Multivariate Uncertainty Characterization for Resilience Planning in Electric Power Systems

Mostafa Nazemi, Student Member, IEEE, Payman Dehghanian, Member, IEEE, Mohannad Alhazmi, Student Member, IEEE, and Fei Wang, Senior Member, IEEE

Abstract—Following substantial advancements in stochastic classes of decision-making optimization problems, scenario-based stochastic optimization, robust (distributionally robust) optimization, and chance-constrained optimization have recently gained an increasing attention. Despite the remarkable developments in probabilistic forecast of uncertainties (e.g., in renewable energies), most approaches are still being employed in a univariate framework which fails to unlock a full understanding on the underlying interdependence among uncertain variables of interest. In order to yield cost-optimal solutions with predefined probabilistic guarantees, conditional and dynamic interdependence in uncertainty forecasts should be accommodated in power systems decision-making. This becomes even more important during the emergencies where high-impact low-probability (HILP) disasters result in remarkable fluctuations in the uncertain variables. In order to model the interdependence correlation structure between different sources of uncertainty in power systems during both normal and emergency operating conditions, this paper aims to bridge the gap between the probabilistic forecasting methods and advanced optimization paradigms; in particular, perdition regions are generated in the form of ellipsoids with probabilistic guarantees. We employ a modified Khachiyan's algorithm to compute the minimum volume enclosing ellipsoids (MVEE). Application results based on two datasets on wind and photovoltaic power are used to verify the efficiency of the proposed framework.

Index Terms—Probabilistic forecasting; uncertainty sets; ellipsoids; resilience; stochastic optimization.

I. INTRODUCTION

The proliferation of renewable energy resources have resulted in a drastic increase in the level of complexity and uncertainty in electric power systems. This calls for advanced developments of (i) highly scalable optimization techniques capable of accommodating remarkably high degree of uncertainty, and (ii) fundamental forecasting foundations to provide a suitable set of inputs to a number of decision-making problems ranging from long-term planning to short-term operation in power systems under uncertainty and risk [1], [2].

This work was supported in part by the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (grant No. LAPS19016).

M. Nazemi, P. Dehghanian, and M. Alhazmi are with the Department of Electrical and Computer Engineering, George Washington University, Washington, DC 20052, USA (e-mails: mostafa_nazemi@gwu.edu; payman@gwu.edu; alhazmi@gwu.edu).

Fei Wang is with the State Key Laboratory of Alternate Electrical Power System With Renewable Energy Sources, North China Electric Power University, Beijing 102206, China, and also with the Department of Electrical Engineering and Hebei Key Laboratory of Distributed Energy Storage and Microgrid, North China Electric Power University, Baoding 071003, China (e-mail: feiwang@ncepu.edu.cn).

The most commonly accepted forecasting methodology in the literature and in practice appears in the form of deterministic or point forecast which comprises of single-valued prediction realization for a variable of interest, location and time individually and independently as the input to various decision making problems [3], [4]. Although easier to implement and interpret, such techniques are always subject to errors and fluctuations as the solution to a deterministic optimization formulation is highly sensitive to small changes in uncertain quantities [5]; if ignored, such uncertainties can result in suboptimal or infeasible solutions. Probabilistic forecasting, e.g., stochastic optimization [6], chance-constrained [7], robust [8], and interval [9] optimization, can accommodate mechanisms to handle various uncertainties in the system. For instance, the well-known stochastic programming optimization techniques utilize so many scenarios to find the optimal solution in uncertain environments [10]. Nevertheless, stochastic programming demands a heavy computational burden and the hard-to-characterize probability distributions of random variables [11].

The existing literature is found at the earliest stage of developments on characterizing multivariate uncertainty sets to be utilized in such classes of stochastic optimization problems. Most probabilistic forecasts are employed in a univariate framework which can provide the uncertainty information for every variable, lead time, and location individually; however, such forecasts are not able to provide optimal inputs to various classes of decision-making problems particularly when spatial, temporal and/or inter-variable dependencies have to be considered. This is, for instance, the case when facing extreme weather emergencies (e.g., hurricanes), when the renewable energy resources such as solar photovoltaic (PV) or wind power (WP) become significantly sensitive to meteorological impacts of the weather. In other words, during the hurricane, while the PV output is low due to the cloudy sky, WP is much higher with higher wind speed. Therefore, such interdependencies are expected to be considered when generating scenarios for structural resilience planning as well as response and recovery solutions to ensure the operational resilience [12]–[21].

The most common approach in the literature for modeling temporal/spatial dependencies between different sources of uncertainties is to generate a huge set of scenarios as random draws of multivariate characterization [22]. However, such classes of decision-making problems based on probabilistic forecast necessitate multivariate probabilistic forecast regions rather than scenarios. Robust optimization as a popular class of probabilistic optimization tends to produce a conservative
solution [11]. While the conservativeness of the solution is highly dependent on the size of the uncertainty sets, controlling the size of the uncertainty set is not a straightforward task. Similar to the case of univariate probabilistic forecasts, multivariate prediction region should satisfy the required level of performance in the system. A multivariate prediction region mostly takes the form of multivariate ellipsoids, polyhedral or boxes which technically defines a region where the realization of multivariate random variable is expected to lie in, with a certain probability or a predefined tolerance level.

Only a few established frameworks are available in the literature to generate multivariate prediction regions [1], [2], [23]. Uncertainty sets are constructed based on the Gaussian assumption in [24], [25] for wind power and for load demands [26]. The size and conservativeness of the wind uncertainty sets in the form of polyhedra presented in [27], [28] and in the form of ellipsoids in [29] are controlled by a parameter called uncertainty budget.

For many of probabilistic decision-making problems, the uncertainty set is characterized in the form of ellipsoids [30], [31], as ellipsoids prediction regions are more flexible and realistic compared to finite or hype-boxes ones [32]. To the best of our knowledge, except a few works in the literature [2], there is no established work to model the correlation of different sources of uncertainty following natural hazards (e.g., hurricanes, earthquakes, wildfires, etc.). The primary goal in this paper is to propose a new class of multivariate prediction regions to characterize the uncertainty of those random variables which are spatially/temporally correlated and dependent to each other during and following the HILP events. A new optimization framework is proposed based on Khachiyan’s algorithm [33]–[35] to find the minimum volume enclosing ellipsoids (MVEE). The empirical data is employed to evaluate the performance of the predicted MVEE regions which can be used as an input to interval, robust, and chance-constrained optimization problems.

The rest of the paper is organized as follows: in Section II, a big picture of the proposed framework for weather-driven uncertainty correlation characterization following natural disasters is explained. Section III presents the formulation of the Khachiyan’s algorithm, while the methodology to generate multivariate ellipsoidal prediction regions is discussed in Section IV. Section V includes the empirical results and finally concluding remarks are given in Section VI.

II. WEATHER-DRIVEN UNCERTAINTY CORRELATION

Natural disasters (e.g., windstorm, hurricane, flood, earthquake, wildfire, tornado, etc.) can have a considerable impact on power system operation. Day-to-day weather parameter forecasts are subject to a huge range of fluctuations during emergencies which should be mathematically accounted for in different planning and operation decision-making problems. Figure 1 presents the big picture of the proposed weather-driven multivariate uncertainty characterization framework which includes three different layers: (i) in the first layer, various daily weather parameters are measured in weather stations across the network. In case of an extreme weather emergency, the fluctuation in weather parameter forecasts can be detected in advance; (ii) the correlations between different sources of uncertainty in power system are mathematically modeled in the second layer. For instance, as shown in Fig. 2 following a severe hurricane, the portion of the generated electricity from wind power becomes larger due to higher-wind conditions and wind speed, while the PV power output decreases due to the cloudy weather; (iii) finally, the correlated uncertainties are considered as an input to the last layer in which various decision-making optimization problems ranging from long-term planning to short-time operation are solved using different types of stochastic optimization algorithms (e.g., robust, distributionally robust, and chance-constrained). One can conclude that the correlation between different sources of uncertainty (e.g., wind power and solar energy) cannot be simply neglected in modern decision-making problems. In this paper, a novel framework is proposed to model the correlation of different uncertain factors in power system during both normal and emergency operating conditions: we utilize the concept of ellipsoidal prediction regions which can be used as an input to different stochastic optimization problems. Among a huge set of prediction regions which can be used to find the optimal solutions of stochastic optimization problems, the proposed framework provides a prediction region with a predefined reliability level.

III. MATHEMATICAL FORMULATION

Inspired by [33], we here provide necessary formulations to generate and assess the volume of a full-dimension ellipsoid. Let $X := \{x^1, ..., x^n\} \subseteq \mathbb{R}$ be a matrix whose affine hull is $\mathbb{R}^d$. Note that the $\mathbf{r}$th column is given by $A^r$, $i = 1, ..., m$. The minimum-volume ellipsoid satisfies:

$$\frac{1}{d!} \text{MVEE}(X) \subseteq \text{conv}(X) \subseteq \text{MVEE}(S),$$

where $\text{conv}(X)$ represents the convex hull of $X$ and the ellipsoid on the left-hand side is obtained by scaling $\text{MVEE}(X)$ around its center by a factor of $1/d$. The factor on the left-hand side is changed to $1/\sqrt{2}$ if $X$ is centrally symmetric (i.e., if $X = -X$). Hence, $\text{MVEE}(X)$ provides a rounding of the full-dimensional $\text{conv}(X)$. Given $\epsilon > 0$, an ellipsoid $E \subseteq \mathbb{R}^d$ can be interpreted as a $(1+\epsilon)/\epsilon$-rounding of $\text{conv}(X)$ if

$$\frac{1}{(1+\epsilon)^d} E \subseteq \text{conv}(X) \subseteq E,$$

where the left-hand side of (2) is replaced by $1/\sqrt{(1+\epsilon)^d}$ while the ellipsoid $E$ is centrally symmetric. Likewise, for $\gamma > 0$, an ellipsoid $E \subseteq \mathbb{R}^d$ is a $(1+\gamma)$-approximation of $\text{MVEE}(X)$ if

$$\text{conv}(X) \subseteq E, \quad \text{Vol}(E) \leq (1 + \gamma)\text{Vol}(\text{MVEE}(X)),$$

where $\text{Vol}(E)$ denotes the volume of the ellipsoid $E$. A full-dimensional ellipsoid $E_{A, c} \subseteq \mathbb{R}^d$ is specified by a $d \times d$ symmetric positive definite matrix $A$ and a center $c \in \mathbb{R}^d$.

$$E_{A, c} = \{x \in \mathbb{R}^d : (x - c)^T A(x - c) \leq 1\}$$

The volume of the ellipsoid $E_{A, c}$ is denoted by $\text{Vol}(E_{A, c}) = \mu |A^{-1}/2|$, where $\mu$ is the volume of the unit ball in $\mathbb{R}^d$ [33],
Fig. 1. The overall architecture of the proposed framework for multivariate weather-driven uncertainty characterization.

If \( X \) is not centrally symmetric, a “lifting” of \( X \) to \( \mathbb{R}^n \) is defined, where \( n := d + 1 \), by

\[
X' := \pm a^1, ..., \pm a^m, \quad \text{where } a^i := \begin{bmatrix} x^i \\ 1 \end{bmatrix} \ i = 1, ..., m,
\]

which is centrally symmetric. One can then conclude that MVEE(\( X \)) and MVEE(\( X' \)) are closely related [35]. Because \( X' \) is centrally symmetric, the center of MVEE(\( X' \)) is located at the origin; therefore, the problem of computing MVEE(\( X' \)) can be formulated as the following convex optimization problem [33]:

\[
\min_Y - \log |Y| \\
\text{s.t. } (a^i)^T Y a^i \leq 1, \quad i = 1, ..., m, \quad Y \in \mathbb{R}^{n \times n} \text{ is symmetric and positive definite}
\]

where \( Y \in \mathbb{R}^{n \times n} \) is the decision variable. The Lagrangian dual of (6) is equivalent to

\[
\begin{align*}
\max_r \Phi(r) &:= \log |\Gamma(r)||Y| \\
\text{s.t. } c^T r &= 1, \\
r &\geq 0,
\end{align*}
\]

where \( r \in \mathbb{R}^m \) is the decision variable and \( \Gamma : \mathbb{R}^m \rightarrow \mathbb{R}^{n \times n} \) is a linear operator function given by the following equation

\[
\Gamma(r) := \sum_{i=1}^m r_i a^i (a^i)^T.
\]

The following sufficient optimality conditions for \( r^* \) should be satisfied to solve the dual program (7).

\[
\begin{align*}
\nu_i(r^*) + s_i^* &= \lambda^*, \quad i = 1, ..., m, \quad (9a) \\
e^T r^* &= 1, \quad (9b) \\
r_i^* s_i^* &= 0, \quad i = 1, ..., m, \quad (9c) \\
r_i^* s_i^* &\geq 0, \quad i = 1, ..., m, \quad (9d)
\end{align*}
\]

where

\[
\nu_i(r) := (a^i)^T \Gamma(r)^{-1} a^i, \quad i = 1, ..., m
\]

For any feasible solution \( r \in \mathbb{R}^m \) of (7) with \( \Phi(r) > -\infty \), we have [36]:

\[
\sum_{i=1}^m r_i \nu_i(r) = n \quad (11)
\]

Therefore, multiplying both sides of (9a) by \( r_i^* \) and summing up for \( i = 1, ..., m \), we obtain \( \lambda^* = n \) by (9b) and (9c). So, the optimality condition of (7) can be equivalently written as

\[
\begin{align*}
\nu_i(r^*) &\leq n, \quad i = 1, ..., m, \quad (12a) \\
e^T r^* &= 1, \quad (12b) \\
r^* &\geq 0. \quad (12c)
\end{align*}
\]

By combining (11) and (12), one can conclude

\[
r_i^* > 0 \quad \text{implies } \nu_i(r^*) = n, \quad i = 1, ..., m
\]

Fig. 2. The weather-dependent correlation of power grid uncertainties during extreme weather events.
which can be interpreted as the complementary slackness conditions of (9c). Khachiyan’s algorithm is driven by computing a feasible solution $\tilde{r}$ of dual program that satisfies the $\epsilon$-related optimality conditions defined by

$$ v_i(\tilde{r}) \geq (1 + \epsilon)n, \quad i = 1, ..., m $$  \hspace{1cm} (14)

For such a solution $\tilde{r}$, let $j \in \{1, ..., m\}$ be such that $\tilde{r}_j > 0$. Using (11), we can have

$$ v_j(\tilde{r}) = \frac{1}{\tilde{r}_j} \left( n - \sum_{i=1, i \neq j}^m \tilde{r}_j v_j(\tilde{r}) \right) \geq \frac{n}{\tilde{r}_j} \left( 1 + \epsilon \right) \left( 1 + \frac{1}{\tilde{r}_j} \right) $$  \hspace{1cm} (15)

where (14) is used to check the feasibility of $\tilde{r}$. Hence, such a solution $\tilde{r}$ satisfies an approximation form of the complementary slackness conditions of (13). Therefore, Khachiyan’s algorithm [36] starts with a feasible solution $\tilde{r} > 0$ of (7) and improves upon the objective function value by increasing only one component of $\tilde{r}$ at each iteration and then re-scaling to the region feasibility. So, in order to relate the optimal solution $r^*\leq n \epsilon (1+\epsilon)$ at each iteration and then re-scaling to the region feasibility. Hence, such a solution $\tilde{r}$ satisfies an approximation form of the complementary slackness conditions of (13). Therefore, Khachiyan’s algorithm [36] starts with a feasible solution $\tilde{r} > 0$ of (7) and improves upon the objective function value by increasing only one component of $\tilde{r}$ at each iteration and then re-scaling to the region feasibility. In this section, we review Khachiyan’s algorithm [36]

$$ A^* := \frac{1}{d} (B^r B^T - Br^r) (Br^r)^{-1} $$  \hspace{1cm} (16)

where $A^* := \frac{1}{d} (B^r B^T - Br^r) (Br^r)^{-1}$.

Besides, we can have

$$ \log \text{vol}(\text{MVEE})(X) = \frac{d}{2} \log d + \frac{1}{2} \log |\Gamma(r^*)| $$  \hspace{1cm} (17)

IV. \text{KHACHIYAN’S ALGORITHM IMPLEMENTATION}

In this section, we review Khachiyan’s algorithm [36] which seeks to find a minimum volume ellipsoid containing $A := \{\pm a^1, ..., \pm a^m\};$ however, it constructs a sequence of ellipsoids as it is a dual program

$$ E_k := \{y \in \mathbb{R}^d : y^T \Gamma(r^*)^{-1} y \leq 1\} $$  \hspace{1cm} (18)

which should satisfy $E_k \subset A$, and it stops when $A \leq \sqrt{(1+\epsilon)n} E_k$. Hence, the polar ellipsoid takes the following form

$$ E_k^0 := \{z \in \mathbb{R}^n : z^T \Gamma(r^*) z \leq 1\} $$  \hspace{1cm} (19)

and the algorithm stops when $\{1/\sqrt{(1+\epsilon)n} E_k^0\}$ is contained in $A^\circ$. However, in a case that $\{1/\sqrt{(1+\epsilon)n} E_k^0\}$ is not contained in $A^\circ$ at a particular iteration, one pair of hyperplanes $(a^i)^T z = \pm 1$ of $A^\circ$ intersects $E_k^0$. So the condition

$$ (a^i)^T \Gamma(r^*)^{-1} a^i = \eta > (1 + \epsilon)n \text{ in Khachiyan’s algorithm} \text{ correspond to the following} $$

$$ A^\circ \subseteq \{z \in E_k^0 : -\psi \sqrt{(a^i)^T \Gamma(r^*)^{-1} a^i} \leq (a^i)^T z \leq \psi \sqrt{(a^i)^T \Gamma(r^*)^{-1} a^i}\} $$  \hspace{1cm} (20)

where $\psi = 1/\sqrt{p} < 1/\sqrt{(1+\epsilon)n}$. In general, Khachiyan’s algorithm can be summarized as follows:

\begin{algorithm}
\caption{Modified Khachiyan’s algorithm to compute a feasible solution of (7) satisfying (14)}
\begin{algorithmic}
\State \textbf{Require:} $X = \{x^1, ..., x^n\} \subset \mathbb{R}^d$, $\epsilon > 0$
\State $k \leftarrow 0$, $n \leftarrow d + 1$, $r^0 \leftarrow (1/m)e$
\State $\epsilon^i \leftarrow ((a^i)^T, 1)^2, i = 1, ..., m$
\State $1$: While $r^k$ does not satisfy (14), do $\text{end loop}$
\State $2$: \textbf{loop}$\text{begin}$
\State $3$: $j \leftarrow \arg \max_{i=1}^{m} (a^i)^T \Gamma(r^k)^{-1} a^i$
\State $4$: $\rho \leftarrow (a^i)^T \Gamma(r^k)^{-1} a^i$
\State $5$: $\beta \leftarrow \min(\epsilon^i, 1)$
\State $6$: $r^{k+1} \leftarrow (1 - \beta) r^k + \beta e^j, k \leftarrow k + 1$
\State $\text{end loop}$
\State $7$: \textbf{Output} $r^k$.
\end{algorithmic}
\end{algorithm}

Note that at each iteration, in order to compute the $(a^i)^T \Gamma(r^k)^{-1} a^i$, the inverse of $\Gamma(r^k)$ should be available. At each iteration, $\Gamma(r^k)$ is updated by adding a rank-one symmetric matrix to it and then scale it by a positive number. In some iterations, the vector $a^i$ and hence possibly the matrix $\Gamma(r^k)$ may be sparse. So, the Cholesky factorization $LDL^T$ of $\Gamma(r^k)$ is used where $L$ is a lower triangular matrix with unit diagonal and $D$ is diagonal with positive diagonal entries. In order to calculate $\rho$ and $j$, we can have $\rho_i = (a^i)^T \Gamma(r^k)^{-1} a^i$ for each $i$.

V. \text{NUMERICAL RESULTS AND DISCUSSIONS}

A. Data

Three different datasets including wind, PV power [37], and hurricane intensity [38] are used in this paper to model the correlation between different sources of uncertainty in case of hurricanes. For all three datasets, the resolution of data is considered as 1000 scenarios for one hour and the forecasts horizon are 1 to 24 hours. The PV and wind forecasts for three different randomly selected days are depicted in Fig. 3. The hurricane intensity is measured by the Saffir-Simpson hurricane wind scale which is from 1 to 5 based on a hurricane’s sustained wind speed [38]. All simulations have been performed in Matlab platform in a Mac-book Pro with a 2.4 GHz Quad-Core Intel Core i5 and 16 GB of memory.

B. Ellipsoidal Prediction Regions Implementation

1) Normal Operating Conditions: The bi-variate prediction ellipsoids are generated using the proposed algorithm to evaluate the correlation of wind power in different time scales. Figure 4 illustrates 5 ellipsoidal prediction regions with probabilities ranging from 0.80 to 0.95 by 0.05 increments and a probability of 0.99 for three randomly-selected time
periods. One can notice that as the probability of enclosing data increases, the size of the obtained ellipsoids becomes larger. Similarly in Fig. 5, four different prediction ellipsoids are generated to model the correlations of PV power output as a random variable in different time scales. Figure 6 represents the correlation between two sources of uncertainty (e.g., wind and PV power) during system normal operating conditions. Comparing different ellipsoidal prediction regions, it can be seen that the shape, rotation, and calibration of the ellipsoids vary depending on the stochastic processes of interests as well as the underlying uncertainty level in different time periods and different days. Moreover, the covariance matrix of the random variables highly impacts both the size and orientation of the prediction regions. Therefore, it is of utmost importance to have an accurate prediction regions.

2) Emergency Operating Conditions: During the weather emergencies, the correlation of different sources of uncertainty is even further highlighted and should be accounted for in different decision-making problems. Take a HILP hurricane as an example, during which, the windy weather would result in more wind power outputs due to higher wind speeds; however, the PV power generation is much less than that during the normal operating conditions because of the cloudy weather. According to [38], we here assume five different categories to represent the HILP hurricane intensity, i.e., Category 1 with 74-95 mph sustained winds speed as very dangerous...
Fig. 6. Optimal ellipsoid prediction regions for Wind power and PV Power with probabilities ranging from 0.85 to 0.95 by 0.05 increments (from the smallest to the largest) for three randomly-selected time intervals.

Fig. 7. Three dimensional optimal ellipsoid prediction regions with tolerance level $\epsilon = 0.1$ for three randomly-selected time intervals with different hurricane intensity level forecasts.

wind. Category 2 with 96-110 mph wind speed as extremely dangerous wind, Category 3 with 111-129 mph wind speed as devastating damage wind, Category 4 with 130-156 mph wind speed as catastrophic damage winds and finally, Category 5 with 157 mph or higher wind speed as very catastrophic damage wind. Note that the location of wind farms are not necessarily subject to these hurricane categories. To further simplify the evaluations and considering the safety operation requirements of the wind farms taken into account during the design and deployment process, We here assume that the location of wind farms are not very close to the hurricane path. Figure 7 reflects the three dimensional correlation of wind, PV power, and hurricane intensity at three randomly-selected time intervals during the HILP hurricane day. Different statistics (e.g., min, max, mean, variance, and standard deviation) of the wind and PV power in the three selected time intervals
are tabulated in Table I. As it can be seen, a higher hurricane intensity results in more wind power output and less PV power generation. one can conclude that the proposed framework not only is able to capture the correlation of different uncertainties in power grids, but it can also effectively correlate the underlying interdependence structure of the hazard intensity with the existing sources of exogenous uncertainties in power grids. Effectively addressing the correlation of such uncertain factors during emergencies will result in generation of plausible scenarios that will feed modern optimization formulations seeking solutions for enhanced structural and operational resilience during HILP extremes.

VI. CONCLUSION

Transitioning from deterministic to stochastic forecasts in which power system operators can confidently harness the full potentials of uncertain renewables calls for developing frameworks that allow for characterization of multivariate uncertainties in forms that suit the best needs of various decision-making problems. This paper has proposed a framework to generate, calibrate, and evaluate the uncertainty information in the form of multivariate ellipsoidal prediction regions which can be used as an input to different classes of stochastic optimization. In the face of extreme hazards, the multivariate uncertainty sets can provide the required information on stochastic variables that are temporally/spatially correlated with severity and intensity of the extreme event. In order to verify the applicability of the proposed approach, two datasets including wind and PV power during both normal and emergency operating conditions are employed. The simulation results revealed that the proposed ellipsoidal uncertainty characterization is able to track and predict the existing uncertainty level in time and capture the temporal/spatial multivariate uncertainty correlation with the desired probability and sharpness level. The proposed framework can be applied to a variety of planning and operation problems which involve correlated random variables in order to elevate the resilience level of the power grid when facing natural or man-cyber threats.

REFERENCES


