

# Polyhedral Predictive Regions for Power System Applications

Faranak Golestaneh , Pierre Pinson, *Senior Member, IEEE*, and Hoay Beng Gooi , *Senior Member, IEEE*

**Abstract**—Despite substantial improvement in the development of forecasting approaches, conditional and dynamic uncertainty estimates ought to be accommodated in decision-making in power system operation and market, in order to yield either cost-optimal decisions in expectation, or decision with probabilistic guarantees. The representation of uncertainty serves as an interface between forecasting and decision-making problems, with different approaches handling various objects and their parameterization as input. Following substantial developments based on scenario-based stochastic methods, robust and chance-constrained optimization approaches have gained increasing attention. These often rely on polyhedra as a representation of the convex envelope of uncertainty. In this paper, we aim to bridge the gap between the probabilistic forecasting literature and such optimization approaches by generating forecasts in the form of polyhedra with probabilistic guarantees. For that, we see polyhedra as parameterized objects under alternative definitions (under  $L_1$  and  $L_\infty$  norms), the parameters of which may be modeled and predicted. We additionally discuss assessing the predictive skill of such multivariate probabilistic forecasts. An application and related empirical investigation results allow us to verify probabilistic calibration and predictive skills of our polyhedra.

**Index Terms**—Probabilistic forecasting, box uncertainty sets, polyhedron, robust optimization, chance-constrained optimization.

## I. INTRODUCTION

THE variability and limited predictability of renewable power generation have introduced new challenges into power systems. With a large-scale uncertain generation, in order to reduce the gap between fail-safe and economical solutions of operational problems, advancement in two areas is essential. First, the development of highly scalable optimization techniques capable of accommodating considerable degree of uncertainty is required. Second, it is of utmost importance to

develop adequate and high-quality representations of the uncertainties involved to be used as input to the aforementioned optimization techniques [1], [2].

Practitioners mostly use so-called deterministic or point forecasts as input to decision making today. These comprise single-valued prediction for the future realization of a variable of interest, disregard of the actual range of potential outcomes [3], [4]. However, the solution to an optimization problem in a deterministic setup may be highly sensitive to small perturbations of uncertain quantities. Hence, ignoring uncertainty can result in suboptimal or infeasible solutions in practice [4].

Uncertainty forecasts can be represented in various forms such as scenario, probabilistic and ramp forecasts [5]. Since wind and Photovoltaic (PV) power both show high cross-correlation in time and space, more recently, forecasting spatial/temporal scenarios has been of interest. For example, temporal uncertainty forecast is a key requirement for multi-period operational problems such as unit-commitment and state of charge of energy storage [6]. Stochastic programming as one of the most common optimization techniques in power systems applications uses scenarios as inputs to find optimal solutions in uncertain environments [7], [8]. However, stochastic programming holds a number of pitfalls in a practical context including heavy computational burden and the need for hard-to-obtain probability distributions [9].

The issues with stochastic programming motivates to move towards more recent approaches to optimization under uncertainty, namely robust, chance-constrained and interval optimization. Recently, these optimization techniques have been deployed in power systems applications [10]–[12]. For these classes of decision-making problems, the required uncertainty representation takes the form of prediction regions rather than scenarios [13]. Robust optimization is a computationally viable methodology providing solutions deterministically immune to any realization of uncertainty within a defined uncertainty set (another term for prediction regions) [1]. Interval optimization derives optimistic and pessimistic solutions based on boundaries of prediction regions [14]. In chance-constrained optimization, the uncertainty sets give a probability guarantee for the coverage of observations from the stochastic process considered [15].

Prediction regions in an univariate case, e.g. modeling uncertainty of a single wind farm at a particular time, can be adequately addressed by prediction intervals [5]. However, when modeling temporal/spatial or multivariable correlations is of interest, prediction regions take the form of multivariate ellipsoids, boxes and polyhedra [13], [16]. We refer to uncertainty sets as

Manuscript received March 10, 2018; revised June 11, 2018; accepted July 15, 2018. Date of publication July 31, 2018; date of current version December 19, 2018. This work was supported by the Energy Innovation Programme Office through the National Research Foundation and Singapore Energy Market Authority under Project LA/Contract NRF2014EWT-EIRP002-005. Paper no. TPWRS-00349-2018. (*Corresponding author: Faranak Golestaneh.*)

F. Golestaneh is with the Australian Energy Market Operator (AEMO), Melbourne, VIC 3000, Australia (e-mail: faranak001@e.ntu.edu.sg).

P. Pinson is with the Department of Electrical Engineering, Technical University of Denmark, 2800 Kongens Lyngby, Denmark (e-mail: ppin@elektro.dtu.dk).

H. B. Gooi is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: ehbgooi@ntu.edu.sg).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TPWRS.2018.2861705

prediction regions to emphasize on the fact that they are predictions in nature. The reader can refer to [1], [15]–[17] for the formulations of the robust and chance-constrained optimization and the applications of multivariate prediction regions.

Although the multivariate prediction regions have been used in several optimization applications, the literature has been almost silent on how to efficiently generate and evaluate them. The parameters of multivariate prediction regions are simply chosen based on assumptions or by trial-and-error without verification of those assumptions in practical applications. Uncertainty sets are constructed based on a Gaussian assumption in [12] for nodal load and in [11], [15] for wind power. The inadequacy of a Gaussian assumption in describing uncertainty of wind and PV power is discussed in [3]. A parameter named uncertainty budget is used to control the size and conservativeness of wind power uncertainty sets in the form of ellipsoids in [18] and in the form of polyhedra in [19]. As a different approach, in [20], convex hulls of spatial/temporal scenarios are defined as prediction regions of wind/PV power. In [21], temporal scenarios are used as inputs to produce multivariate prediction intervals (MPIs) to characterize the dependency of wind power forecast errors over a time horizon.

Robust optimization tends to produce conservative solutions. The conservativeness of a robust solution is directly linked to the size of uncertainty sets [8]. However, controlling size of uncertainty sets is not a trivial task to be determined arbitrarily. As any other type of prediction, uncertainty prediction should provide a certain level of required performance. Similar to the case of univariate probabilistic forecasts, multivariate prediction regions are assessed based on their calibration and sharpness [3]. Calibration is linked to conservativeness and it shows how close the empirical coverage rate of a prediction region is to its nominal one. In contrast to [11], [12], [15], [18]–[20], we emphasize on generating prediction regions with predefined coverage rates. This helps the decision-maker to know in advance what the degree of constraint violation is upon obtaining the solution of the optimization problem. Sharpness relates to how small the spread of uncertainty is for the required probability guarantee. Too large prediction regions increase the decision-making complexity as they give a larger range of likely outcomes.

In [3], we proposed a framework to produce skilled ellipsoidal prediction regions. However, various decision-making problems demand for different forms of uncertainty characterization. For example, the robust counterpart of a linear programming problem with polyhedral uncertainty sets is a linear programming problem while the same with the ellipsoidal uncertainty sets is a Second Order Cone Programming (SOCP) problem. Although SOCP problems are convex and computationally tractable, their nonlinearity can be a practical drawback [17], [22]. Consequently, in this work we focus on multivariate prediction polyhedra. Our underlying motivation is to propose a data-driven approach capable of generating highly skilled prediction polyhedra. We study evaluation methodologies for verification of the proposed methods using real data. Two formulations for prediction polyhedra are developed. In addition, due to recent interests in prediction convex hulls [20], their relevance and limitations are discussed and a verification framework for their quantitative

assessment is developed. Because any forms of multivariate prediction is prone to be affected by outliers, we propose an idea to make convex hulls more robust to outliers. The robustness of the prediction regions to outliers is also examined and compared. All techniques output convex polyhedra and suit the requirements of robust and chance-constrained optimization. Also, theoretically they can be employed for both spatial and temporal uncertainty prediction. Their performances in practice, however, will be assessed over empirical results in Section V. The efficiency of the proposed frameworks is evaluated for wind and PV power. Temporal and spatial polyhedral prediction regions of dimensions 2, 3, 6, 12 and 24 with the probability of 5% to 95% in 5% increments are generated and evaluated.

The rest of the paper is organized as follows: in Section II, the proposed methodology and formulations to generate polyhedral prediction regions are discussed. The proposed skill assessment techniques are provided in Section III. The framework to estimate the parameters of the proposed prediction regions is explained in Section IV. Section V contains the empirical results and finally concluding remarks are given in Section VI.

## II. METHODOLOGY

Due to growing interest in characterizing uncertainty information in forms of polyhedra and multivariate boxes, in this section, four frameworks to produce such prediction geometries are proposed.

### A. Prediction Polyhedra

At every time step  $t$ , one aims at predicting the random variable, e.g. wind/PV power, for future times  $t + 1, t + 2, \dots, t + K$  at  $Z$  contiguous locations. Denote  $\mathbf{X}$  as an uncertain variable of dimension  $D = K \times Z$ ,  $\mathbf{X}_t = [X_{t+1}, \dots, X_{t+D}]$ . Denote  $\mu = E(\mathbf{X})$  as the expected value  $\mathbf{X}$  and  $\Sigma = E[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T]$  as its the covariance matrix. Inspired by [22], we propose the following two formulations for prediction polyhedra.

$$P_{t,\alpha}^1 := \{\mathbf{X} \in \mathbb{R}^D : \|\Lambda_t(\mathbf{x}_t - \mu_t)\|_1 \leq \Gamma_t^\alpha\} \quad (1)$$

$$P_{t,\alpha}^\infty := \{\mathbf{X} \in \mathbb{R}^D : \|\Lambda_t(\mathbf{x}_t - \mu_t)\|_\infty \leq \Delta_t^\alpha\} \quad (2)$$

where  $\alpha$  is the nominal coverage rate of prediction polyhedra,  $\Delta$  and  $\Gamma$  are called scale or robust parameters. With the assumption that  $\Sigma^{-1}$  is a symmetric and positive definite matrix,  $\Lambda$  as the Cholesky decomposition of  $\Sigma^{-1}$  is an upper triangular matrix with positive diagonal elements. For vector  $\mathbf{x} \in \mathbb{R}^{D \times 1}$ ,  $\|\mathbf{x}\|_1$  denotes the first norm as  $\sum_{d=1}^D |x_d|$  and  $\|\mathbf{x}\|_\infty$  denotes the infinity norm given by  $\max_{d=1, \dots, D} |x_d|$ . Henceforth, the upper case letters symbolize random variables while lower case letters express their realizations.

A special case of (1) is when  $\Sigma_t$  is diagonal with diagonal elements equal to 1. Then, (1) is defined as the sum of deviations from  $\mu_t$  and it is controlled by  $\Gamma_t^\alpha$ . If the deviation in one dimension is large, the deviations in the other dimensions are forced to be small to have a sum of maximum  $\Gamma_t^\alpha$ . If  $\Sigma_t$  is diagonal with diagonal elements equal to the standard deviation of the random variable in each dimension, then, the share of

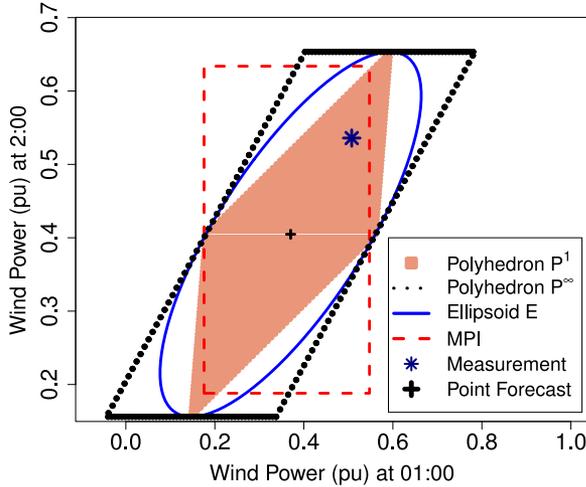


Fig. 1. Typical prediction geometries of dimension two.

deviation in each dimension from the sum of deviations is determined by its standard deviation. If the standard deviation in one dimension is small, it should not deviate a lot from  $\mu_t$  and if the standard deviation in one dimension is large, the random variable in that dimension should consume more of  $\Gamma_t^\alpha$  [22]. When  $\Sigma_t$  is a covariance matrix with non-zero elements, the  $\Gamma_t^\alpha$  is shared between the dimensions based on the standard deviation in each dimension and the correlations between different dimensions.

The polyhedron given by (1) is inscribed in a ellipsoid defined by the following formulation [3], [22].

$$E_{t,\alpha} := (\mathbf{x}_t - \mu_t)^\top \Sigma^{-1} (\mathbf{x}_t - \mu_t) \leq (\Gamma_t^\alpha)^2 \quad (3)$$

The predictive performance of the uncertainty sets in forms  $P^1$ ,  $P^\infty$  is directly linked to how accurate and optimal their predicted parameters are. The parameters to be predicted include a location parameter  $\mu$  (mean vector), a shaping parameter  $\Sigma$  (covariance matrix), and scaling parameters  $\Gamma^\alpha$  and  $\Delta^\alpha$  (being a function of the nominal coverage rate). In robust optimization literature, the scale parameter is commonly known as the uncertainty budget and it controls the conservativeness. It is worth noting that even though the first and the second-order moment information (i.e., mean and covariance) are classically used for Gaussian objects, considering them as a basis for defining polyhedra does not necessarily means one can assume the underlining distribution is Gaussian.

Assuming equal values for  $\Gamma$  and  $\Delta$  in (1)–(3), typical  $P^1$ ,  $P^\infty$  and  $E$  with probability level of 85% are illustrated in Fig. 1. In this figure, the predicted scale parameter is 2.210 while the location  $\mu$ , and shape  $\Sigma$  for  $P^1$ ,  $P^\infty$  and  $E$  are

$$\Sigma = \begin{bmatrix} 0.01762222 & 0.01135601 \\ 0.01135601 & 0.01265258 \end{bmatrix}, \quad \mu = [0.370 \quad 0.405] \quad (4)$$

It is to be emphasized that  $\Gamma$  and  $\Delta$  are not expected to be equal in general. In Section IV, the proposed ideas to determine the correct values for  $\Gamma$  and  $\Delta$  are explained. What Fig. 1 illustrates is that in case  $\Gamma$  and  $\Delta$  take equal values in (1)–(3), how  $P^1$ ,  $P^\infty$  and  $E$  relate to one another. Robust or interval optimization go

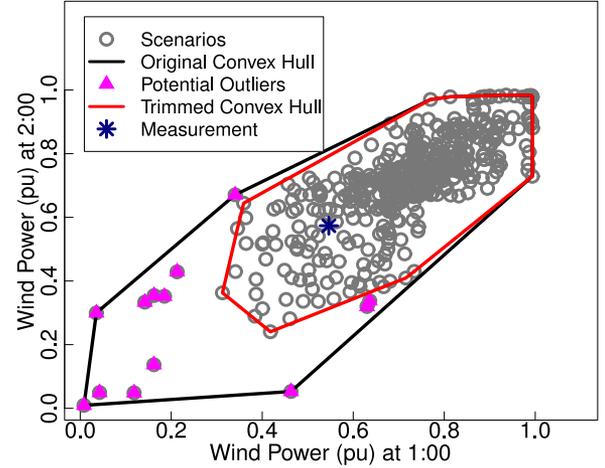


Fig. 2. The convex hull containing all the predicted scenarios. The original convex hull against the trimmed convex hull generated by excluding the potential outliers.

along the faces/edges to find the optimal solution. From Fig. 1, it can be inferred that  $P^1$  and  $P^\infty$  impose similar computational cost in optimization because they actually have equal number of edges/facets with a difference that  $P^1$  tends to be sharper. As shown in Fig. 1, the measurement is included in all  $P^1$ ,  $P^\infty$  and  $E$ . If having many more sample observations, one would expect close to 85% of the observations to be covered by the prediction regions.

### B. Prediction Convex Hulls

The convex hull of a set of points,  $S$ , is the smallest convex set containing all the points. The idea is to find the convex hull of spatial/temporal scenarios [23]. Spatial/temporal scenarios are generated by uniformly sampling from multivariate predictive distributions. At each time  $t$ ,  $S$  scenarios are produced where each scenario is a vector of dimension  $D$ . Among few methods available to find a non-ambiguous and efficient representation of required convex hulls, we use Quickhull algorithm. Quickhull algorithm is fast and efficient in most cases and tends to perform well in practice [24]. Time complexity of this algorithm for most cases is  $O(n \log n)$  and in the worst case is  $O(n^2)$ . Theoretically, Quickhull algorithm can work in high dimensions. For a straightforward explanation and implementation guide of Quickhull algorithm, one can refer to [25]. The original convex hull illustrated in Fig. 2 represents the convex hull of predicted temporal scenarios of wind power for a randomly selected day.

### C. Trimmed Prediction Convex Hulls

As stated in subsection II-B, the predicted spatial/temporal scenarios are produced by uniformly sampling from multivariate predictive distributions. Although in the Monte Carlo-based analysis, all scenarios are considered to come with an equal likelihood of occurrence, some of them might be far from the center of cloud. We label those scenarios as outliers. Outliers are marked in Fig. 2. As can be observed, they grossly impact on the size of prediction convex hulls. Discarding the outliers results

in the trimmed convex hull in Fig. 2 which is much sharper than the original one.

Although there is a wealth of techniques available to detect outliers in univariate datasets, only limited options are at hand for multivariate data. It is to be noted that a multivariate outlier does not necessarily have to be an outlier in any of its univariate coordinates. Mahalanobis distance is one of the most widely used metrics for multivariate outlier identification. Basically, it quantifies how far away a point is from the center of the cloud, taking into account the shape of the cloud as well. Those scenarios with Mahalanobis distance larger than the critical chi-square values at a significance level of 0.001 are labeled as outliers [26].

#### D. Benchmark Method 1: Multivariate Prediction Intervals

The multivariate prediction literature still is in a primitive stage and there are not many data-driven benchmarks available to conduct a comparative study on the performance of the proposed techniques. Among few works available, the adjusted intervals approach is found to be a relevant benchmark [21], [27], [28]. This technique uses the Univariate (marginal) Prediction Intervals (UPIs) and the multivariate scenarios as inputs to generate MPIs. In the adjusted intervals technique, the boundaries of MPIs are determined based on the proportion of the spatio-temporal scenarios which they include. It is based on the assumption that if an interval includes  $\alpha\%$  of the scenarios, it also provides similar coverage rate for the measured trajectories. The MPI with nominal coverage rate  $\alpha\%$  is generated as

- 1) The boundaries of MPI are considered to be equal to the UPI with the same coverage rate.
- 2) For each dimension, the width of the MPI is increased by changing its upper and lower boundaries to the closest upper and lower scenarios, respectively.
- 3) The coverage rate of the MPI is calculated by counting the proportion of the scenarios which are within the boundaries of the MPI.
- 4) If the coverage is less than  $\alpha\%$ , steps 2 and 3 are repeated, otherwise, the boundaries of the MPI are set equal to those determined in the previous step.

For the details of the approach to generate MPIs, the reader is referred to [21]. Typical MPIs are illustrated in Fig. 1.

#### E. Benchmark Method 2: Robust Uncertainty Sets

The most widely used framework for polyhedral uncertainty sets is given in (5) [1], [29].

$$P^R := \{\mathbf{X} \in \mathbb{R}^D : \sum_{d=1}^D \frac{|x_{t+d} - c_{t+d}|}{w_{t+d}} \leq \gamma$$

$$x_{t+d} \in [c_{t+d} - w_{t+d}, c_{t+d} + w_{t+d}]\} \quad \forall t \quad (5)$$

$c_{t+d}$  is the center of the  $P^R$  and  $w_{t+d}$  impacts on the size of the uncertainty sets. Usually,  $c_{t+d}$  is considered to be the point forecast generated for time  $t$  and dimension  $d$  and  $w_{t+d}$  is determined based on the standard deviation of dimension  $d$  [29].  $\gamma$  is called the uncertainty budget and can take any value in the range  $[0, D]$ . When  $\gamma = 0$ , the uncertainty in point

forecasts are ignored. As  $\gamma$  increases, a higher uncertainty level is assumed and a larger total deviation from the point forecasts are considered, leading to more robust and conservative solutions. When  $\gamma = D$ , the uncertainty sets are equal to the entire hyper-rectangles defined in (5). We refer to the region enclosed by the inequality in (5) as robust polyhedrons and to the hyper-rectangles in (5) as robust boxes. Then,  $P^R$  is the intersection of the robust polyhedrons and the boxes in (5).

The robust polyhedron in (5) can be considered as a special case of (1) when  $\Sigma_t$  is diagonal with diagonal elements equal to  $w_{t+d}$ . The uncertainty budget is determined by the user arbitrarily based on his aversion to uncertainty. One does not expect to get uncertainty sets with predefined probability levels based on the common approaches available for determination of the uncertainty budget [18], [22].

### III. PREDICTIVE SKILL ASSESSMENT

The predictive performance of probabilistic forecasts are commonly examined based on their two properties, namely calibration and sharpness [30]. Calibration is a joint property of forecasts and observations, and it is decided based on the statistical consistency between them. Sharpness refers to concentration of forecasts [31]. Following the probability and statistics literature, we refer to calibration as the proximity of the nominal coverage rate of a prediction region to its empirical coverage rate. The coverage rate of a prediction region is the proportion of times that the region contains the materialized events (observations). Similarly, a nominal coverage rate refers to the expected coverage while an empirical coverage represents the empirical coverage of that region calculated based on real data. Sharpness is examined based on the size of prediction regions, e.g. area in dimension two, and volume in higher dimensions. The aim is to generate sharp and concentrated prediction regions subjected to calibration.

#### A. Prediction Polyhedra

As can be observed in Fig. 1, both  $P^1$  and  $P^\infty$  are simple, convex and have few edges. In geometry, based on the definition, each vertex of a  $D$ -dimensional simple polyhedron is adjacent to exactly  $D$  edges. Robust and interval optimization go along the edges of uncertainty sets to find the optimal solution. In general, fewer number of faces is an advantage in the sense that it imposes less computation to optimization.

There is no limitation to represent uncertainty in higher dimensions in the forms of  $P^1$  and  $P^\infty$  as long as the correlation matrix  $\Sigma$  can be predicted and Cholesky decomposition of  $\Sigma^{-1}$  can be calculated. The discussion regarding estimation of the covariance matrix is provided in Section IV. The proposed approach and formulations are competent at generating prediction polyhedra with any desired probability levels.

**Volume:** Since to the best of our knowledge there is no straightforward approach to calculate the volume of  $P^1$  and  $P^\infty$  analytically, in Section V, a Monte Carlo-based technique is explained to estimate their volumes numerically.

**Calibration:** To evaluate calibration of prediction polyhedra, one needs to calculate the empirical coverage of each predicted

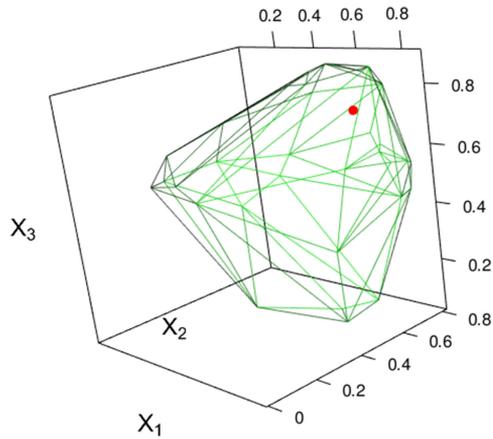


Fig. 3. A typical convex hull generated for PV power data.  $X_1$ ,  $X_2$  and  $X_3$  denote PV power output at zones 1, 2 and 3, respectively. The red dot represents the measurement.

polyhedron and compare that with the corresponding nominal coverage. Let  $\xi_t^{\alpha_i}$  be a binary variable taking 1 if the prediction polyhedron with nominal probability  $\alpha_i$  contains the observed value at time  $t$  and 0 otherwise. Then the empirical coverage is given by

$$\hat{\alpha}_i = \frac{1}{T} \sum_{t=1}^T \xi_t^{\alpha_i} \quad (6)$$

with  $T$  as the length of the evaluation set.  $P^1$  and  $P^\infty$  include the measurement  $\mathbf{x}_t$  if it satisfies (1) and (2), respectively.

### B. Prediction Convex Hulls

As one can notice in Fig. 2, the number of faces in convex hulls is higher than simple prediction polyhedra. This is more noticeable in higher dimensions as shown in Fig. 3. A higher number of faces/edges imposes a higher computation into optimization because as mentioned before robust or interval optimization go along the faces/edges to find the optimal solution. In addition, a major limitation of the prediction convex hulls is that they cannot be generated for predefined nominal coverage rates. They just represent the smallest convex region including the predicted scenarios. When verifying them on real measurements, they can show any empirical coverage, ranging from zero to one. Prediction convex hulls comparing to  $P^1$  and  $P^\infty$  have the complexity of generating multivariate scenarios as their inputs first. In addition, it might be required to use more scenarios to generate them as the dimension increases.

**Volume:** One advantage of convex polyhedra is that their volumes can be calculated by subdividing them into smaller pieces. To do that a common approach is by triangulation methods where the polyhedron is decomposed into simplices. A simplex is a generalization of triangle to arbitrary dimensions. The volume of simplices can easily be computed. Eventually, the volume of polyhedron is computed by summing up the volumes of all simplices [24], [32].

**Calibration:** In convex geometry, given points  $C = \{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_S\}$ , the point  $\theta_1 \hat{\mathbf{x}}_1 + \theta_2 \hat{\mathbf{x}}_2, \dots, \theta_S \hat{\mathbf{x}}_S$  is called

their convex combination if  $\theta_i \geq 0, i = 1, 2, \dots, S$  and  $\sum_{i=1}^S \theta_i = 1$ . Therefore, a convex combination of points can be viewed as a weighted average of the points, with  $\theta$  as the weight of each point in the mixture. The convex hull of set  $C$  ( $\text{conv } C$ ) contains the arbitrary point  $\mathbf{y}$ , if  $\mathbf{y}$  is a convex combination of  $C$  [33]. We use this definition to identify if a prediction convex hull generated for time  $t$  includes the measurement recorded at the same time.  $\mathbf{y} \in C$  if there is a solution for the following linear programming problem

$$\arg_{\theta} \min e^T \theta \quad (7)$$

subject to

$$\begin{aligned} A\theta &= \mathbf{y} \\ B\theta &= 1 \\ \theta &\geq 0 \end{aligned} \quad (8)$$

with  $\mathbf{y} \in \mathbb{R}^D$ ,  $e \in \mathbb{R}^D$  arbitrary cost vector,  $B = (1, \dots, 1) \in \mathbb{R}^S$ ,  $A = (\mathbf{x}_1, \dots, \mathbf{x}_S) \in \mathbb{R}^{D \times S}$  and  $\theta \in \mathbb{R}^S$ .

Following (7) and (8), one can determine if an arbitrary point  $\mathbf{y}$  is inside  $\text{conv } C$  directly with no need to generate  $\text{conv } C$  first. The observed coverage of prediction convex hulls can be computed according to (6) once the inclusion of each  $\mathbf{y}$  in the evaluation set is examined.

### C. Multivariate Prediction Intervals

Similar to the prediction convex hulls, the MPIs have also the complexity of generating multivariate scenarios as their inputs. While theoretically, scenarios and consequently MPIs can be generated for arbitrary dimensions, the performance of the MPIs in various dimensions is discussed in Section V.

**Calibration:** For MPIs, the empirical coverage rate is computed by counting the number of measured scenarios which fully lie within their boundaries [21].

**Volume:** The volume  $V_t$  of a MPI at time  $t$  is calculated as

$$V_t = \prod_i (h_{i,t} - l_{i,t}) \quad \forall t \quad (9)$$

with  $h_{i,t}$  and  $l_{i,t}$  as the upper and lower bounds of the MPI at dimension  $i$ , respectively.

## IV. PARAMETER ESTIMATION

The parameters of  $P^1$  and  $P^\infty$  are determined as

**$u_t$ :** The location parameter  $u_t$  is the center of prediction polyhedra and is considered to be the point forecasts for the multivariate random variable  $\mathbf{X}$  at time  $t$ . Denote  $\hat{\mathbf{x}}_t = [\hat{x}_{1,t} \ \hat{x}_{2,t} \ \dots \ \hat{x}_{D,t}]$ , with  $\hat{x}_{i,t}, \forall i$  as the point forecast for time  $t$  and dimension  $i$  where  $\hat{x}_{i,t}$  for each dimension is generated independently. We refer to  $\hat{\mathbf{x}}$  as predictions and  $\mathbf{x}$  as the measurement or the materialized trajectory.

**$\Sigma_t$ :** The  $\Sigma_t$  is defined as the covariance matrix of point forecast errors estimated using data up to time  $t$ . We suggest to use the Dynamic Conditional-Correlation-GARCH (GARCH-DCC) technique to predict the covariance matrix [34]. In econometrics literature, GARCH-DCC has been widely implemented

and is shown to be capable of estimating time-varying covariance matrices [34], [35]. GARCH-DCC is mostly suitable for those random processes like forecast errors of renewable power generation for which the covariance matrix changes noticeably over time. In case the random process presents a slow-moving covariance matrix, rolling historical correlations and exponential smoothing as less complicated techniques can be deployed [36]. GARCH-DCC is known to be efficient at estimation of potentially very large correlation matrices as in this model the number of parameters to be estimated is independent of the correlated time series [34]. As stated in subsection III-A, if the covariance matrix can be estimated and the Cholesky decomposition of  $\Sigma^{-1}$  can be calculated,  $P^1$  and  $P^\infty$  can be generated for arbitrary dimensions.

$\Gamma_t^\alpha$ : In [3], for ellipsoidal prediction regions, we proposed an approach to find the optimal scale parameters by making a compromise between volume of the ellipsoids and their calibration. To the best of our knowledge, for  $P^1$  and  $P^\infty$  polyhedra, there is no straightforward closed form formulation to calculate the volume. Therefore, we propose a data-driven technique to find the minimum scale parameter which provides the required coverage rate over the most recent historical data. The scale parameter is updated whenever new measurements are received. In the proposed method, a window of size  $\omega$  of the most recent measurements, point forecasts and predicted Cholesky decomposition of covariance matrices are input to the following equation and  $\Upsilon$  for those values is calculated.

$$\Upsilon_i = \|\Lambda_i(\mathbf{x}_i - \mu_i)^\top\|_1 \quad i = t - \omega, \dots, t - 2, t - 1 \quad (10)$$

where  $\mu_i$  is  $\hat{\mathbf{x}}_i$  and  $\mathbf{x}_i \forall i$  are the measured trajectories. Then,  $\Upsilon_i, \forall i$  are sorted ascending. For the desired probability level  $\alpha$ ,  $\Gamma_t^\alpha$  is considered as the  $N^{th}$  smallest  $\Upsilon_i, \forall i$  or in other words the  $N^{th}$  element of the sorted vector, where  $N$  is

$$N = \text{round}(\omega \times \alpha) \quad (11)$$

with  $\text{round}(x)$  as a function which returns the closest integer to  $x$ . Following the proposed method,  $\Gamma_t^\alpha$  is updated for each  $t$  on a rolling base. This technique is based on this expectation that if prediction polyhedra envelop a window of most recent historical data, they should present a similar coverage for the future observations. After obtaining  $u_t, \Sigma_t$  and  $\Gamma_t^\alpha, P_{t,\alpha}^1$  is readily available.

$\Delta_t^\alpha$ : The  $\Delta_t^\alpha$  can be estimated similar to  $\Gamma_t^\alpha$  as explained above, with the only difference that  $\Upsilon_i, \forall i$  are calculated as

$$\Upsilon_i = \|\Lambda_i(\mathbf{x}_i - \mu_i)^\top\|_\infty \quad i = t - \omega, \dots, t - 2, t - 1 \quad (12)$$

For techniques to generate spatial/temporal scenarios, MPIs and convex hulls, the reader is referred to [23], [21] and [25], respectively.

## V. EMPIRICAL RESULTS

### A. Data Description and Models Setup

In order to evaluate the proposed forecasting frameworks, the wind power and PV power datasets provided for the Global Energy Forecasting Competition (GEFCOM) 2014 are used here.

The datasets are available online [37]. We use wind power data to predict temporal dependency and PV power data to study spatial dependency.

The wind power dataset includes wind power measurements of 10 wind farms in Australia. The data for farm three is used here for analysis. The data includes four explanatory variables which are zonal and meridional wind components forecasts at two heights, 10 and 100 m above ground level provided by the European Centre for Medium-range Weather Forecasts (ECMWF). Weather forecasts are issued every day at midnight. The resolution of data is of one hour and forecast horizons are 1- to 24-hour ahead. Data for January 2012 to the end of April 2013 is used to train the models while the out of sample subset covers May 2013 to December 2013.

PV power data includes 12 independent variables as the output of Numerical Weather Predictions (NWP) used as the predictors and PV power generation as the predictand. The available data covers the period from April 2012 to the end of June 2014 for three contiguous zones. Data for April 2012 to the end of May 2013 is used to train the model and the evaluation subset covers data from June 2013 to the end of May 2014. Analysis are carried out to predict the simultaneous stochastic behavior of PV power at three zones at 12:00 pm for spatial dependency studies. Power measurements are normalized by the nominal capacity of their corresponding generation unit.

A support vector machine (SVM) from package “e1071” in R is used to generate wind/PV power point forecasts. The parameters of SVM, i.e. cost of constraints violation and gamma parameter for kernels are tuned based on 5-fold cross-validation using the tuning functions available in the same package. The SVM model yields 15.43%, 14.1% root mean square error for wind and PV power (12:00 pm only), respectively. We tried ELM [38], linear regression and a combination of k-nearest neighbors and SVM [3] and found SVM more accurate with regard to the root mean square error for the data used in the study. However, using SVM is just a suggestion and it can be replaced with any other deterministic forecast method which provides a better accuracy. Because all wind farms are adjacent to each other, the weather forecasts available for the first six wind farms are used as the explanatory variables to generate forecasts for farm 3. The covariance matrices are predicted using GARCH-DCC functions from “rmgarch” package in R. Univariate quantiles with the nominal probability 2.5% to 97.5% in 2.5% increments are produced by the quantile regression from package “quantreg” in R. 500 scenarios [23] are generated as the inputs for the adjusted interval technique. The upper limit of intervals is considered to be 99.5% quantile given by the quantile regression and the lower limit is considered to be zero.

The proposed  $P^1$  and  $P^\infty$  involve three different forecasts, i.e. point forecast, covariance matrix forecast and the scale parameter forecast. We used SVM as a batch and off-line learning algorithm which is trained only once using the training data. The GARCH-DCC is implemented as a semi online learning as the model is re-estimated every 20 periods. The re-estimation period is determined based on a simple 3-fold cross-validation approach. The scale parameter is re-estimated online for every new measurement as it is explained in Section IV. A 3-fold

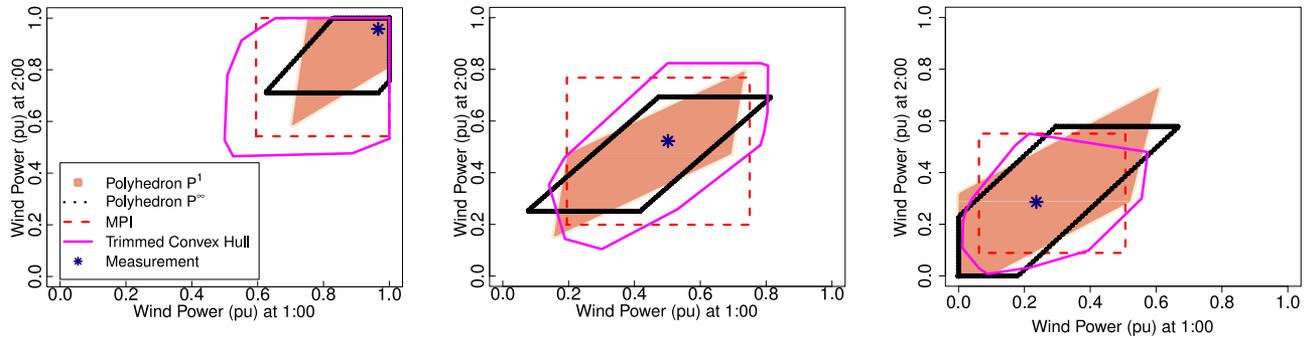


Fig. 4. Visual comparison of polyhedral prediction regions of dimension 2 for wind power data.

cross-validation approach is conducted to find the optimal value of  $\omega$  in (10) and (12) and it is found to be 60.

Polyhedral prediction regions of dimensions 2, 6, 12 and 24 for wind power data are generated and evaluated. Dimension 2 includes wind power data at 01:00 am and 2:00 am. Dimension 6 covers 1- to 6-hour head predictions from 01:00 am to 6:00 am. Dimension 12 represents data from 01:00 am to 12:00 pm and dimension 24 includes all 24 hourly lead times from 01:00 am to 24:00 midnight. Polyhedral prediction regions produced for PV power data are of dimension 3, describing the correlated uncertainty of PV power at three zones under the study at 12:00 pm. Throughout this section, all the analysis in dimension 3 are based on spatial prediction regions produced for PV power data while the results provided in other dimensions are based on temporal prediction regions of wind power.

### B. Performance Evaluation

Fig. 4 shows the polyhedral prediction regions for three randomly selected days for out of sample data. The regions are limited to the feasible range of normalized wind power data [0, 1]. Comparing MPIs with  $P^1$ ,  $P^\infty$  and convex hulls, one can notice that the later ones present a correlated pattern between generation at two successive hours while MPIs show a uniform relation between them. All polyhedra have a fairly reasonable size and follow the variations in wind power generation.

In the following, we will compare the various prediction regions in dimensions  $D \geq 2$ . The following simulation results suggest that both  $P^1$  and  $P^\infty$  have better predictive skill than MPIs in terms of both calibration and sharpness (volume). In addition,  $P^1$  tends to be sharper and less conservative than  $P^\infty$ . It is to be noted that one should expect to see more improvements in the area of verification of such forecasts in the future. We still have a minimum sound basis here to analyze our forecasts and conclude.

1) *Calibration*: Fig. 5 reports the deviations between empirical coverage rate and nominal coverage rate of polyhedral prediction regions for dimensions 2, 3, 6 and 24. The nominal coverage rates ranging from 0.05 to 0.95 in 0.05 increments are included in the figure. As it is expected, UPIs fail to capture the dependent and correlated uncertainty of wind/PV power output over successive hours and at adjacent locations. The calibration and reliability of UPIs decline as the dimension increases. The

calibration of MPIs is also woefully inadequate. The  $P^1$  and  $P^\infty$  polyhedra maintain a fairly stable calibration in all dimensions and for all nominal coverage rates.

Convex hulls are not included in Fig. 5 because as discussed in Section III, they do not provide prediction regions with predetermined nominal coverage rates. When using prediction convex hulls, one expects to get the smallest convex region with the highest probability guarantee. The untrimmed temporal convex hulls return 90.4%, 56%, 10% and almost 0% empirical coverage rates in dimensions 2, 6, 12 and 24 for wind power, respectively. The spatial prediction convex hulls contain 90% of the PV power measurements. Our empirical results suggest that the prediction convex hulls perform poorly in higher dimensions. We produced temporal prediction convex hulls of dimension 4 for wind power data (1:00 am to 4:00 am) and obtained 72% empirical coverage rate. Thus, we do not recommend prediction convex hulls in dimensions higher than 4. Additionally, although it is straightforward and computationally efficient to compute calibration of convex hulls following (7) and (8), the algorithm to find convex hulls themselves becomes very slow for dimensions higher than 8 and it does not converge in dimensions more than 9.

2) *Volume*: For the bounded random variables, the size of the prediction polyhedra is determined by the intersection of two polyhedra. The first one is the prediction polyhedron itself and the second one is formed by the feasible range of the random variable. For the case of wind/PV power, the second polyhedron is a hyper-cube with edges of length equal to the maximum capacity of generation unit. Because there is no simple formulation to calculate the intersection analytically, we use a Monte Carlo-based method for estimation of the volume of prediction regions [3]. The idea is to generate  $N'$  random samples in the feasible range and then calculate the proportion of those points which lie in the prediction polyhedron. The volume of prediction polyhedra  $V^P$  is calculated as

$$V^P = N''V^c / N' \quad (13)$$

with  $N''$  as the number of  $D$ -dimensional points enveloped by the prediction polyhedron and  $V^c$  is the volume of the bounded hyper-cube. Fig. 6 illustrates the size of the proposed prediction polyhedra in comparison with MPI for ten randomly selected days from the evaluation data. The prediction polyhedra of sizes

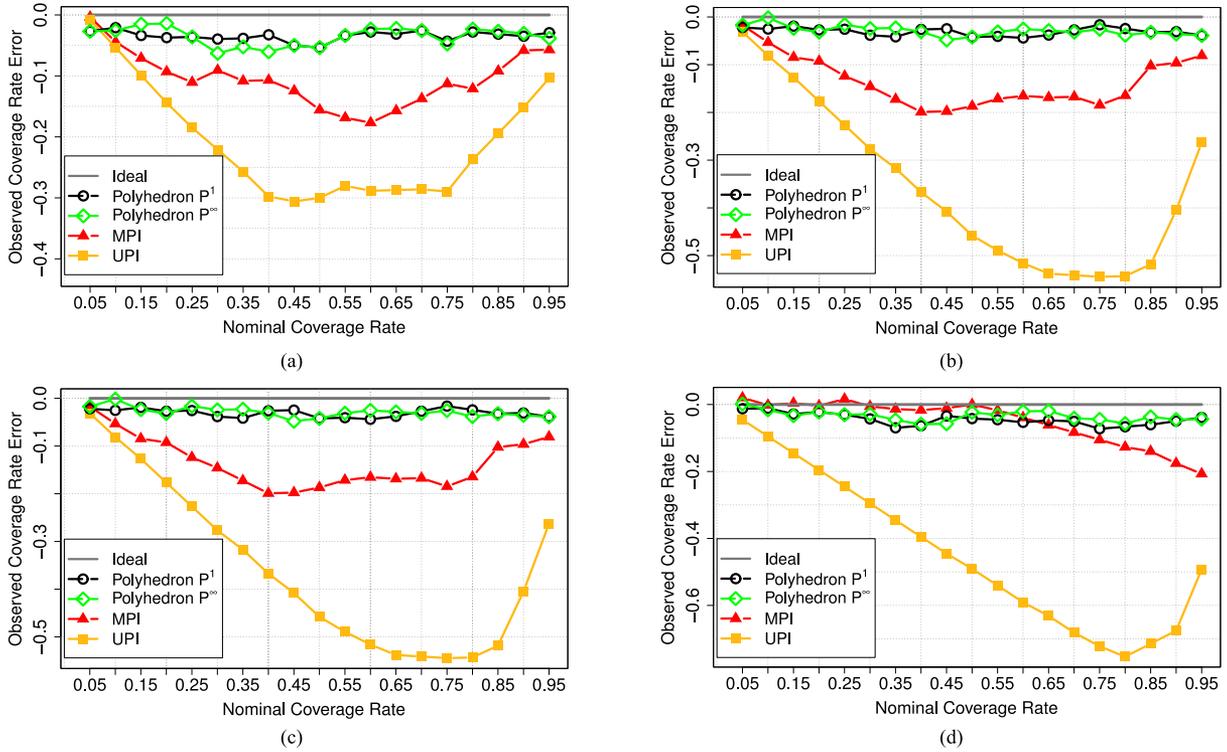


Fig. 5. The difference between empirical coverage and nominal coverage rates (Empirical-Nominal). (a) Temporal polyhedra, dimension 2. (b) Spatial polyhedra, dimension 3. (c) Temporal polyhedra, dimension 6. (d) Temporal polyhedra, dimension 24.

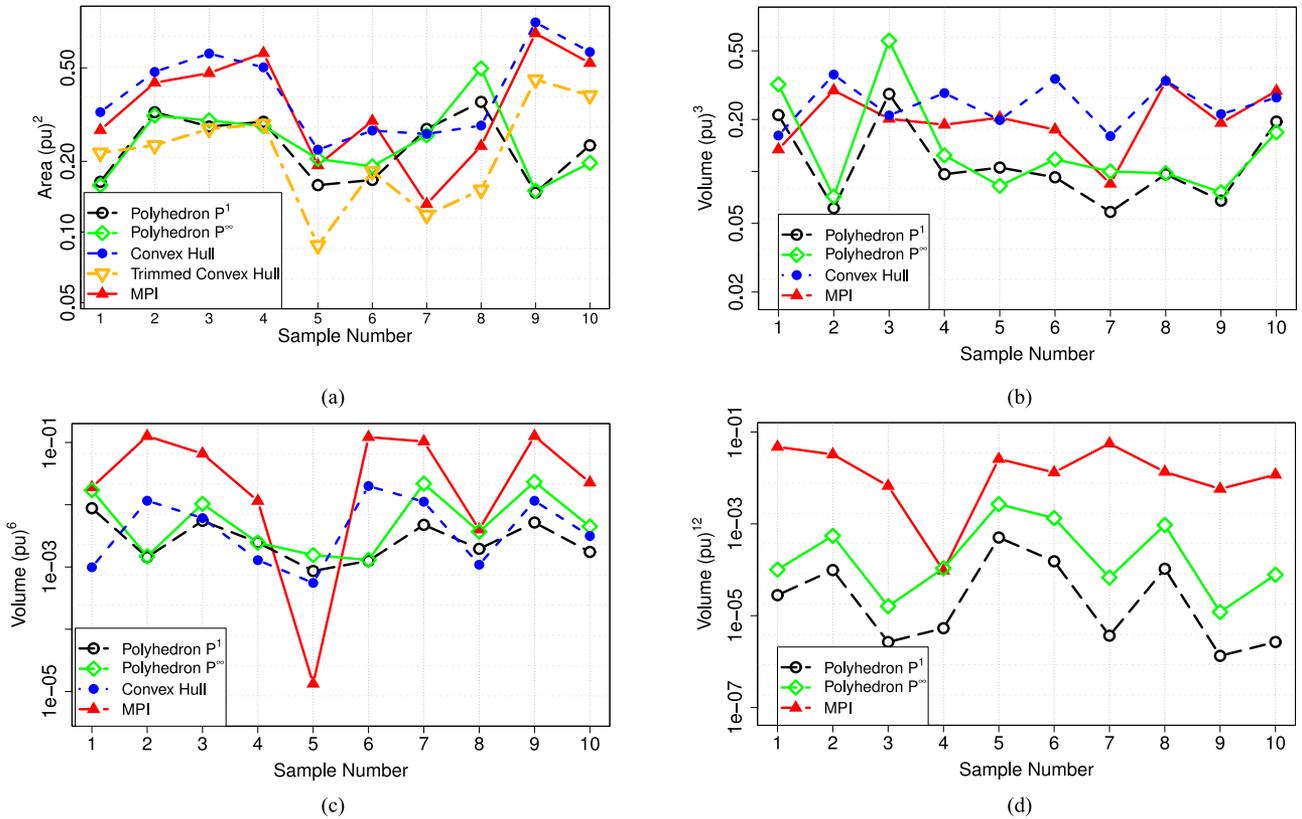


Fig. 6. Estimated volume of polyhedral prediction regions with nominal coverage rates (a) 95%, (b) 90%, (c) 85% and (d) 80%. (a) Dimension 2. (b) Dimension 3. (c) Dimension 6. (d) Dimension 12.

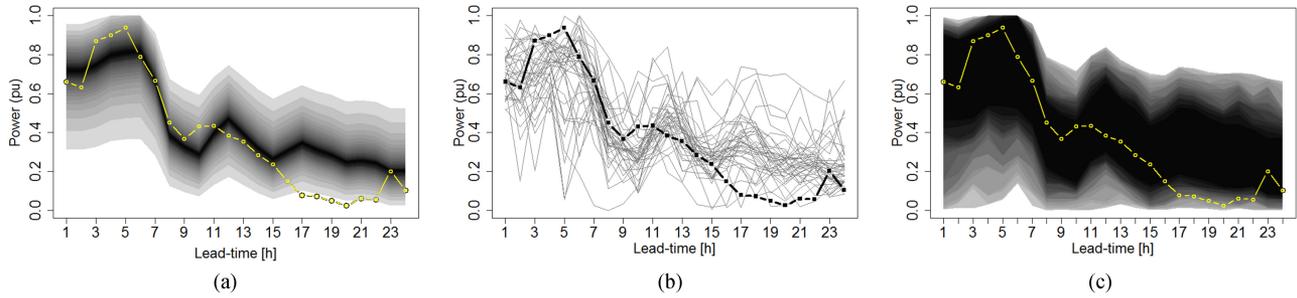


Fig. 7. (a) PV observations (yellow color curves) along with 19 univariate prediction intervals with coverage ranging from 0.05 to 0.95 in 0.05 increments (from the darkest to the lightest), (b) PV observations (dark black color curves) along with 40 generated space-time trajectories (gray color curves), (c) PV power observations (yellow color curves) along with 19 MPIs with nominal coverage ranging from 0.05 to 0.95 in 0.05 increments (from the darkest to the lightest).

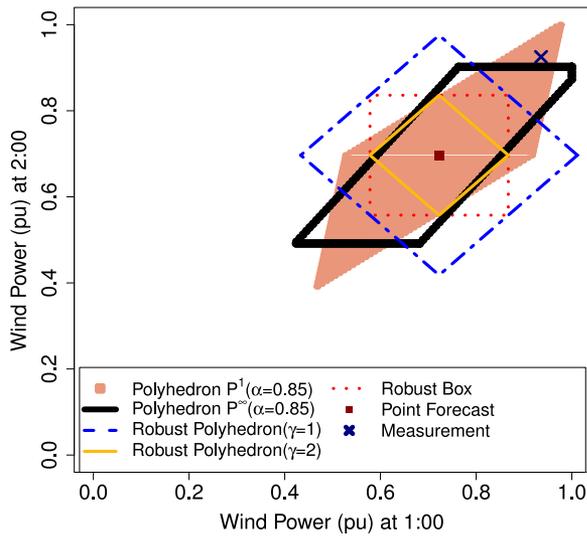


Fig. 8. Temporal scenarios predicted for a randomly selected day from the wind power data. Few scenarios are detected as outliers. 6 Synthetic outliers are generated for analysis.

2, 3, 6 and 12 and nominal probabilities 95%, 90%, 85% and 80% are included in the figure. It is to be noted that to study the results for different days, the selected days are not the same for all dimensions shown in the figure. The vertical axes are logarithmic for a clearer illustration. The empirical coverage of original prediction convex hulls is 90% and it reduces to 87% for trimmed convex hulls. However, as trimming the convex hulls in dimension 2 reduces their sizes significantly, we recommend discarding outliers regardless.

Among the four techniques,  $P^1$  shows the overall best performance in terms of conservativeness and sharpness. As the dimension increases, the MPIs become wider and more conservative. For example, in dimension 12 for the first day, 80% MPI is more than 100 times larger than 80%  $P^1$ . In order to better visualize the size of MPIs in higher dimensions, Fig. 7 provides the MPIs for a randomly selected day from the evaluation data along with the marginal prediction intervals and multivariate trajectories used as their inputs. As shown in the figure, MPIs especially those with low coverage rates are wide and low in sharpness.

TABLE I  
EMPIRICAL COVERAGE RATE OF ROBUST BOXES

Dimension	D=2	D=6	D=12	D=24
Empirical Coverage Rate	0.25	0.10	0.02	0

### C. Comparison With Robust Uncertainty Sets

We follow [29] and consider  $c_{t+d}$  as the point forecast generated for time  $t$  and dimension  $d$  and  $w_{t+d} = 0.2 \times c_{t+d}$ . The robust polyhedrons with  $\gamma = 1$  and  $\gamma = 2$ , and the robust box of dimension 2 for a randomly selected day are illustrated in Fig. 8. In the same figure,  $P^1$  and  $P^\infty$  with nominal coverage rate 85% are also illustrated. When  $\gamma = 1$ , the robust polyhedron determines the boundaries of  $P^R$  and when  $\gamma = 2$ , the robust box is the limiting factor and the  $P^R$  will be the entire robust box.

As stated in subsection II-E,  $\gamma = D$  is the most conservative uncertainty region defined by (5). Table I gives the empirical coverage rates calculated for this most conservative case.  $P^R$  are very poorly calibrated and they cover only 25% of the real measurements and almost 0% of the measurements in dimensions 12 and 24. As the table implies,  $P^R$  show even worse calibration than UPIs. This is of course expected because as UPIs are the output of probabilistic forecasts, they take into account the univariate uncertainties. This is while  $P^R$  are simply the conventional confidence intervals around the point forecasts and they are even more constrained by the robust polyhedrons controlled by the uncertainty budget. It is to be emphasized that a lower calibration does not necessarily mean that the uncertainty regions are smaller. As can be seen in Fig. 8, a lower calibration can be a result of enclosing a wrong region where there is a low likelihood that the future random variable is realized in.

Fig. 9 gives more examples of the robust polyhedrons with  $\gamma = 2$ . In Fig. 9 a,  $P^1$  and  $P^\infty$  with 85% nominal coverage rate are compared with the robust polyhedron generated for the same time. Assuming  $P^R$  are limited by only robust polyhedrons, they cover about 36% of the measurements. In Figs. 9 b and 9 c, 35%  $P^1$  and  $P^\infty$  along with the robust polyhedrons are illustrated. As  $w_{t+d}$  is a function of point forecast for  $t$  and  $d$ , the robust polyhedrons tend to be larger for higher deterministic expected values.

As stated in subsection II-E, the framework in (5) does not output uncertainty sets with predefined probability levels. One

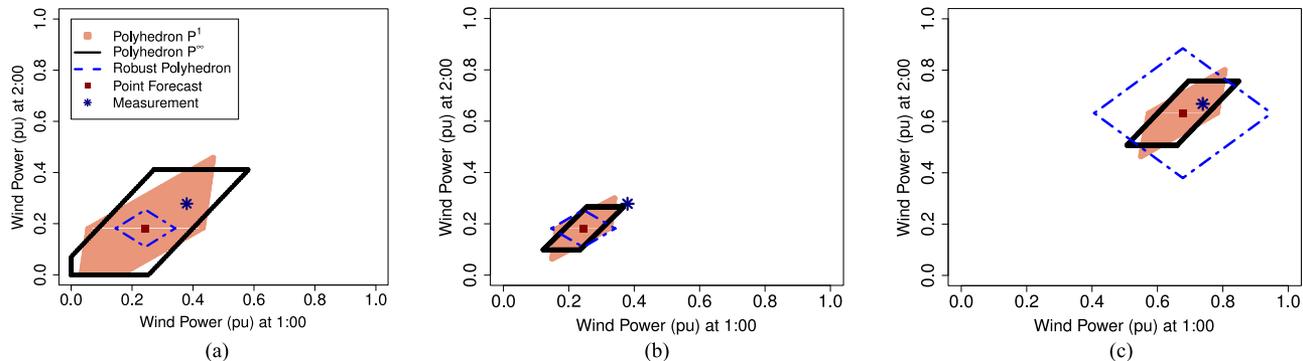


Fig. 9. Visual comparison of prediction polyhedra and robust uncertainty sets of dimension 2 for wind power data. (a)  $\alpha = 85\%$  (b)  $\alpha = 35\%$  and (c)  $\alpha = 35\%$  nominal coverage rate.

other fundamental drawback of  $P^R$  is that they are incapable of incorporating correlation between the random variables at different dimensions. As can be induced from Fig. 8,  $P^1$  and  $P^\infty$  both clearly show the correlated uncertainty between wind power at two successive hours as both are aligned with  $x = y$ -plane. When a random variable is temporally correlated, a large forecast error at time  $t$  usually is followed by a similar forecast error at the next time step. This is while, the geometry defined by  $P^R$  is unable to model such a dependency. This is the reason that in Table I, as the dimension increases the calibration degrades. When the dimension increases, the alignment between a higher number of random variables is assumed while this basically is not considered in how the  $P^R$  is designed.

Although formulation in (5) is been widely used in the literature recently, it naturally has very poor predictive performance when the random variables are correlated. This implies that using them as a representation of uncertainty in optimization does not output robust solutions to the true underlying uncertainty.

#### D. Comparison With Ellipsoidal Prediction Regions

It is worth comparing the prediction polyhedrons with the prediction ellipsoids proposed in [3]. We generated prediction ellipsoids with the same point and covariance matrix forecasts used to design  $P^1$  and  $P^\infty$ . The scale parameter of the prediction ellipsoids is designed based on the optimization framework proposed in [3] and it is updated every 30 steps (every 30 days). Fig. 10 illustrates the calibration errors of prediction ellipsoids,  $P^1$  and  $P^\infty$  of dimension 6. The prediction geometries with nominal coverage rate 85% are also compared with respect to their volumes in Fig. 11. As it can be deduced from the figures, prediction ellipsoids and polyhedrons show comparative predictive performance.

With regard to the complexity of implementation and the computational cost, the method proposed in this work to design the prediction polyhedrons is more efficient as it does not involve optimization. Since updating the scale parameter as proposed in this work requires very low computational cost, it is updated at every step (every day) while for the case of prediction ellipsoids, we limited updating the scale parameter to every 30 steps. It is worth noting that designing the parameters of  $P^1$  and  $P^\infty$  and the prediction ellipsoids are not limited to the methods proposed here and in [3]. We expect that the research community

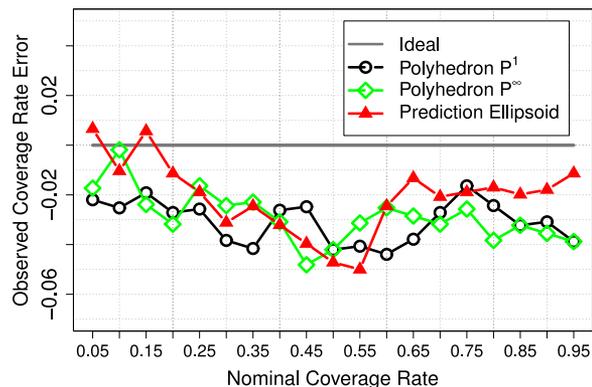


Fig. 10. The difference between empirical coverage and nominal coverage rates (Empirical-Nominal), dimension 6.

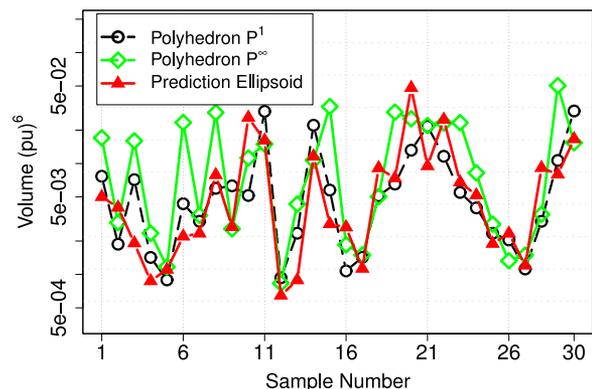


Fig. 11. Estimated volume of 85% prediction regions of dimension 6 for 30 randomly selected days.

will actively contribute to improve these classes of multivariate forecasts in the near future and develop competitive alternative approaches.

While both prediction geometries are able to forecast correlated uncertainty with a high accuracy, as it is explained in Section I, the choice between these two mostly depends on the formulation of the decision-making problem and the optimization solver available. The use of prediction polyhedrons is appealing in robust linear optimization. However, one should not assume that ellipsoidal prediction regions can be fully replaced by the prediction polyhedrons in all

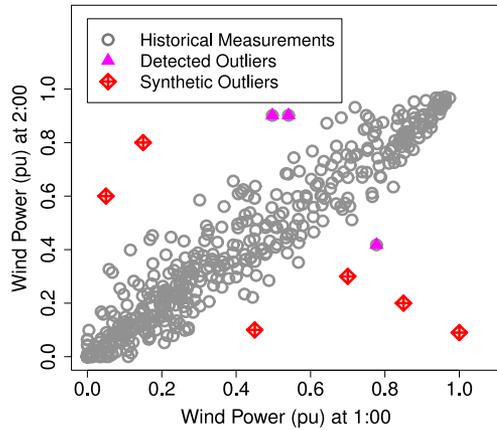


Fig. 12. Temporal scenarios predicted for a randomly selected day from the wind power data. Few scenarios are detected as outliers. 6 Synthetic outliers are generated for analysis.

robust optimization problems. For example, in Quadratically Constrained Quadratic Programs (QCQP), if the uncertainty set is a single prediction ellipsoid, then the robust counterpart is a semidefinite optimization problem. However, in case the uncertainty sets are in the form of polyhedrons, then the robust counterpart is NP-hard [17]. We intend to provide input to the various optimization formulations in a data-driven manner, so as to reconcile forecasting and decision-making under uncertainty.

#### E. Robustness to Outliers

In almost all real world datasets, there is a possibility of outlier occurrence. Outliers might come from error in measurement, collection or communication of data. Outliers can pose serious problems in statistical analysis and grossly distort and mislead them. The first measure to deal with outliers is to identify them, then either they can be discarded or replaced with more consistent data. Detection of multivariate outliers is discussed in subsection II-C. It is important to assess the robustness of various regression models to outliers.

Fig. 12 shows all the historical wind power measurements available in the training subset recorded at 1:00 am and 2:00 am. Following the approach discussed in subsection II-C, those observations with Mahalanobis distance higher than  $1.1\chi_2^2(\alpha = 0.001)$  are detected as outliers. We generate six more synthetic outliers as shown in Fig. 12 and substitute them for 6 randomly selected measurements. Then, we compare the performance of prediction polyhedra with and without the presence of those synthetic outliers. Based on empirical results, outliers on average change the empirical coverage of MPIs,  $P^1$ ,  $P^\infty$  and convex hulls by 1.5%, 0.7%, 0.9% and 3%, respectively. The occurrence of outliers also changes the volumes of 90% polyhedral prediction regions by 0.048, 0.012, 0.020 and 0.044 for MPIs,  $P^1$ ,  $P^\infty$  and convex hulls, respectively. The results indicate that  $P^1$  provides the highest robustness to outliers followed by  $P^\infty$ .

## VI. CONCLUSION

In order to facilitate the transition from deterministic forecasts to the point where end-users can confidently harness uncertainty

information, it is required to develop frameworks allowing to characterize uncertainty in forms that suit best the needs of various decision-making communities. Due to growing interests in polyhedral uncertainty sets, this work proposed frameworks to generate, calibrate and evaluate uncertainty information in the form of multivariate polyhedra for PV and wind power and within various temporal and spatial scales.

Two of the proposed techniques use point and correlation matrix forecasts as inputs and predict the uncertainty budget such that prediction polyhedra provide the desired probability levels and conservativeness. Two other techniques work based on finding convex boundaries of spatial/temporal scenarios. The proposed approaches together with multivariate prediction intervals as a benchmark are compared based on their calibration and conservativeness. The empirical results suggest that prediction convex hulls are not recommended for wind/PV predictions in dimensions higher than four. In addition, MPIs degrade with regard to sharpness as the dimension increases.  $P^1$  shows overall the best performance followed by  $P^\infty$ . Both  $P^1$  and  $P^\infty$  are promising formulations for skilled uncertainty characterization in convex forms and their performance does not degrade as the dimension increases, provided that their parameters are predicted appropriately.

We also compared  $P^1$  and  $P^\infty$  with ellipsoidal prediction regions and concluded that they are able to provide comparative predictive skill. It is discussed that the choice between ellipsoidal and polyhedral prediction regions depends on the optimization problem and the solvers available.  $P^1$  and  $P^\infty$  also are compared with the robust uncertainty sets. The poor performance of the robust uncertainty sets in characterizing the correlated uncertainty is elaborated both intuitively and numerically. It is shown that robust uncertainty sets have very poor calibration and their calibration dramatically worsens as the dimensions increases.  $P^1$  and  $P^\infty$  have the lowest computational complexity comparing to the other approaches as they require only point and covariance matrix forecasts and the scale parameter is estimated with very low computational cost.

## ACKNOWLEDGMENT

The authors thank Dr. R. Bessa for insightful discussions and for sharing code on multivariate prediction intervals and the editor and the anonymous reviewers for improving this manuscript by their constructive comments.

## REFERENCES

- [1] 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, Feb. 2013.
- [2] R. J. Bessa *et al.*, "Towards improved understanding of the applicability of uncertainty forecasts in the electric power industry," *Energies*, vol. 10, no. 9, p. 1402, 2017.
- [3] F. Golestaneh, P. Pinson, R. Azizipanah-Abarghooee, and H. B. Gooi, "Ellipsoidal prediction regions for multivariate uncertainty characterization," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4519–4530, Jul. 2018.
- [4] A. Ben-Tal and A. Nemirovski, "Robust solutions of linear programming problems contaminated with uncertain data," *Math. Program.*, vol. 88, no. 3, pp. 411–424, 2000.

- [5] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets: Operational Problems*, vol. 205. Berlin, Germany: Springer Science & Business Media, 2013.
- [6] P. Haessig, B. Multon, H. B. Ahmed, S. Lascaud, and P. Bondon, "Energy storage sizing for wind power: Impact of the autocorrelation of day-ahead forecast errors," *Wind Energy*, vol. 18, no. 1, pp. 43–57, 2015.
- [7] C. Lowery and M. O'Malley, "Reserves in stochastic unit commitment: An Irish system case study," *IEEE Trans. Sustain. Energy*, vol. 6, no. 3, pp. 1029–1038, Jul. 2015.
- [8] Q. P. Zheng, J. Wang, and A. L. Liu, "Stochastic optimization for unit commitment? A review," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1913–1924, Jul. 2015.
- [9] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Oper. Res.*, vol. 62, no. 6, pp. 1358–1376, 2014.
- [10] M. Doostizadeh, F. Aminifar, H. Ghasemi, and H. Lesani, "Energy and reserve scheduling under wind power uncertainty: An adjustable interval approach," *IEEE Trans. Smart Grid*, vol. 7, no. 6, pp. 2943–2952, Nov. 2016.
- [11] W. Wei, F. Liu, and S. Mei, "Distributionally robust co-optimization of energy and reserve dispatch," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 289–300, Jan. 2016.
- [12] B. Hu and L. Wu, "Robust SCUC with multi-band nodal load uncertainty set," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2491–2492, May 2016.
- [13] Y. Guan and J. Wang, "Uncertainty sets for robust unit commitment," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1439–1440, May 2014.
- [14] L. Wu, M. Shahidehpour, and Z. Li, "Comparison of scenario-based and interval optimization approaches to stochastic SCUC," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 913–921, May 2012.
- [15] A. Venzke, L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis, "Convex relaxations of chance constrained ac optimal power flow," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2829–2841, May 2018.
- [16] A. Velloso, A. Street, D. Pozo, J. M. Arroyo, and N. G. Cobos, "Scenario-based uncertainty set for two-stage robust energy and reserve scheduling: A data-driven approach," unpublished paper, 2018. [Online]. Available: <https://arxiv.org/abs/1803.06676v1>
- [17] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM Rev.*, vol. 53, no. 3, pp. 464–501, 2011.
- [18] P. Li, X. Guan, J. Wu, and X. Zhou, "Modeling dynamic spatial correlations of geographically distributed wind farms and constructing ellipsoidal uncertainty sets for optimization-based generation scheduling," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1594–1605, Oct. 2015.
- [19] C. Shao, X. Wang, M. Shahidehpour, X. Wang, and B. Wang, "Security-constrained unit commitment with flexible uncertainty set for variable wind power," *IEEE Trans. Sustain. Energy*, vol. 8, no. 3, pp. 1237–1246, Jul. 2017.
- [20] T. Soares, R. J. Bessa, P. Pinson, and H. Morais, "Active distribution grid management based on robust AC optimal power flow," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6229–6241, Nov. 2018.
- [21] R. J. Bessa, "From marginal to simultaneous prediction intervals of wind power," in *Proc. 18th Int. Conf. Intell. Syst. Appl. Power Syst.*, 2015, pp. 1–6.
- [22] D. A. Pachamanova, "A robust optimization approach to finance," Ph.D. dissertation, Massachusetts Inst. Technol., Operations Research Center, 2002.
- [23] F. Golestaneh, H. B. Gooi, and P. Pinson, "Generation and evaluation of space-time trajectories of photovoltaic power," *Appl. Energy*, vol. 176, pp. 80–91, 2016.
- [24] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, "The quickhull algorithm for convex hulls," *ACM Trans. Math. Softw.*, vol. 22, no. 4, pp. 469–483, 1996.
- [25] E. Mücke, "Quickhull: Computing convex hulls quickly," *Comput. Sci. Eng.*, vol. 11, no. 5, pp. 54–57, 2009.
- [26] M. Werner, "Identification of multivariate outliers in large data sets," Ph.D. dissertation, Dept. Math., University of Colorado at Denver, 2003.
- [27] D. Kolsrud, "Time-simultaneous prediction band for a time series," *J. Forecasting*, vol. 26, no. 3, pp. 171–188, 2007.
- [28] J. S.-H. Li and W.-S. Chan, "Simultaneous prediction intervals: An application to forecasting US and Canadian mortality," in *Proc. Living 100 Symp.*, 2011, pp. 1–21.
- [29] A. Attarha, N. Amjadi, and A. J. Conejo, "Adaptive robust ac optimal power flow considering load and wind power uncertainties," *Int. J. Elect. Power Energy Syst.*, vol. 96, pp. 132–142, 2018.
- [30] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets: Operational Problems*. Berlin, Germany: Springer, 2014.
- [31] T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," *J. Roy. Statist. Soc.: Series B*, vol. 69, no. 2, pp. 243–268, 2007.
- [32] B. Büeler, A. Enge, and K. Fukuda, "Exact volume computation for polytopes: A practical study," in *Polytopes: Combinatorics and Computation*. Berlin, Germany: Springer, 2000, pp. 131–154.
- [33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [34] R. Engle, "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *J. Bus. Econ. Statist.*, vol. 20, no. 3, pp. 339–350, 2002.
- [35] R. F. Engle and K. Sheppard, "Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH," Nat. Bureau Econ. Res., Cambridge, MA, USA, Tech. Rep., 2001.
- [36] V. Zakamulin, "A test of covariance-matrix forecasting methods," *J. Portfolio Manage.*, vol. 41, no. 3, pp. 97–108, 2015.
- [37] Global energy forecasting competition 2014 probabilistic solar power forecasting. 2014. [Online]. Available: <https://crowdanalytix.com>
- [38] F. Golestaneh, P. Pinson, and H. B. Gooi, "Very short-term nonparametric probabilistic forecasting of renewable energy generation—With application to solar energy," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3850–3863, Sep. 2016.

Authors' photographs and biographies not available at the time of publication.