_{j1i} \\ \vdots \\ \sum\limits_{j=1}^{N} z_{j|\mathcal{W}|i} \end{array} \right) \right) \\[4mm] \text{s.t. } C\xi \leq D \\ t_{jki} - z_{jki}\widehat{\xi}_{kj} \geq 0 \ \forall k, j \\ t_{jki} \geq z_{jki}\xi_k^{\min} \ \forall k, j \\ t_{jki} \geq \xi_k + z_{jki}\xi_k^{\max} - \xi_k^{\max} \ \forall k, j \\ t_{jki} \leq \xi_k + z_{jki}\xi_k^{\min} - \xi_k^{\min} \ \forall k, j \\ t_{jki} \geq z_{jki}\xi_k^{\max} \ \forall k, j \\ 0 \leq z_{jki} \leq 1 \ \forall k, j \\ \frac{1}{N} + v_{jki} - \sigma_{kji}\widehat{\xi}_{kj} - \pi_{kji} \leq 0 \ \forall k, j \\ v_{jki} \geq \sigma_{kji}\xi_k^{\min} \ \forall k, j \\ v_{jki} \geq \overline{V}\xi_k + \sigma_{kji}\xi_k^{\max} - \overline{V}\xi_k^{\max} \ \forall k, j \\ v_{jki} \leq \overline{V}\xi_k + \sigma_{kji}\xi_k^{\min} - \overline{V}\xi_k^{\min} \ \forall k, j \\ v_{jki} \geq \sigma_{kji}\xi_k^{\max} \ \forall k, j \\ \sigma_{kji}, \pi_{kji} \geq 0 \ \forall k, j \\ \frac{1}{N} \sum\limits_{j=1}^{N} z_{kji} = \sum\limits_{j=1}^{N} \pi_{kji} \ \forall k \end{array} \right\} \ \forall i. \tag{25a}$$

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2022.3221348

13

Note that the vectors $\widetilde{\xi}^{\max} \in \mathbb{R}^{|\mathcal{W}|}$ and $\widetilde{\xi}^{\min} \in \mathbb{R}^{|\mathcal{W}|}$ correspond to the maximum and minimum thresholds for the uncertain parameters $\widetilde{\xi}$. In the case of renewable energy sources, $\widetilde{\xi}^{\min}$ and $\widetilde{\xi}^{\max}$ are related to zero and the installed capacity, respectively. We now focus on getting rid of the outer maximization operator. The final step of the reformulation is to dualize (25), such that the problem becomes a minimization and the operator can be dropped from the constraint. The following problem is a tractable reformulation of (8):

$$\min_{\Pi} \quad \alpha\theta_1 + \beta\theta_2 + \frac{1}{N}\sum_{i=1}^{N} y_i \tag{26a}$$

$$\text{s.t.} \quad y_i \geq b\left(x\right) - \zeta_i^{(1)\top}\widehat{\xi}_i - \zeta_i^{(2)\top}F\left(\widehat{\xi}_i\right) + \mu^{(0)\top}d$$

$$+ \sum_{k=1}^{|\mathcal{W}|}\sum_{j=1}^{N}\left(\mu_{kji}^{(3)}\widetilde{\xi}_k^{\max} - \mu_{kji}^{(4)}\widetilde{\xi}_k^{\min} + \mu_{kji}^{(6)} + \frac{1}{N}\mu_{kji}^{(7)} + \right.$$

$$\left. \mu_{kji}^{(9)}\overline{V}\xi_k^{\max} - \mu_{kji}^{(10)}\overline{V}\xi_k^{\min}\right) \quad \forall i \tag{26b}$$

$$a_k\left(x\right) + \zeta_{ik}^{(1)} - C_k^\top\mu^{(0)}$$

$$+ \sum_{j=1}^{N}\left(\mu_{kji}^{(3)} + \mu_{kji}^{(4)} + \overline{V}\mu_{kji}^{(10)} - \overline{V}\mu_{kji}^{(9)}\right) = 0 \ \forall k, i \tag{26c}$$

$$\frac{1}{N}\zeta_{ik}^{(2)} - \mu_{kji}^{(1)}\widehat{\xi}_{kj} - \mu_{kji}^{(2)}\widetilde{\xi}_k^{\min} - \mu_{kji}^{(3)}\widetilde{\xi}_k^{\max}$$

$$+ \mu_{kji}^{(4)}\widetilde{\xi}_k^{\min} + \mu_{kji}^{(5)}\widetilde{\xi}_k^{\max} - \mu_{kji}^{(6)} + \frac{1}{N}\lambda_k \leq 0 \ \forall k, j, i \tag{26d}$$

$$\mu_{kji}^{(1)} + \mu_{kji}^{(2)} + \mu_{kji}^{(3)} - \mu_{kji}^{(4)} - \mu_{kji}^{(5)} = 0 \ \forall k, j, i \tag{26e}$$

$$-\mu_{kji}^{(7)} + \mu_{kji}^{(8)} + \mu_{kji}^{(9)} - \mu_{kji}^{(10)} - \mu_{kji}^{(11)} = 0 \ \forall k, j, i \tag{26f}$$

$$\mu_{kji}^{(7)} - \lambda_k = 0 \ \forall k, j, i \tag{26g}$$

$$\mu_{kji}^{(7)} - \mu_{kji}^{(8)}\xi_k^{\min} - \mu_{kji}^{(9)}\xi_k^{\max} + \mu_{kji}^{(10)}\xi_k^{\min} + \mu_{kji}^{(11)}\xi_k^{\max} \leq 0 \tag{26h}$$

$$\left\|\zeta_i^{(1)}\right\|_* \leq \alpha \ \forall i \tag{26i}$$

$$\left\|\zeta_i^{(2)}\right\|_* \leq \beta \ \forall i. \tag{26j}$$

References

[1] The North American Electric Reliability Corporation (NERC), "Special report: Potential reliability impacts of emerging flexible resources," 2010, [Online]. Available: https://www.nerc.com/files/IVGTF_Task_1_5_Final.pdf.

[2] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets*. Springer, 2013.

[3] J. M. Morales, A. J. Conejo, and J. Perez-Ruiz, "Economic valuation of reserves in power systems with high penetration of wind power," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 900–910, 2009.

[4] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 52–63, 2013.

[5] M. Lubin, Y. Dvorkin, and S. Backhaus, "A robust approach to chance constrained optimal power flow with renewable generation," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3840–3849, 2016.

[6] C. Ordoudis, V. A. Nguyen, D. Kuhn, and P. Pinson, "Energy and reserve dispatch with distributionally robust joint chance constraints," *Oper. Res. Lett.*, vol. 49, no. 3, pp. 291–299, 2021.

[7] D. Kuhn, P. Mohajerin Esfahani, V. A. Nguyen, and S. Shafieezadeh-Abadeh, "Wasserstein distributionally robust optimization: Theory and applications in machine learning," *INFORMS Tut. Oper. Res. Manage. Sci. Age Anal.*, pp. 130–166, 2019.

[8] J. M. Keynes, *A Treatise on Probability*. MacMillan & Co., 1921, London, UK.

[9] 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, 2017.

[10] A. Arrigo, C. Ordoudis, J. Kazempour, Z. De Grève, J.-F. Toubeau, and F. Vallée, "Wasserstein distributionally robust chance-constrained optimization for energy and reserve dispatch: An exact and physically-bounded formulation," *Eur. J. Oper. Res.*, vol. 296, no. 1, pp. 304–322, 2022.

[11] B. Li, R. Jiang, and J. L. Mathieu, "Integrating unimodality into distributionally robust optimal power flow," *TOP*, vol. 30, pp. 594–617, 2022.

[12] A. Esteban-Pérez and J. M. Morales, "Partition-based distributionally robust optimization via optimal transport with order cone constraints," *4OR-Q. J. Oper. Res.*, 2021.

[13] 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, 2018.

[14] A. Arrigo, J. Kazempour, Z. De Grève, J.-F. Toubeau, and F. Vallée, "Enhanced Wasserstein distributionally robust OPF with dependence structure and support information," in *IEEE PowerTech 2021 Conference*, 2021, Madrid, Spain.

[15] R. Gao and A. Kleywegt, "Data-driven robust optimization with known marginal distributions," 2017. [Online]. Available: https://faculty.mccombs.utexas.edu/rui.gao/copula.pdf

[16] G. C. Pflug and M. Pohl, "A review on ambiguity in stochastic portfolio optimization," *Set-Valued Var. Anal.*, vol. 26, p. 733–757, 2018.

[17] R. B. Neslen, *An Introduction to Copulas*. Springer, 1999, New York, USA.

[18] A. Sklar, "Fonctions de répartition à n dimensions et leurs marges," *Publications de l'institut de statistiques de l'université de Paris*, vol. 8, pp. 229–231, 1959.

[19] R. Gao and A. J. Kleywegt, "Distributionally robust stochastic optimization with dependence structure," 2017. [Online]. Available: https://arxiv.org/abs/1701.04200

[20] 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, 2007.

[21] Y. Guo, K. Baker, E. Dall'Anese, Z. Hu, and T. H. Summers, "Data-based distributionally robust stochastic optimal power flow – Part I: Methodologies," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1483–1492, 2019.

[22] L. V. Kantorovich and G. S. Rubinshtein, "On a space of totally additive functions," *Vestnik Leningradskogo Universiteta*, vol. 13, pp. 52–59, 1958.

[23] A. Arrigo, J. Kazempour, Z. De Grève, J.-F. Toubeau, and F. Vallée, "Online companion: Embedding dependencies between wind farms in distributionally robust optimal power flow," 2021, available: https://doi.org/10.5281/zenodo.6627137.

[24] R. D. Christie, B. F. Wollenberg, and I. Wangensteen, "Transmission management in the deregulated environment," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 170–195, 2000.

[25] K. Baker, "Solutions of DC OPF are never AC feasible," *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, p. 264–268, 2021.

[26] 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, July 2018.

[27] ARPA-E, *GO Competition*, 2022. [Online]. Available: https://gocompetition.energy.gov/about-competition

[28] D. Kuhn, W. Wiesemann, and A. Georghiou, "Primal and dual linear decision rules in stochastic and robust optimization," *Math. Prog.*, vol. 130, no. 1, pp. 177–209, 2011.

[29] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Trans. on Power Del.*, vol. 4, no. 1, pp. 725–734, January 1989.

[30] 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.

[31] C. Ordoudis, P. Pinson, J. M. Morales, and M. Zugno, "An updated version of the IEEE RTS 24-bus system for electricity market and power system operation studies," 2016. [Online]. Available: http://orbit.dtu.dk/files/120568114/An

Adriano Arrigo received the Master degree and the Ph.D. degree in electrical engineering from the University of Mons in 2018 and 2022, respectively. He is currently a postdoctoral research fellow in the Power Systems and Markets Research group of the University of Mons. His research interests include risk management and decision-making in electricity systems and economics.

Jalal Kazempour (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Castilla-La Mancha, Ciudad Real, Spain, in 2013. He is currently an Associate Professor with the Department of Wind and Energy Systems, Technical University of Denmark, also heading the "Energy Markets and Analytics" Section. His research interests include intersection of multiple fields, including power and energy systems, electricity markets, optimization, game theory, and machine learning.

Zacharie De Grève (Member, IEEE) received the Electrical and Electronics Engineering degree from the Faculty of Engineering, University of Mons, Mons, Belgium, in 2007. He was a Research Fellow of the Belgian Fund for Research (F.R.S/FNRS) until 2012, when he got the Ph.D. degree in electrical engineering from the University of Mons, where he is currently an Associate Professor with the Electrical Power Engineering Unit. His main research interests deal with the application of Machine Learning and Operations Research to electric power systems, and energy systems more generally. He also develops an expertise in computational electromagnetics.

Jean-François Toubeau (Member, IEEE) received the Master degree and the Ph.D. degree in electrical engineering from the University of Mons (Belgium) in 2013 and 2018, respectively. He is currently a Senior Researcher with the University of Leuven (KU Leuven) Energy Institute, TME Branch (energy conversion). His research interests include machine learning and decision-making in modern power systems.

François Vallée (Member, IEEE) received the degree in civil electrical engineering and the Ph.D. degree in electrical engineering from the Faculty of Engineering, University of Mons, Belgium, in 2003 and 2009, respectively. He is currently vice-dean of the faculty of engineering and Professor and leader of the "Power Systems and Markets Research Group" at the University of Mons. His Ph.D. work has been awarded by the SRBE/KBVE Robert Sinave Award in 2010. His research interests include PV and wind generation modeling for electrical system reliability studies in presence of dispersed generation.