An Improved Integrated Cumulant Method by Probability Distribution Pre-Identification in Power System With Wind Generation

RUILIN CAO\textsuperscript{1}, JIE XING\textsuperscript{1}, BINGYAN SUI\textsuperscript{2}, AND HONGYAN MA\textsuperscript{1}

\textsuperscript{1}College of Information Science and Technology, Donghua University, Shanghai 201620, China
\textsuperscript{2}PowerChina Shanghai Electric Power Engineering Corporation Ltd., Shanghai 200025, China

Corresponding author: Jie Xing (xingj@dhu.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2019YFE012784.

\textbf{ABSTRACT} The increase of wind generation (WG) has challenged the conventional way of probabilistic load flow (PLF) calculation. A reliable and efficient PLF method is required to face the stochastic nature of various power systems with WG. Firstly, the paper analyzes several typical cumulant methods (CMs) for PLF, such as Gram-Charlier expansion of type A (GCA), Gram-Charlier expansion of type C (GCC), and maximum entropy (ME). Then, an improved integrated CM by probability distribution pre-identification is proposed for power systems with WG based on doubly fed induction generations (DFIGs). The skewness and kurtosis are used as probability distribution pre-identification indices in the CM framework. Meanwhile, the influence of the DFIG control strategy on reactive power is considered in the load flow model and the moment calculation. Finally, the accuracy and efficiency of the proposed method are validated with the IEEE test system. In various scenarios, suitable CM is selected and applied to the PLF based on pre-identifying distribution characteristics. Results reveal that probabilistic density functions (PDFs) of bus voltages and line flows obtained by the proposed method have both accuracy and efficiency.

\textbf{INDEX TERMS} Probabilistic load flow (PLF), integrated cumulant method (integrated CM), doubly fed induction generation (DFIG), Gram-Charlier expansion of type A (GCA), Gram-Charlier expansion of type C (GCC), maximum entropy (ME).

\section{I. INTRODUCTION}

In recent years, the development of renewable energy is growing rapidly worldwide, which plays an important role in alleviating fossil energy depletion and environmental pollution [1], [2]. However, intermittence, randomness, and weak controllability are the main features of renewable energy, such as wind and photovoltaic. The high penetration of renewable energy leads to increasing the uncertainties of power systems [3], [4]. The uncertainties affect the operation and stability of power systems [5]–[8]. In order to consider the uncertainties arising from the stochastic nature of loads and generators, the method of probabilistic load flow (PLF) is proposed [9]. Currently, wind generation (WG) is one of the premier renewable power generations [10]–[12]. Considering the uncertainties arising from WG in steady-state operation, the PLF method is further studied to analyze the power system integrated with WG [13], [14].

The methods for PLF include simulation method, analytical method, and approximate method [15], [16]. Among them, the analytical method obtains probabilistic distributions of variables with high computational efficiency and the method has attracted wide attention from scholars [6]. The most widely used method in the analytical method is the cumulant method (CM) that replaces the complex convolution operations with simple cumulants arithmetic operations [1]. In the research of CMs, Gram-Charlier expansion (GC) [17]–[20], Cornish-Fisher expansion (CF) [21]–[23], and maximum entropy (ME) [24]–[26] are used more.

In traditional power systems, uncertainties mainly come from common units and loads. The generation satisfies 0-1 distribution and the load characteristic satisfies normal distribution in steady-state operation [17], which makes the state variables of power systems such as bus voltages and
Gram-Charlier expansion of type A (GCA) is used in the PLF method to fit the probability density function (PDF) of the state variables in power systems with enough accuracy and efficiency [17], [18]. Moreover, the form of GCA is simple and easy to implement.

However, the distribution characteristic of WG is different from that of common generators because wind speed in nature generally approximates Weibull distribution [6]. The WG’s output characteristic deviates from a normal distribution, which causes state variables not to meet quasi-normal distributions. To calculate non-quasi-normal distribution based on cumulants, CF, Gram-Charlier expansion of type C (GCC), and ME have attracted researchers’ attention for their good performances in PLF calculation. CF is the inverse function expansion of the variables’ cumulative distribution function (CDF) based on GC. In [23], with few iterations and low time consumption, CF is more effective for evaluating the impact of photovoltaic generators on voltage profiles in an IEEE 33-bus radial system. In [19], GCC that adopts an exponential form is introduced to the power grid integrated with large-scale wind farms. In the IEEE 30-bus test system, the computing time of GCC is significantly faster than that of the Monte Carlo simulation (MCS). Taking the MCS results as the comparative baseline, the average root mean square (ARMS) of the results of GCC is mostly less than 0.01%. ME is used in PLF to make a reasonable inference on the unknown probability distribution of random power flow. In [25], one bus is integrated with 30 MW wind turbines based on the IEEE 118-bus system. As the change of scenarios, ME always gets accurate PDFs. For time consumptions, PLF with 50000 times MCS is 242.6 s and with 8th ME is 1.04 s in a scene. In [26], results show that ME is superior to GCA in accuracy and better than MC in computational efficiency. The effectiveness of the CMs is closely related to whether the variable satisfies the normal distribution [27]. Most current studies focus on improving a certain PDF fitting method, with the purpose to obtain results from the method closer to MCS results. But the calculational efficiency and time-consuming of PLF are not fully taken into account. In addition, doubly fed induction generation (DFIG) is widely used in power systems due to flexible control strategies and good performance [28], [29]. But in references above, the relevance between reactive power and active power of wind power is determined by fixed power factor, which cannot reflect the effect of WG’s output characteristic in the PLF method.

In this paper, an improved integrated CM framework is proposed with fully considering the accuracy and efficiency of the above-mentioned three CMs that are GCA, GCC, and ME. The improved integrated CM is based on a probability distribution pre-identification and is used for PLF in a power system with WG. In the method framework, there are two steps. In the first step, two judgment indices (kurtosis and skewness) are calculated. In the second step, more effective CM from GCA, GCC, or ME is selected in PLF by judging whether the indices meet the threshold. To make the PLF method more suitable for the power system with DFIGs, according to the reactive power functions of DFIG under different control strategies, the corresponding elements of the Jacobian matrix are revised in the load flow model. And the PDFs of DFIG outputs are derived under two control strategies including the constant power factor control and the constant voltage control.

The paper is organized as follows. In Section II, several typical CMs are introduced, such as GCA, GCC, and ME. In Section III, PDFs and cumulants of wind power are derived based on the mathematical theory of cumulant and the output characteristics of DFIG. In Section IV, an improved integrated CM for PLF with DFIGs is proposed, where the Kurtosis and skewness of probabilistic distribution are used as pre-identification indices. In Section V, the accuracy and efficiency of the proposed method are verified by the IEEE test system in different operation scenarios. Finally, the conclusions are drawn in Section VI.

II. TYPICAL CUMULANT METHODS IN PROBABILISTIC LOAD FLOW

A. GRAM-CHARLIER EXPANSION OF TYPE A

GCA expresses the PDF of a random variable as a combination of normal distribution and its derivatives [30]. \( p(x) \) is defined as the PDF of a random variable and its expression is as follows:

\[
p(x) = \sum_r c_r H_r(x) \alpha(x)
\]

\[
H_r(x) = (-1)^r \frac{d^r[\alpha(x)]}{dx^r} / \alpha(x)
\]

\[
c_r = \frac{1}{r!} \int_{-\infty}^{\infty} p(x)H_r(x)dx
\]

where \( r \) is the order; \( \alpha(x) \) is the PDF of the standardized normal frequency function; \( x \) is a random variable in standard measure; \( H_r(x) \) is called as Chebyshev-Hermite polynomial; the coefficient \( c_r \) is determined by the cumulants \( \kappa_r \) of \( x \) [30].

A great advantage of GCA is that its coefficients have a very simple relationship with the cumulants of the variables. However, GCA has the uncomfortable drawbacks due to the poor accuracy of the strong non-normal distribution and negative values in certain regions of the PDF.

B. GRAM-CHARLIER OF TYPE C

To overcome drawbacks of GCA sometimes, GCC as an exponential variant of GCA is expressed as follows [31]:

\[
p(x) = \frac{\exp \left[ \sum_r \frac{1}{2} \gamma_r H_r(x) \right]}{\int \exp \left[ \sum_r \frac{1}{2} \gamma_r H_r(x) \right] dx}
\]

where \( H_r(x) \) is same as (1) and \( \gamma_r \) is the unknown coefficient of the series expansion.
The exponential form of GCC ensures that the integral of the variables’ PDF in the defined domain is 1 and the values of the PDF are non-negative.

When \( r = 1, 2, \ldots, k \), the coefficient vector \( \gamma = [\gamma_1, \gamma_2, \ldots, \gamma_k]^T \) is solved by the linear equation as shown in (3):

\[
A\gamma = B
\]  

(3)

where \( B = [b_1, b_2, \ldots, b_k]^T \), \( b_1 = 0 \), \( b_m = -(m - 1) \), \( \psi_{m-2}(m = 2, 3, \ldots, k) \), and Hermite moment \( \psi_m \) is expressed by the moment \( \mu_r \) of \( x \) as shown in (4).

\[
\begin{align*}
\psi_0 &= 1 \\
\psi_1 &= \mu_1 \\
\psi_2 &= \mu_2 - 1 \\
\psi_3 &= \mu_3 - 3\mu_1 \\
\psi_4 &= \mu_4 - 6\mu_2 + 3 \\
\psi_5 &= \mu_5 - 10\mu_3 + 15\mu_1 \\
\ldots
\end{align*}
\]  

(4)

In (3), \( A \) is a \( k \times k \) dimensional symmetric matrix and its element expression is shown as (5), where \( \Delta_{i-1-j-1,m} \) is expressed as (6), when \( i+j+m = \text{even} \) and \( q = (i+j+m)/2 \geq i,j,m \), otherwise, \( \Delta_{i-1-j-1,m} = 0 \).

\[
a_{ij} = \sum_{m=0}^{i+j-2} \frac{1}{m!} \Delta_{i-1-j-1,m} \psi_m
\]  

(5)

\[
\Delta_{i-1-j-1,m} = \prod_{p=i,j,m} \frac{\Gamma(p+1)}{\Gamma(q-p+1)}
\]  

(6)

C. MAXIMUM ENTROPY

ME is a nonlinear PDF reconstruction method that also ensures that the sum of probabilities is equal to unit and the probability values do not exist negative. Under certain constraints, when the entropy reaches the maximum, the probability distribution of the variables is estimated to be the closest to the true distribution. The model of ME is formulated as [26]:

\[
\begin{align*}
\max h(x) &= -\int p(x) \ln p(x) \, dx \\
\text{s.t.} \quad &\int \phi_r(x)p(x) \, dx = \mu_r \\
&\int p(x) \, dx = 1
\end{align*}
\]  

(7)

where \( h(x) \) is the entropy of the random variable \( x \) and \( \mu_r \) is the moment of \( x \). Based on the available input moments, ME fits PDF \( p(x) \) according to the measurement of distribution entropy to achieve the best distribution fit. \( p(x) \) is the PDF of \( x \) which is generally given as (8).

\[
p(x) = \exp \left[ -\sum_r \lambda_r \phi_r(x) \right]
\]  

(8)

In (8), \( \lambda_r \) is the Lagrange multiplier that is solved by combing (7) and (8) with setting the function \( \phi(x) = x^r \). The solving formula of \( \lambda_r \) is shown in (9).

\[
\int x^r \exp \left[ -\sum_r \lambda_r \phi_r(x) \right] \, dx = \mu_r
\]  

(9)

Newton-Raphson iteration method is used to solve the nonlinear equation (9) to obtain \( \lambda_r \).

III. CUMULANTS AND PDFS OF WIND GENERATION BASED ON DFIGS

The cumulant \( \kappa_r \) of the random variables \( x \) is defined as (10) [32].

\[
\exp \left\{ \kappa_1 x + \kappa_2 x^2/2! + \cdots + \kappa_r x^r/r! \right\} = 1 + x + x^2/2! + \cdots + x^r/r!
\]  

(10)

If \( x \) is a continuous random variable satisfying the CDF \( F(x) \), its moment \( \mu_r \) about the mean is obtained by:

\[
\mu_r = \int_{-\infty}^{\infty} x^r dF(x)
\]  

(11)

If \( x_i (i = 1, 2, \ldots, n) \) is a discrete random variable, and its PDF is \( p(x_i) (i = 1, 2, \ldots, n) \), its moment \( \mu_r \) about the mean is obtained by:

\[
\mu_r = \sum_{i=1}^{n} p(x_i) x_i^r
\]  

(12)

In the PLF method, PDFs of power system injected variables (power generation and load) are known, then the moments of the injected are obtained from (11) or (12). According to the equations of moments and the cumulants as shown in (13), the injected cumulants are obtained by the injected moments [30].

\[
\begin{align*}
\kappa_1 &= \mu_1 \\
\kappa_2 &= \mu_2 - \mu_1^2 \\
\kappa_3 &= \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 \\
\kappa_4 &= \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4 \\
\kappa_5 &= \mu_5 - 5\mu_4\mu_1 - 10\mu_3\mu_1 + 20\mu_3\mu_1^2 + 30\mu_2\mu_1^3 - 60\mu_2\mu_1^2 + 24\mu_1^5 \\
\ldots
\end{align*}
\]  

(13)

Assume that the injected variables in power systems are independent of each other. Defining \( \Delta W^r \) as the sum of cumulants of the injected variables in the power system. \( \Delta W^r_0 \) is the cumulant of power generation and \( \Delta W^r_L \) (negative) is the cumulant of load. According to the properties of cumulant, their relationship is shown in (14).

\[
\Delta W^r = \Delta W^r_0 + \Delta W^r_L
\]  

(14)

The PDFs of wind power are defined as \( p(P_{Gw}) \) and \( p(Q_{Gw}) \), and defining \( F(P_{Gw}) \) and \( F(Q_{Gw}) \) are the integral of \( p(P_{Gw}) \) and \( p(Q_{Gw}) \), respectively. In this paper, the total power generated by a wind farm is assumed as the sum
of power generated by each independent DFIG in the farm regardless of the interaction between two wind turbines. The active power $P_{Gw}$ of the wind farm with DFIGs is a function of wind speed $v$ as shown in (15):

$$P_{Gw}(v) = \begin{cases} 0, & v < v_{ci} \\ a_2 v^2 + a_1 v + a_0, & v_{ci} \leq v < v_N \\ P_{GNW}, & v_N \leq v < v_{co} \\ 0, & v \geq v_{co} \end{cases}$$

(15)

where $v_N$ is the rated wind speed; $P_{GNW}$ is the rated active power; $a_0$, $a_1$, and $a_2$ are the coefficients that are obtained by the least square curve fitting method.

Wind speed $v$ satisfies Weibull distribution which is expressed as:

$$F(v) = 1 - e^{-(\frac{v}{\mu})^k}$$

(16)

where the constant $k$ and $c$ are the scale and shape parameters which are obtained by analyzing the historical data [14].

The distributions of $P_{Gw}$ are expressed as follows:

$$p(P_{Gw}) = \begin{cases} F(v_{ci}), & P_{Gw} = 0 \\ F(v_{co}) - F(v_N), & P_{Gw} = P_{GNW} \\ F(P_{Gw}^{-1}(v)), & 0 < P_{Gw} < P_{GNW} \end{cases}$$

(17)

where $P_{Gw}^{-1}(v)$ is the inverse function of (15).

The reactive power from a DFIG is a function of voltage $V_w$ at the bus where the wind farm is located, and is determined by the control strategies of DFIGs [33]:

a) When DFIGs are under the constant power factor control on the stator-side with power factor $\cos \varphi$, the expression of the reactive power $Q_{Gw}$ of the wind farm with DFIGs is shown as:

$$Q_{Gw} \approx Q_1 = P_1 \tan \varphi$$

(18)

$$\cos \varphi = P_1/\sqrt{(P_1)^2 + (Q_1)^2}$$

where $P_1$ and $Q_1$ are the active and reactive power on the DFIG stator-side. The relationship of $Q_{Gw}$, $P_1$, and $P_{Gw}$ is obtained by (19):

$$P_{Gw} = \frac{R_2^2 (X_1 + X_m)^2 (P_1^2 + Q_{Gw}^2)}{\frac{X_2^2 V^2}{X_m^2} Q_{Gw} + (1 - s)P_1 + \frac{R_2^2 V^2}{X_m^2}}$$

(19)

where $P_{Gw}$ is obtained by (15); $R_1$ and $X_1$ are the stator resistance and reactance; $R_2'$ and $X_2'$ are the reduction values of the rotor resistance and reactance; $X_m$ is the excitation reactance; $s$ is the slip.

b) When DFIGs are controlled at a constant voltage $V_w$, $Q_{Gw} \approx Q_1$ satisfies:

$$P_1^2 + (Q_{Gw} + \frac{V_w^2}{X_1 + X_m})^2 = \frac{V_w^2 X_m^2}{(X_1 + X_m)^2} I_2'^2$$

(20)

where $I_2'$ is the reduction value of the rotor-side current.

c) When the reactive power exceeds the adjustable range under the constant voltage control, it satisfies the inequation as shown in (21) where $I_2'$ is the maximum current.

$$P_1^2 + (Q_{Gw} + \frac{V_w^2}{X_1 + X_m})^2 > \frac{V_w^2 X_m^2}{(X_1 + X_m)^2} I_2'^2$$

(21)

Then the function $Q_{Gw}$ is obtained by (19) and (20) with $I_2' = I_{2m}$.

According to the above equations $P_{Gw}$ and $Q_{Gw}$, $Q_{Gw}$ is expressed by $v$ and $V_w$. When defining the equation $Q_{Gw}(v, V_w)$, the distributions are expressed as follows:

$$p(Q_{Gw}) = \begin{cases} F(v_{ci}) + 1 - F(v_{co}), & |Q_{Gw}| = |Q_{Gw_{min}}| \\ F(v_{co}) - F(v_N), & |Q_{Gw}| = |Q_{Gw_{max}}| \end{cases}$$

$$F(Q_{Gw}) = \frac{Q_{Gw} - Q_{Gw_{min}}}{Q_{Gw_{max}} - Q_{Gw_{min}}},$$

(22)

where $Q_{Gw_{min}}$ is the minimum reactive power and $Q_{Gw_{max}}$ is the maximum.

Applying the aforementioned theoretical background and derivation, the moments of the wind farm with DFIGs outputs are expressed as (23)–(24).

$$\mu_r(P_{Gw}) = \int_0^{P_{GNW}} (P_{Gw})^r dF(P_{Gw}) + p(P_{GNW}) \cdot (P_{GNW})^r$$

(23)

$$\mu_r(Q_{Gw}) = p(Q_{Gw_{min}}) \cdot (Q_{Gw_{min}})^r$$

$$+ \int_{Q_{Gw_{min}}}^{Q_{Gw_{max}}} (Q_{Gw})^r dF(Q_{Gw})$$

$$+ p(Q_{Gw_{max}}) \cdot (Q_{Gw_{max}})^r$$

(24)

IV. AN IMPROVED INTEGRATED CM BY PROBABILITY DISTRIBUTION PRE-IDENTIFICATION

At present, CM widely used in literature is mostly based on quasi-normal distribution, such as GCA. When the stochastic nature of the state variable is similar to normal distribution, GCA can obtain accurate probability distribution. But when GCA is used to fit a non-quasi-normal distribution, the negative values appear at the tail of the PDF. However, although GCC and ME are suitable for fitting non-quasi-normal distributions, their calculations are complicated and their time consumptions are relatively long.

In order to make use of the advantages of various typical CMs, an integrated CM for PLF analysis method with WG based on quasi-normal distribution pre-identification is proposed in this section.

A. PRE-IDENTIFICATION INDICES FOR CM SELECTION

The skewness $S_k$ and kurtosis $K_\alpha$ of the distribution reflect the extent of departure from the symmetry of the normal distribution [34]. Through mathematical derivation, the $S_k$ and $K_\alpha$ can be expressed by the variable’s third cumulant $\kappa_3$ and fourth cumulant $\kappa_4$ in standard measure, usually as shown in (25).

$$\begin{cases} S_k = \kappa_3 \\ K_\alpha = \kappa_4 + 3 \end{cases}$$

(25)
The skewness and kurtosis of the standard normal distribution are 0 and 3. When the skewness and kurtosis of the distribution are within a certain range, it can be regarded as a quasi-normal distribution. Therefore, the $S_k$ and $K_u$ are used as pre-identification indices for the next step in the proposed CM framework and the ranges $[S, S_0]$ and $[K, K_0]$ are set as the threshold ranges to pre-identify whether the distribution meets the quasi-normal distribution.

**B. FRAMEWORK OF THE PROPOSED METHOD**

The process framework of the proposed integrated CM for PLF with WG by pre-identification of probability distributions is shown in Fig. 1.

![Flow chart of the proposed method.](image)

In step a of Fig. 1, the cumulants of the DFIGs outputs are obtained by their moments that are derived as (23)–(24). Meanwhile, the Jacobian matrix is calculated to use the state variables' cumulants with the injected variables' cumulants and is obtained by the load flow model with DFIGs. The Jacobian matrix $J$ in this paper is modified as follows according to the output characteristics of DFIG:

In PLF calculation, the relationship between the cumulant of the injected $\Delta W$ and the cumulant of the state variable $\Delta X$ is built as a load flow model as shown in (26) [14].

$$
\Delta X = J^{-1} \Delta W
$$

$$
\Delta X = [\Delta \delta \Delta V \delta \Delta w V w]^{T}
$$

$$
\Delta W = [\Delta P \Delta Q \Delta P_w Q_w]^{T}
$$

In (26), to obtain $J$ at the injected expected value, $\Delta X$ and $\Delta W$ are regard as the correction and error during the load flow iteration; $\Delta \delta$ and $\Delta V$ are the corrections of voltage magnitude and angle; $\Delta P$ and $\Delta Q$ are the errors of active power and reactive power; the subscript “w” indicates the bus where the DFIGs are located.

In a power system with $n$ conventional buses $i$ and $n_w$ buses $w_k$ integrated for the DFIGs, for conventional bus $i$, the errors $\Delta P_i$ and $\Delta Q_i$ are specifically the difference between the constant of injected power and the power function of grid state variables, so the expressions of part elements in the $J$ are shown as:

$$
\frac{\partial \Delta P_i}{\partial \Delta V_i} = - \partial \left[ V_i \sum_{j=1}^{n} V_j \left( G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij} \right) \right] \frac{\partial \Delta V_i}{\partial \Delta V_i}
$$

$$
\frac{\partial \Delta Q_i}{\partial \Delta V_i} = - \partial \left[ V_i \sum_{j=1}^{n} V_j \left( G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij} \right) \right] \frac{\partial \Delta V_i}{\partial \Delta V_i}
$$

(27)

$$
\frac{\partial \Delta Q_{w_k}}{\partial \Delta V_{w_k}} = \partial \left[ V_{w_k} \sum_{j=1}^{n} V_j \left( G_{w_k} \sin \delta_{w_k} - B_{w_k} \cos \delta_{w_k} \right) \right] \frac{\partial \Delta V_{w_k}}{\partial \Delta V_{w_k}}
$$

$$
\frac{\partial \Delta Q_{w_k}}{\partial \Delta V_{w_k}} = - \partial \left[ V_{w_k} \sum_{m=1}^{n_w} V_{w_m} \left( G_{w_k} \sin \delta_{w_k} - B_{w_k} \cos \delta_{w_k} \right) \right] \frac{\partial \Delta V_{w_k}}{\partial \Delta V_{w_k}}
$$

(28)

where $G_{ij} + jB_{ij} = Y_{ij}$ and $Y_{ij}$ is the bus admittance matrix.

For a bus $w_k$ with DFIGs, the reactive power $Q_{Gw_k}$ is as a function of bus voltage $V_{w_k}$ and not a constant. The expressions of $Q_{Gw_k}$ are shown in (18)–(21), so the expression of elements in $J$ is:

$$
\frac{\partial \Delta Q_{Gw_k}}{\partial \Delta V_{w_k}} = \partial \left[ V_{w_k} \sum_{j=1}^{n} V_j \left( G_{w_k} \sin \delta_{w_k} - B_{w_k} \cos \delta_{w_k} \right) \right] \frac{\partial \Delta V_{w_k}}{\partial \Delta V_{w_k}}
$$

$$
\frac{\partial \Delta Q_{Gw_k}}{\partial \Delta V_{w_k}} = - \partial \left[ V_{w_k} \sum_{m=1}^{n_w} V_{w_m} \left( G_{w_k} \sin \delta_{w_k} - B_{w_k} \cos \delta_{w_k} \right) \right] \frac{\partial \Delta V_{w_k}}{\partial \Delta V_{w_k}}
$$

(29)

According to the above, the Jacobian matrix of the power load calculation with DFIGs is obtained.

In step b, to obtain the state variable’s cumulant $\Delta W^r$ in each order for the CM to calculate the PLF result, $\Delta W^r$ is calculated by the injected cumulant $\Delta X^r$ and the Jacobian matrix $J$, which is expressed as:

$$
\Delta X^r = \left( J^{-1} \right)^T \Delta W^r
$$

(30)

In step c, $S_k$ and $K_u$ of a state variable’s distribution with their threshold ranges are used to select a CM to fit the variable’s PDF. In the proposed method, the threshold ranges are derived by quasi-normal distribution [34]. Within the threshold ranges, the variable’s PDF is considered to satisfy quasi-normal distribution and GCA is used,
otherwise, GCC and ME are used which are more suitable for fitting a non-quasi-normal distribution.

In step d, according to the CMs’ introductions in Section II of this paper, the selected CM in step c is to fit the distribution function of the state variable as the PLF result.

V. CASE STUDY

The proposed method described in the preceding sections is applied to an IEEE-14 test system. The network diagram is shown in Fig. 2. A wind farm with DFIGs is connected to bus 2. All the case studies are performed with the MATLAB platform.

![Diagram of IEEE 14-bus test system integrated with DFIGs.](image)

The bus active and reactive basic loads in power system are generally considered to satisfy the normal distribution and their standard deviations are assumed to be 10% [19], [27] or 30% [25] of their expectations. In the below case studies, assume that the standard deviations of all bus active and reactive basic loads are 30% of their expectations.

The capacity of each DFIG is 2MW with the cut-in wind speed \( v_{ci} = 5 \) m/s, the rated wind speed \( v_N = 18 \) m/s, and the cut-out wind speed \( v_{co} = 25 \) m/s. The DFIG’s motor parameters are \( R_1 = 0.0062 \) \( \Omega \), \( X_1 = 0.0024 \) \( \Omega \), \( R_s = 0.0024 \) \( \Omega \), \( X_s = 0.0694 \) \( \Omega \), \( X_m = 4.2233 \) \( \Omega \), and \( s = -0.208 \).

As the comparative baseline, the sampling number \( N \) of MCS is set to be 20000 and ARMS is calculated for comparing the accuracy of different calculation methods. The value of ARMS is defined as:

\[
ARMS = \sqrt{\frac{\sum_{i=1}^{N} (PLF_i - MCS_i)^2}{N}} \quad (31)
\]

where \( PLF_i \) is the \( i^{th} \) point’s value on the CDF calculated by the proposed PLF method; \( MCS_i \) is the \( i^{th} \) point’s value on the CDF calculated by MCS.

The orders of GCA, GCC, and ME are selected as \( 7^{th}, 4^{th}, \) and \( 7^{th} \) in the proposed method, considering both accuracy and efficiency. The \( 7^{th} \) GCA and \( 7^{th} \) ME use from \( 1^{st} \) to \( 7^{th} \) cumulants, and the \( 4^{th} \) GCC uses from \( 1^{st} \) to \( 4^{th} \) cumulants.

### Table 1. Probabilistic results of the line active power distributions.

| Line     | \( 7^{th} \) GCA | \( 4^{th} \) GCC | \( 7^{th} \) ME |
|----------|----------------|----------------|--------------|
| 1-2      | 0.01356        | 0.01340        | 0.01342      |
| 1-5      | 0.02339        | 0.02331        | 0.02323      |
| 2-3      | 0.01607        | 0.01609        | 0.01603      |
| 2-4      | 0.00920        | 0.00919        | 0.00918      |
| 2-5      | 0.00965        | 0.00954        | 0.00954      |
| 3-4      | 0.03081        | 0.03080        | 0.03079      |
| 6-12     | 0.00508        | 0.00503        | 0.00520      |
| 12-13    | 0.00395        | 0.00395        | 0.00389      |

### Table 2. Operate time of each CM.

| Operate time (s) | \( 7^{th} \) GCA | \( 4^{th} \) GCC | \( 7^{th} \) ME |
|-----------------|----------------|----------------|--------------|
| 0.002           | 0.004          | 0.581          |

### A. VALIDITY OF THE PROPOSED METHOD

This case is to verify the necessity of the pre-identification method in this paper. The verification scenario is in case a): bus 2 is integrated with \( 20 \times 2 \) MW DFIGs under a constant power factor (\( \cos \phi = 0.98 \), \( \phi < 0 \)) and wind speed satisfies Weibull distribution with parameters \( k = 2.84 \) and \( c = 7.35 \).

From Table 1, the ARMS values of \( 7^{th} \) GCA, \( 4^{th} \) GCC, and \( 7^{th} \) ME are close, although the calculation accuracy of GCA is generally considered not as good as the other two CMs [27]. It can be seen that the accuracy of GCA results is only slightly lower than other CMs in some cases and the accuracy for reactive power distributions in line 2-3 and line 6-12 is slightly better than that of GCC or ME.

Table 2 shows the time consumption of each CM to obtain bus voltages’ PDFs in the power system.

It is found that the \( 7^{th} \) GCA is the fastest, the \( 4^{th} \) GCC takes twice as long as the \( 7^{th} \) GCA, and the \( 7^{th} \) ME is the slowest due to iterative operations.

In order to clearly show the results of the three CMs, Fig. 3 and Fig. 4 show the PDFs of voltage \( V_7 \) at bus 7 and active power \( P_{1-2} \) in line 1-2, respectively.

From Fig. 3, the fitting errors of the three CMs are all small and similar. But from Fig. 4, the tail part of the PDF fitted by GCA appears negative, which is unreasonable and indicates that the third and fourth cumulants of the variables exceed a certain range according to the above analysis.

In general, GCC and ME have better applicability and accuracy, and GCA has a high calculation speed. In order to realize precision and efficiency in the PLF with the full use of the advantages of CMs, an improved integrated
CM framework is proposed that selects suitable CM for certain variable distribution based on cumulants by pre-identification.

**B. ACCURACY AND EFFICIENCY OF THE PROPOSED METHOD IN DIFFERENT SCENARIOS**

The following verifies the accuracy and efficiency of the proposed method, when the wind farm with DFIGs is under different scenarios, such as different control strategies, different wind speed distributions, and different wind power penetration levels. The case a) in the following scenarios is the same as the situation of the power system in the validity verification above case study. The quasi-normal distribution pre-identification indices are that $S_k$ is within $\pm 1.0493$ and $K_u$ is between 3 and 7 [34].

1) SCENARIO 1: CONTROL STRATEGIES

In this scenario, bus 2 is integrated with $20 \times 2$ MW DFIGs. Wind speed satisfies Weibull distribution with parameters $k = 2.84$ and $c = 7.35$. The DFIGs are under three control strategies:

a) the constant power factor ($\cos \varphi = 0.98$, $\varphi < 0$) control;

b) the constant power factor ($\cos \varphi = 0.8$, $\varphi < 0$) control;

c) the constant voltage control.

Table 3 shows pre-identification indices, selected CM, and ARMS results of part of state variable distributions under different control strategies.

From Table 3, GCA, GCC, or ME is selected to be used in the proposed PLF by range judgment of pre-identification indices. GCC has greater applicability in PLF for power grids with DFIG and the calculation speed is better than ME. Take the first row for example. Under control strategies a) and b), the $K_u$ of the $V_2$'s PDF does not meet the GCA applicable conditions, hence GCC is selected in the proposed CM framework, whereas under control strategies c) the $S_k$ and $K_u$ of the $V_2$'s PDF are both within their threshold ranges so GCA is selected.

It can also be seen from Table 3 that control strategies of DFIGs have different influences on distributions of load flow. Under the control strategy c), most of the state variables’ PDFs are obtained by GCA that is easy implemented. The control strategy a) and b) adopt constant power factor control. In contrast, under the control strategy b) the power factor lag of WG is more serious, hence some state variables need to be fitted by GCC or ME to obtain more accurate PDFs.

To show the method effect more clearly, PDFs of the active power $P_{1-2}$ and reactive power $Q_{1-2}$ in line 1-2 under control strategy b) are drawn in Fig. 5 and Fig. 6.

From Fig. 5 and Fig. 6, the selected CMs that are ME and GCC respectively are closer to MCS. GCA is not selected that appears negative tail part and deviates from MCS. In the same way, Fig. 7 shows PDFs of $P_{1-2}$ in line 1-2, when the
DFIGs are under control strategy c) where the selected method is ME.

Fig. 8 shows PDFs of $Q_{2-4}$ in line 2-4, when the DFIGs are under control strategy c).

In Fig. 8, the selected method is CCA. Although, GCA is considered that the accuracy is slightly lower in many cases. By the pre-identification, GCA is applied to the case of quasi-normal distribution, the calculation accuracy is nearly the same with other CMs, meanwhile, the calculation time is saved.

2) SCENARIO 2: WIND SPEED DISTRIBUTION

In this scenario, bus 2 is integrated with $20 \times 2$ MW DFIGs under constant power factor ($\cos \varphi = 0.98, \varphi < 0$). Wind speed satisfies Weibull distribution with the values of parameters which are as follows:

a) $k = 2.84$ and $c = 7.35$;

b) $k = 2$ and $c = 8.5$.

Table 4 shows pre-identification indices, selected CM, and ARMS results of part of state variable distributions under different wind speeds.

Fig. 9 and Fig. 10 show the PDFs of $V_2$ and $P_{2-4}$ in line 2-4 respectively, while wind speed satisfies Weibull distribution with $k = 2$ and $c = 8.5$.

In scenario 2, it is found that the accuracy of PLF is influenced by the wind speed distribution which directly affects wind power injection. From Table 4, Fig. 9 and Fig. 10, under wind speed b), the distributions of state variables in power system deviate from the quasi-normal distribution to a greater extent and the selected CM gives enough reliable results in the case without negative values appearing at the tail part of PDF.

3) SCENARIO 3: WIND POWER PENETRATION

In this scenario, wind speed satisfies the Weibull distribution with parameters $k = 2.84$ and $c = 7.35$. The control strategy
TABLE 4. Probabilistic results of the state variable distributions.

| State variable | Wind speed | Pre-identification indices | Selected CM | ARMS (%) |
|----------------|------------|----------------------------|-------------|----------|
|                |            | $S_k$ | $K_a$ |                |          |          |
| $V_2$          | a)         | 0.046 | 2.993 | GCC           | 0.0255   |
|                | b)         | 0.063 | 2.988 | GCC           | 0.0325   |
| $V_3$          | a)         | 0.003 | 2.998 | ME            | 0.0171   |
|                | b)         | 0.005 | 3     | GCA           | 0.0194   |
| $V_7$          | a)         | 0     | 3     | GCA           | 0.0086   |
|                | b)         | 0     | 3     | GCA           | 0.0082   |
| $P_{t-2}$      | a)         | -0.035 | 2.996 | GCC           | 0.0144   |
|                | b)         | -0.049 | 2.993 | GCC           | 0.0234   |
| $P_{t+4}$      | a)         | 0.001 | 3     | GCA           | 0.0072   |
|                | b)         | 0.001 | 3     | GCA           | 0.0069   |
| $Q_{t-2}$      | a)         | 0.054 | 2.987 | GCC           | 0.0076   |
|                | b)         | 0.073 | 2.979 | GCC           | 0.0093   |
| $Q_{t+4}$      | a)         | 0.007 | 3     | GCA           | 0.0089   |
|                | b)         | 0.010 | 2.999 | ME            | 0.0074   |

FIGURE 9. The PDFs of voltage $V_2$ under wind speed b).

The installed capacity of the wind farm integrated at bus 2 is as follows:

a) $20 \times 2$ MW;
b) $40 \times 2$ MW.

Table 5 shows pre-identification indices, selected CM, and ARMS results of part of state variable distributions under different wind power penetration.

Table 5. Probabilistic results of the state variable distributions.

| State variable | Wind power penetration | Pre-identification indices | Selected CM | ARMS (%) |
|----------------|------------------------|----------------------------|-------------|----------|
|                |                        | $S_k$ | $K_a$ |                |          |          |
| $V_2$          | a)                     | 0.046 | 2.993 | GCC           | 0.0255   |
|                | b)                     | 0.191 | 2.967 | GCC           | 0.0269   |
| $V_3$          | a)                     | 0.003 | 2.998 | ME            | 0.0171   |
|                | b)                     | 0.025 | 2.998 | GCC           | 0.0212   |
| $V_7$          | a)                     | 0     | 3     | GCA           | 0.0060   |
|                | b)                     | 0.002 | 3     | GCA           | 0.0095   |
| $P_{t-2}$      | a)                     | -0.035 | 2.996 | GCC           | 0.0144   |
|                | b)                     | -0.255 | 2.977 | GCC           | 0.0226   |
| $P_{t+4}$      | a)                     | 0.001 | 3     | GCA           | 0.0072   |
|                | b)                     | 0.013 | 3     | GCA           | 0.0063   |
| $Q_{t-2}$      | a)                     | 0.054 | 2.987 | GCC           | 0.0076   |
|                | b)                     | 0.31  | 2.875 | GCC           | 0.0395   |
| $Q_{t+4}$      | a)                     | 0.007 | 3     | GCA           | 0.0089   |
|                | b)                     | 0.029 | 2.987 | GCC           | 0.0129   |

FIGURE 10. The PDFs of $P_{t-4}$ under wind speed b).

of DFIG is constant power factor ($\cos\phi = 0.98$, $\phi < 0$). The installed capacity of the wind farm integrated at bus 2 is as follows:

a) $20 \times 2$ MW;
b) $40 \times 2$ MW.

In scenario 3, when the penetration rate of wind power increases, the deviation degree of the state variable distribution from the quasi-normal distribution intensifies. Under b) with more wind power penetration, GCC is selected to fit the PDFs of more state variables by pre-identification indices. From Fig. 11 and Fig. 12, the PDFs of $V_2$ and $Q_{1-2}$ obtained by GCA have obvious negative values at the tail while the PDFs obtained by selected method have better accuracy. Overall, in the proposed PLF method, the state variable’s distribution is predicted before calculating, and then a CM is selected to fit the distribution to get accurate results.
VI. CONCLUSION

In order to obtain results of PLF with DFIGs with both accuracy and efficiency, this paper proposes an improved integrated CM by probability distribution pre-identification. Before fitting the distributions of the state variables of power systems, the state variables’ cumulants are used to calculate the pre-identification indices that assess whether the distributions of the state variables meet the quasi-normal distribution. In the proposed CM framework, GCA is used to fit the quasi-normal distribution, and GCC or ME is used to fit the non-quasi-normal distribution. The proposed framework method makes full use of the advantages of various CMs for PLF with DFIGs. It can not only give reliable and efficient PLF results under various random variables but also avoid the negative values at the tail of the non-quasi-normal PDFs. Compared with several typical CMs, the proposed method has enough accuracy and efficiency. In addition, based on the output characteristics of DFIG, the influence of control strategy on the DFIG’s reactive power is considered in the load flow model and the PDF of wind power.

In case studies, the accuracy and efficiency of the proposed method are validated under different WG conditions. It is concluded that when the skewness and kurtosis of the variable’s distribution within the threshold ranges, the selection of GCA can halve the calculation time, and by comparison, the ARMS of the three CMs are almost the same at this time. Therefore, the applicable pre-identification is introduced to predict the distributions of variables, so that the advantage of each CM is flexibly applied to make PLF with DFIG have both computational accuracy and efficiency.

Large-scale wind power integration also affects the stability of the power system, such as voltage stability, low-frequency oscillation, and sub-synchronous oscillation of the power system. The effect of wind uncertainty on stability will be carried out in the future.

APPENDIX

In Table 6, bus 2 is integrated with $20 \times 2$ MW DFIGs under a constant power factor ($\cos \varphi = 0.98$, $\varphi < 0$). Wind speed satisfies Weibull distribution with parameters $k = 2.84$ and $c = 7.35$.

TABLE 6. ARMS of bus voltages’ CDFs under different CMs.

| Order | GCA (%) | GCC (%) | ME (%) |
|-------|---------|---------|--------|
| $4^\text{th}$ | 0.014931 | 0.14931 | 0.014980 |
| $5^\text{th}$ | 0.014963 | 0.14967 | 0.015015 |
| $6^\text{th}$ | 0.014963 | - | 0.014978 |
| $7^\text{th}$ | 0.014957 | - | 0.014962 |
| $8^\text{th}$ | 0.014957 | - | 0.014956 |

REFERENCES

[1] C. Wang, C. Liu, F. Tang, D. Liu, and Y. Zhou, “A scenario-based analytical method for probabilistic load flow analysis,” Electr. Power Syst. Res., vol. 181, Apr. 2020, Art. no. 106193.
[2] M. M. Sany, M. I. Mosaad, M. F. El-Naggar, and S. Barakat, “Reliability support of dependable grid using green energy systems: Economic study,” IEEE Access, vol. 9, pp. 14528–14539, 2021.
[3] K. Wu, J. Li, B. Zhang, Z. Yu, and X. Liu, “Preventive dispatch strategy against FDIA induced loads overloads in power systems with high wind penetration,” IEEE Access, vol. 8, pp. 210452–210461, 2020.
[4] G. Ren, J. Liu, J. Wan, Y. Guo, and D. Yu, “Overview of wind power intermittency: Impacts, measurements, and mitigation solutions,” Appl. Energy, vol. 204, pp. 47–65, Oct. 2017.
[5] M. Ebeed, A. Ali, M. I. Mosaad, and S. Kamel, “An improved lightening attachment procedure optimizer for optimal reactive power dispatch with uncertainty in renewable energy resources,” IEEE Access, vol. 8, pp. 168721–168731, 2020.
[6] J. Xu, P. K. Kanyingi, K. Wang, G. Li, B. Han, and X. Jiang, “Probabilistic small signal stability analysis with large scale integration of wind power considering dependence,” Renew. Sustain. Energy Rev., vol. 69, pp. 1258–1270, Mar. 2017.
[7] S. D. Ahmed, F. S. M. Al-Ismail, M. Shafiuullah, F. A. Al-Sulaiman, and I. M. El-Amin, “Grid integration challenges of wind energy: A review,” IEEE Access, vol. 8, pp. 10857–10878, Jan. 2020.
[8] X. Guo, H. Bao, J. Xiao, and S. Chen, “A solution of interval power flow considering correlation of wind power,” IEEE Access, vol. 9, pp. 78915–78924, 2021.
[9] B. Borkowska, “Probabilistic load flow,” IEEE Trans. Power App. Syst., vol. PAS-93, no. 3, pp. 752–759, May 1974.
[10] A. A. E. Tawfiq, M. O. A. El-Raouf, M. I. Mosaad, A. F. A. Gawad, and M. A. E. Farahat, “Optimal reliability study of grid-connected PV systems using evolutionary computing techniques,” IEEE Access, vol. 9, pp. 42125–42139, 2021.
[11] M. Hossain, A. Abu-Siadah, and S. Muyeen, “Methods for advanced wind turbine condition monitoring and early diagnosis: A literature review,” Energies, vol. 11, no. 5, p. 1309, May 2018.
[12] R. Ma, X. Li, W. Gao, P. Lu, and T. Wang, “Random-fuzzy chance-constrained programming optimal power flow of wind integrated power considering voltage stability,” IEEE Access, vol. 8, pp. 217957–217966, 2020.
[13] Y. Zhang, J. Wang, and X. Wang, “Review on probabilistic forecasting of wind power generation,” Renew. Sustain. Energy Rev., vol. 32, pp. 255–270, Apr. 2014.
[14] N. Gupta, “A review on the inclusion of wind generation in power system studies,” Renew. Sustain. Energy Rev., vol. 59, pp. 530–543, Jun. 2016.
[15] U. H. Ramadhani, M. Shepero, J. Munkhammar, J. Widén, and N. Etherden, “Review of probabilistic load flow approaches for power distribution systems with photovoltaic generation and electric vehicle charging,” Int. J. Electr. Power Energy Syst., vol. 120, Sep. 2020, Art. no. 106003.
[16] C. Wang, D. Liu, F. Tang, and C. Liu, “A clustering-based analytical method for hybrid probabilistic and interval power flow,” Int. J. Electr. Power Energy Syst., vol. 126, Mar. 2021, Art. no. 106005.
[17] B. R. Prusty and D. Jena, “A critical review on probabilistic load flow studies in uncertainty constrained power systems with photovoltaic generation and a new approach,” Renew. Sustain. Energy Rev., vol. 69, pp. 1286–1302, Mar. 2017.

[18] P. Zhang and S. T. Lee, “Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion,” IEEE Trans. Power Syst., vol. 19, no. 1, pp. 676–682, Feb. 2004.

[19] X. Zha, W. Liu, and J. Zhang, “Probabilistic load flow method considering large-scale wind power integration,” Proc. CSEE, vol. 33, no. 7, pp. 77–85, Mar. 2013.

[20] G. Tyl, J. Baldyga, M. Bouaifi, and M. Jasińska, “Population balance approach to model owald ripening of silica using Gram-Charlier series expansion based closure,” Chem. Eng. Res. Des., vol. 159, pp. 491–504, Jul. 2020.

[21] J. Usaola, “Probabilistic load flow with wind production uncertainty using cumulants and Cornish–Fisher expansion,” Int. J. Electr. Power Energy Syst., vol. 31, no. 9, pp. 474–481, Oct. 2009.

[22] J. Usaola, “Probabilistic load flow in systems with wind generation,” IET Gener. Transmiss. Distrib., vol. 3, no. 12, pp. 1031–1041, Dec. 2009.

[23] F. J. Ruiz-Rodriguez, J. C. Hernández, and F. Jurado, “Probabilistic load flow for photovoltaic distributed generation using the Cornish–Fisher expansion,” Electr. Power Syst. Res., vol. 89, pp. 129–138, Aug. 2012.

[24] T. Williams and C. Crawford, “Probabilistic load flow modeling comparing maximum entropy and Gram-Charlier probability density function reconstructions,” IEEE Trans. Power Syst., vol. 28, no. 1, pp. 272–280, Feb. 2013.

[25] B. Sui, K. Hou, H. Jia, Y. Mu, and X. Yu, “Maximum entropy based probabilistic load flow calculation for power system integrated with wind power generation,” J. Mod. Power Syst. Clean Energy, vol. 6, no. 5, pp. 1042–1054, Sep. 2018.

[26] Q. Bian, Y. Qiu, W. Wu, H. Xin, and X. Fu, “Generation dispatch method based on maximum entropy principle for power systems with high penetration of wind power,” J. Mod. Power Syst. Clean Energy, vol. 6, no. 6, pp. 1213–1222, Nov. 2018.

[27] Y. Gao and C. Wang, “Limitation analysis and improvement for series expansion methods to fit the distribution function of probabilistic power flow,” Proc. CSEE, 2021. [Online]. Available: http://kns.cnki.net/kcms/detail/11.2107.TM.20210118.1854.003.html

[28] A. M. S. Yunus, A. Abu-Siada, M. A. S. Masoum, M. F. El-Naggar, and J. X. Jin, “Enhancement of DFIG LVRT capability during extreme short-wind gust events using SMES technology,” IEEE Access, vol. 8, pp. 47264–47271, 2020.

[29] M. I. Mosaad, “Comparative study between the electrical generators used in wind energy conversion systems,” Int. J. Energy, vol. 14, pp. 88–92, Dec. 2020.

[30] M. G. Kendall and A. Stuart, “Systems of distributions,” in The Advanced Theory of Statistics, vol. 1, 4th ed. London, U.K.: C. Griffin and Co., Ltd., 1977, pp. 159–192.

[31] G. Muscolino and G. Ricciardi, “Probability density function of MDOF structural systems under non-normal delta-correlated inputs,” Comput. Methods Appl. Mech. Eng., vol. 168, nos. 1–4, pp. 121–133, Jan. 1999.

[32] M. G. Kendall and A. Stuart, “Moments and cumulants;,” in The Advanced Theory of Statistics, vol. 1, 4th ed. London, U.K.: C. Griffin and Co., Ltd., 1977, pp. 57–96.

[33] D. Ming, L. Binbin, and H. Pingping, “Impacts of doubly-fed wind turbine generator operation mode on system voltage stability,” in Proc. 2nd Int. Symp. Power Electron. Distrib. Gener. Syst. (PEDG), Hefei, China, Jun. 2010, pp. 602–606.

[34] E. Jondeau and M. Rockinger, “Gram-Charlier densities,” J. Econ. Dyn. Control, vol. 25, no. 10, pp. 1457–1483, Oct. 2001.