

Article



Probabilistic Load Flow Method Based on Modified Latin Hypercube-Important Sampling

Quan Li^{1,*}, Xin Wang² and Shuaiang Rong³

- ¹ Huanggang Power Supply Company, Hubei Electric Power Co., Ltd., State Grid Corporation of China, Huanggang 438000, China
- ² Center of Electrical & Electronic Technology, Shanghai Jiao Tong University, Shanghai 200240, China; wangxin26@sjtu.edu.cn
- ³ College of Electrical Engineering, Shanghai University of Electric Power, Shanghai 200082, China; jessica0320@mail.shiep.edu.cn
- * Correspondence: liquan9388@gmail.com; Tel.: +86-021-3420-4622

Received: 21 October 2018; Accepted: 13 November 2018; Published: 15 November 2018



Abstract: The growing amount of distributed generation has brought great uncertainty to power grids. Traditional probabilistic load flow (PLF) algorithms, such as the Monte-Carlo method (MCM), can no longer meet the needs of efficiency and accuracy in large-scale power grids. Latin Hypercube Sampling (LHS) develops a sampling efficiency and solves the correlation problem of distributed generation (DG) access nodes for accuracy analyses. In this paper, a modified Latin Hypercube-Important Sampling method is proposed for higher efficiency and precision by using the importance sampling method before LHS and the Cholesky decomposition method in correlation calculations. The simulation results are presented using a modified IEEE 30-bus system and are compared with traditional MCM and LHS.

Keywords: probabilistic load flow; important sampling; non-parametric estimation construction; Latin hypercube sampling

1. Introduction

With the gradual depletion of fossil fuels, discussions regarding renewable energy have dominated research in recent years. Distributed generations (DGs), represented by wind power systems and solar energy systems, have developed rapidly. More and more DGs are connected to the modern power grid, which can, not only make full use of natural resources and reduce the cost of power generation, but also improve the power supply capacity of the power grid [1]. However, the output of DGs is seriously affected by uncertain weather conditions. Hence, DGs will bring many uncertainties to a power system after accessing the grid [2]. These uncertainties will especially affect the accuracy of power flow calculations. Traditional power flow algorithms cannot evaluate the uncertain impacts of DGs [3]. Hence, load flow algorithms dealing with uncertainties with DGs in power networks has been a hot issue for decades.

The probabilistic load flow (PLF) algorithm [4] was first proposed in 1974 by Borkowskaand. The PLF algorithm aims to determine the statistical data of power flow distribution, including the mean, variance of flow distribution, and other probabilistic characteristics. Different from traditional power flow algorithms, the PLF algorithm regards the power inputs of DGs as uncertain factors. With the probability theory being used to process the uncertainties of a power system, PLF presently has an irreplaceable advantage in the steady-state analysis of power networks.

After decades of development, the PLF algorithm is now classified into three main categories: The approximate method, analytic method and Monte-Carlo method.

The approximate method generally reduces repeated sampling with mathematical methods. The point estimation method (PEM) [5] is a typical approximate method, and is usually classified into two-point estimation [6] and three-point estimation [7]. Though PEM can be the fastest among the various PLF methods in small-scale calculations, its error is high in large-scale power grids. In the analytic method, convolution calculations are adopted to obtain the probability distribution of state quantities and, similarly, it is very complex and inefficient in a large system. In Reference [8], the fast Fourier transformation method was applied to improve the computation speed of convolution. Some other methods put forward a semi-invariant convolution processing method to develop efficiency [9]. However, the calculation speeds in [8,9] decreased remarkably as the number of DG access increases. This indicates that the approximate method and the analytic method are not applicable in a large-scale power system.

However, the Monte-Carlo method (MCM) is more applicable to probabilistic calculations in large power networks. MCM adopts large-scale sampling to obtain experimental samples, which are then randomly simulated several times for analysis [10]. Therefore, the calculation accuracy of MCM is usually high, regardless of the scale of the power system [11]. However, the traditional MCM method has some defects. The error of MCM is large when the sampling times is low, and the calculation speed of MCM is slow when the sampling times is high because of the iterations. More importantly, traditional MCM does not consider correlations among DGs' nodes, which reduces the accuracy of the algorithm.

Therefore, some improved algorithms, such as Gaussian mixture model algorithms [12], the K-means data clustering algorithm [13,14] and ninth-order polynomial transformation algorithms [15] have emerged to improve the calculation speed of MCM. However, these algorithms are not accurate enough for low sampling times. In Reference [16], a sequence method was applied to reduce the iteration times; however, correlation was not considered in this method. In Reference [17], a uniform sampling method was adopted to solve the correlation problem, but the Copula function became very complex when the correlation nodes exceeded two and vitally limited the computation efficiency. The Latin hypercube sampling (LHS) method referred to in Reference [18] can yield a high accuracy with few sampling times, regardless of the size of the power grid. The LHS method is mainly conducted in two steps; the first step is to determine the range of sampling, and obtain the probability distribution of random variables, and the second step is to deal with the correlation problem.

In this paper, the Latin hypercube-important sampling method (LHISM) is presented as an improved LHS algorithm. Firstly, the important sampling method (ISM) is used to narrow the sampling range of LHS. Thus, the efficiency of LHS is greatly improved with the pre-sampling progress. Then, the distribution function of new random variables is obtained using nonparametric estimation. This provides a more precise and practical description of input variables for LHS. Finally, LHS is carried out and the Cholesky decomposition method is adopted to simplify the correlation calculations. Simulations under two kinds of constraints in practical situations were carried out in an IEEE 30-bus system. The results demonstrated that LHISM has a high accuracy with low sampling times and costs less time for the same precision, with regards to expectation and variance, compared with the traditional MCM and LHS methods.

This paper varies from Reference [19], which is simpler when dealing with more than two correlative random variables. This paper also differs from References [20–22] by using the important sampling method to narrow the sampling range and build the distribution function of random variables in advance for LHS. In addition, the Cholesky method is used to ensure an effective process for multiple codes.

The rest of this paper is organized as follows. Section 2 provides an introduction to the ISM, LHS, and Cholesky methods. Section 3 describes the procedure of applying LHISM to the PLF problem in two situations and presents the simulation results of the proposed scheme. Finally, Section 4 concludes the paper.

2. Latin Hypercube-Important Sampling Method

LHISM is an approach that combines LHS and the important sampling method to deal with PLF problems. In this section, the principle of ISM is presented, and then the nonparametric estimation based on the kernel function is introduced to obtain the distribution function of new variances in ISM. Next, the details of the Cholesky decomposition method are given for improving the LHS method.

2.1. Important Sampling Method

The principle of ISM is to avoid random sampling and to concentrate the samples in a desired range, so as to improve the efficiency of an algorithm. There are two problems to solve in ISM; one is the selection of the desired range, and the other is to build a new probability density function for sampling in the desired range.

Most MCM algorithms sample the whole random variable distribution interval. This paper presents the concept of a desired range. A desired range is a reduced sampling space and samples collected in this sampling space are more helpful to PLF calculations than those collected outside the range. Usually, the selection of the desired range needs to consider the actual situation, such as the condition of random variables or calculation purpose. For example, a wind turbine is often shut down or disconnected when wind speeds exceed the threshold. Suppose the limited speed is V_{ci} to V_{co} , Figure 1 shows some samples of wind speeds collected from two wind farms (WF) in the same period of time. The abscissa and ordinate denote the wind speeds as WF1 and WF2, respectively.



Figure 1. Samples of wind speeds.

Obviously, many invalid samples are located outside the range of V_{ci} to V_{co} . Therefore, in one case, we put the sampling range in the range of V_{ci} to V_{co} or even in a smaller range to improve the efficiency of the sampling. If a group of discrete samples is in the desired range, these samples are used as important samples. In addition, we are often concerned with whether the current in important lines exceeds the limitation after DG access. If the current exceeds the limitation, the operation of grid needs to be adjusted. Therefore, in order to analyze the over-limit probability of important lines, the desired range can be placed near the line limit value. Similarly, if the resulting current value of PLF is in the vicinity of the limit value, this group of discrete samples is selected as the "important samples". The specific range is set out in the simulation in Section 3.

Here, we address the second problem in ISM. A new probability distribution function with the same probability properties of the original one is needed. In particular, the following principle was

introduced to keep the new random variable with the same expectation and smaller variance of the old one.

The probability of the original random variable *Z* was p(z). h(z) is another probability density function regarding *z*, which we call important distribution function (IDF). According to the characteristics of probability density function, $\sum_{z \in Z} h(z) = 1$. Supposing the new function $g^*(z)$:

$$g^*(z) = \frac{g(z)p(z)}{h(z)} \tag{1}$$

where g(z) is used to represent the power flow equation (the Newton–Raphson method).

Theorem 1. g(z) and $g^*(z)$ share the same mean but different variances.

Proof of Theorem 1.

$$E(g^*(z)) = \sum_{z \in Z} \frac{g(z)p(z)}{h(z)} \cdot h(z) = \sum_{z \in Z} g(z)p(z) = E(g(z))$$
(2)

$$D(g^{*}(z)) = E[g^{*}(z) - E(g^{*}(z))]^{2} = E\left[\frac{g(z)p(z)}{h(z)} - E(g^{*}(z))\right]^{2}$$

= $h(z) \cdot \sum_{z \in Z} \left[\frac{g(z)p(z)}{h(z)} - E(g(z))\right]^{2}$ (3)

Equations (2) and (3) show that we can use a new random variable, $g^*(z)$, to replace the original random variable, g(z). Here, $h_z(z) = \frac{g(z)p(z)}{E(g(z))}$ is an ideal situation of h(z), and we named it as the optimal IDF. In this situation, $D(g^*(z)) = 0$, so that the means of $g^*(z)$ would be obtained by only one sampling. However, E(g(z)) cannot be solved because variable *Z* is a continuous random variable. Thus, we use sample reconstruction to approach the optimal IDF. Next, a non-parametric reconstruction method based on kernel function is adopted.

2.2. Reconstruction of Optimal Important Distribution Function Based on Non-Parametric Estimation

This section introduces the non-parametric reconstruction based on the kernel function to approach the optimal IDF.

 $Z_1, Z_2, \overline{Z}_3, \ldots, Z_M$ are the important samples collected from variable Z. $\hat{h}_z(Z) = \frac{1}{Mh_n} \sum_{i=1}^M K\left(\frac{z-Z_i}{h_M}\right)$ is called the reconstructed distribution function (RDF), where function K(z) is the kernel function and h_M is the bandwidth of K(z). M is the number of samples. The properties of the kernel function are given in Reference [23], and proved that the value of h_M determines the similarity between RDF and optimal IDF. We used a Gaussian distribution to replace K(z):

$$\hat{h}_{z}(z) = \frac{1}{\mathrm{Mh}_{M}} \sum_{i=1}^{M} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{z-Z_{i}}{h_{M}}\right)^{2}} \right)$$
(4)

The method of integral square-error (ISE) was adopted for the choosing optimal bandwidth. Define ISE as:

$$ISE(\hat{h}_{z}(z)) = \int (\hat{h}_{z}(z) - h_{z}(z))^{2} dz = \int (\hat{h}_{z}^{2}(z) - 2\hat{h}_{z}(z)h_{z}(z) + h_{z}^{2}(z)) dz$$
(5)

Optimal bandwidth h_M is what made ISE $(\hat{h}_z(z))$ the smallest. In order to simplify the calculation, we performed some approximate processing for Equation (5). Firstly, the last item $h_z^2(z)$ was independent from h_M ; thus, we focused on the first two items of ISE $(\hat{h}_z(z))$. Supposing:

$$R(\hat{h}_{z}(z)) = \int \hat{h}_{z}^{2}(z)dz - 2\int h_{z}(z)\hat{h}_{z}(z) = R_{1} - 2R_{2}$$
(6)

where $R_1 = \frac{1}{M^2} \sum_{i,j}^{M} \frac{1}{\sqrt{2\pi h_M^2}} e^{-\frac{1}{2} \left(\frac{Z_i - Z_j}{h_M}\right)^2}$, and the second term, R_2 , can be derived from the samples collected by the small-scale Monte-Carlo sampling in the desired range:

$$R_{2} = \hat{h}_{z}(z)h_{z}(z) = E\left[\hat{h}(z)\right] \approx \frac{1}{M}\sum_{i=1}^{M}\hat{h}_{-i}(Z_{i})$$
(7)

where $\hat{h}_{-i}(Z_i) = \frac{1}{M-1} \sum_{j=1, j \neq i}^{M} \frac{1}{\sqrt{2\pi}h_M} e^{-\frac{1}{2}(\frac{Z_i - Z_j}{h_M})^2}$. The -i that exists in the subscript represents that it is not contained in the *i*th item in the sum. Details of the approximation in Equation (8) are in the following:

$$E\left[\frac{1}{M}\hat{h}_{-i}(Z_i)\right] = \frac{1}{M}\sum_{i=1}^{M} E_{Z_i}\left[\hat{h}_{-i}(Z)\right] = E_z\left[\hat{h}_{-i}(Z)\right] = \int \hat{h}(z)h(z)dz$$
(8)

Then, bringing these results to Equation (6),

$$R(\hat{h}(z)) = \frac{1}{M^2} \sum_{i,j}^{M} \frac{1}{\sqrt{2\pi h_M^2}} e^{-\frac{1}{2}(\frac{Z_i - Z_j}{h_M})^2} - \frac{2}{M} \sum_{i=1}^{M} \hat{h}_{-i}(Z_i)$$
(9)

Finally, we get an ideal bandwidth by minimizing the value of Equation (9). The RDF was also calculated in sequence.

To study the effect of M on the accuracy of RDF, some samples in different M were selected from 2000 wind speed samples in a certain range, and constructed according to Equations (5)–(9). The probability distribution curves of the IDF constructed using different M are compared in Figure 2, and we took the spline fitting curve as the standard distribution curve.



Figure 2. Comparison of the cumulative distribution curves.

As shown in Figure 2, the curve in M = 100 is very close to the standard distribution curve. Therefore, we used M = 100.

2.3. Latin Hypercube Sampling with Cholesky Decomposition Method

After obtaining the RDF of each new variable, we adopted LHS to sample the new random variable. Then, Cholesky decomposition was used to minimize the correlation of samples, which was sampled using LHS.

LHS is a stratified sampling method based on the Monte-Carlo method; the range of cumulative distribution function $F_k(Z)$ is divided into N intervals, in which the representative sample is extracted to replace the whole sampling range. LHS can greatly improve the efficiency of sampling, especially the accuracy of an algorithm under low sampling times. The details of the sampling steps of LHS has been given in Reference [18]. In this way, an original sampling matrix, $Z_0(n \times N)$, which contains a correlation is obtained. The number of input random variables was n, and the times of LHS sampling was N. The next process is to reorder the Z_0 for its minimal correlation by using the Cholesky decomposition method. The detailed theory and related proofs of the Cholesky decomposition method are as follows:

Let C_Z be the correlation coefficient matrix of input random variables $Z_1, Z_2, Z_3, \ldots, Z_M$.

$$C_{Z} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{12} \\ \rho_{12} & 1 & \cdots & \rho_{12} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{12} & \rho_{12} & \cdots & 1 \end{bmatrix}$$
(10)

where ρ_{ij} is the correlation coefficient of Z_i and Z_j .

There is a set of standard normal random variables, $E_1, E_2, ..., E_n$, with a one-to-one mapping between E_n and Z_n :

$$E_k = \Phi^{-1}(F_k(Z_k)), \ k = 1, 2, \cdots, n$$
(11)

Accordingly, let the correlation coefficient matrix of E_1, E_2, \ldots, E_n be C_E .

$$C_{E} = \begin{bmatrix} 1 & \rho_{12}' & \cdots & \rho_{1n}' \\ \rho_{21}' & 1 & \cdots & \rho_{2n}' \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1}' & \rho_{n2}' & \cdots & 1 \end{bmatrix}$$
(12)

Off-diagonal elements of C_E also have a one-to-one mapping with C_X :

$$\rho_{ij} = G\left(\rho'_{ij}\right) \cdot \rho'_{ij} \tag{13}$$

where conversion coefficient $G(\rho'_{ij})$ can be obtained by numerical integration and the search method in Reference [24]. The empirical expression of $G(\rho'_{ij})$ was given in Reference [25].

Then, perform Cholesky decomposition for C_E :

$$C_E = BB^T \tag{14}$$

Assuming that vector Y consists of n independent random variables, Y_i (i = 1, 2, ..., n), that obey standard normal distribution, let:

$$M = \begin{bmatrix} M_1 \\ \vdots \\ M_n \end{bmatrix} = BY = \begin{bmatrix} b_{11} & & \\ \vdots & \ddots & \\ b_{n1} & \dots & b_{nn} \end{bmatrix} \times \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$$
(15)

Theorem 2. In Equation (15), M_1, M_2, \ldots, M_n obey the standard normal distribution.

Proof of Theorem 2. The linear combination of standardized normal random variables also obeys the standard normal distribution. For any variables, Z_i (i = 1, 2, ..., n), there is

$$E(Z_i) = E\left(\sum_{j=1}^i b_{ij} \cdot Y_j\right) = \sum_{j=1}^i b_{ij} \cdot E(Y_j) = 0$$
(16)

The correlation coefficient matrix of Z_1, Z_2, \ldots, Z_n is:

$$\rho\left(\mathbf{Z}\mathbf{Z}^{T}\right) = \rho\left(\mathbf{B}\mathbf{Y}\mathbf{Y}^{T}\mathbf{B}^{T}\right) = B \cdot \rho\left(\mathbf{Y}\mathbf{Y}^{T}\right) \cdot B^{T} = \mathbf{B}\mathbf{B}^{T} = \mathbf{C}_{E}$$
(17)

Therefore, we performed random sampling for M, obeying the standardized normal distribution and obtained M^* as the order matrix. Then, the original sampling matrix using LHS was arranged according to the order of elements in M^* , and the final sample matrix was obtained.

3. Performance Analysis

To investigate the performance of the LHISM method, a modified IEEE-30 bus system [26] was used in the MATLAB platform (2016b, The MathWorks, Inc., Natick, MA, USA). Firstly, the steps of LHISM were given and then we set the input variables and the desired range for two different situations. The simulation results and analyses are presented.

3.1. Steps

• STEP 1:

The first step was to set parameters. The wind speed of each wind farm was taken as a random variable. The desired range and input the value of *M* of the required samples was set.

• STEP 2:

The second step was to obtain important samples. Single random sampling was done for random variables. If the samples or according load flow values of samples were located in the desired range, these samples were taken as important samples. This step was repeated until each random variable collected *M* important samples.

• STEP 3:

The third step was to obtain reconstructed distribution functions (RDFs). The discrete means of each random variable were obtained by taking the important samples in the power flow calculation equation. Then, the optimal bandwidth was calculated and the RDFs for each input variable were acquired.

• **STEP 4**:

The fourth step was to use the LHS and Cholysky method to sample the data and address the correlation problems. Each RDF was processed using LHS to generate $Z_0(n \times N)$, and vector M was formed according to the Cholesky decomposition method, and M was randomly sampled to form order matrix M^* , then the order matrix was used to reorder Z_0 and get the final sample matrix.

• STEP 5:

The last step was to carry out a deterministic power flow calculation and data analysis of the results. The power flow was calculated using the final sample matrix. Then, the statistical results and

the probability distribution of the power flow were obtained. Finally, the corresponding conclusions of power flow were given.

A flowchart of the LHISM algorithm is shown in Figure 3.



Figure 3. Flowchart of the Latin hypercube-important sampling method (LHISM) algorithm.

3.2. The Setting of Desired Range

The active output (P_W) model of a single wind turbine can be simplified as a piecewise function, as shown in Equation (18):

$$P_{W} = \begin{cases} 0 & V_{s} < V_{ci} \text{ or } V_{s} > V_{co} \\ P_{WR} \frac{V_{s} - V_{ci}}{V_{r} - V_{ci}} & V_{ci} \le V_{s} < V_{r} \\ P_{WR} & V_{r} \le V_{s} \le V_{co} \end{cases}$$
(18)

where P_{WR} is the active output of the wind turbine, and P_{WR} is the rated power. V_s is the wind speed; V_{ci} , V_r , V_{co} are the cut-in wind speed, rated wind speed, cut-out wind speed, respectively. In practice, a Weibull distribution is usually used to model instantaneous wind speeds. The curve of the probability distribution of the instantaneous wind speeds, which is modeled using a Weibull distribution, is given in Figure 4. Its scale parameter is 10 and the shape parameter is 2.15.



Figure 4. Distribution curve of instantaneous wind speeds.

As shown in Figure 5, an IEEE-30 bus system is modified for simulation analyses.



Figure 5. Structure of modified IEEE-30 bus system.

Nodes 14, 15, 16, 17, 18 are linked to a wind farm (WF), and the rated capacity of each WF is $P_{\text{WR}} = 40 \text{ MV}$. V_{ci} , V_r , V_{co} are the cutting wind speed, rated wind speed, and cutting speed. Their values are 2.5, 12, and 20 m/s, respectively.

The correlation coefficient matrix of the five WFs is:

$$C_Z = \begin{bmatrix} 1.0 & 0.5 & 0.8 & 0.7 & 0.6 \\ 0.5 & 1.0 & 0.7 & 0.6 & 0.8 \\ 0.8 & 0.7 & 1.0 & 0.5 & 0.6 \\ 0.7 & 0.6 & 0.5 & 1.0 & 0.9 \\ 0.6 & 0.8 & 0.6 & 0.9 & 1.0 \end{bmatrix}$$
(19)

1. Situation I:

As described in Section 2, when the wind speed is lower or higher than a threshold, power generation is often shut down or disconnected. Therefore, this situation determines whether a sample of wind speed is in the desired range, determined by its output power:

$$0.3P_{\rm WR} \le P_W \le P_{\rm WR} \tag{20}$$

2. Situation II:

In situation II, the main consideration was the current values of the main lines after DG access. For example, this simulation only considered the limitation of the current of the line between node 3 and node 4.

$$\begin{cases} I_L = 1.3 p.u. \\ 0.9 I_L \le I_{34} \le 1.1 I_L \end{cases}$$
(21)

 I_{34} is the actual current of the line between node 3 and node 4, and I_L is maximum allowable current for this line. If the actual current conforms to Equation (21), the corresponding wind samples are regarded as important samples.

3.3. Simulation Result of Situation I

The aims of a power flow calculation usually include a line's active power (P), a line's reactive power (Q), a node's voltage (V), and the angle of voltage and current, θ .

This example takes the result of MCM simulation with 20,000 sampling as the standard value of PLF. The standard value of the statistical mean and standard deviation are expressed as μ_a and σ_a . The statistical mean and standard deviation of sample are expressed as μ_s and σ_s , respectively. Accordingly, the accuracy of the LHISM algorithm can be estimated by the relative error rate (RET):

$$\operatorname{RET}_{\mu} = \left| \frac{\mu_a - \mu_s}{\mu_a} \right| \times 100\%$$
(22)

$$\operatorname{RET}_{\sigma} = \left| \frac{\sigma_a - \sigma_s}{\sigma_a} \right| \times 100\%$$
(23)

Set M = 100, and follow the steps described in Figure 4 under the Situation I. In order to show the performance of LHISM, MCM sampling in the desired range was adopted for comparison. The contrast algorithm is referred to as the Monte-Carlo important sampling method (MCISM). Ten simulations were carried out using the same data and settings. The average RETs of these simulations are shown in Figure 6.

The following three conclusions are drawn from Figure 6:

 LHISM had a higher accuracy compared with MCISM, especially for the results of *P* and *θ*. The mean accuracy of LHISM was more than 60% higher than that of MCISM in the case of small sampling times (N = 100) and was about 10% higher in the case of large sampling times (N = 1000).

- Compared with the traditional MCM, LHISM has a very high accuracy in calculating the standard deviation of tidal current results. The precision of the standard deviation obtained using LHISM was more than 40% higher than that of the traditional method, both in terms of small or large sampling times.
- LHISM had a more stable accuracy than MCISM. The variation in LHISM errors between high sampling times and low sampling times was not more than 30%, and was under 35–50% in MCISM.



Figure 6. Error comparison of power flow results between two methods.

3.4. Simulation Result of Situation II

In Situation II, the cumulative distribution curve of current I_{34} was obtained by sampling LHS 500 times. The curve was also compared with the results of the MCM simulation with 20,000 sampling.

Figure 7 shows that the probability of exceeding the limit of I_{34} was 14.9% using LHISM. If the limiting probability obtained by the MCM simulation is regarded as the standard value, the RET is only 5.65%.

Assume the reliable average RET of limiting probability is 0.1%. The LHS algorithm in Reference [19] was used as a comparison. The sampling times and the computation times required for the two methods for achieving a reliable average error are given in Table 1.

Table 1 shows that LHISM took less time to compute compared with the LHS method, and the computing speed of LHISM was about twice as fast as that of the LHS method.



Figure 7. Cumulative distribution curve of *I*₃₄.

Table 1. Comparison of sampling times and calculation times of the two methods.

Method	Sampling Times	Calculation Time (s)
LHISM	407	732
LHS-Copula	789	1503

4. Conclusions

A novel PLF method called LHISM has been proposed in this paper, based on ISM and the modified LHS method. In ISM, the sampling space is narrowed to the desired range. In addition, a new random variable that retains the probability characteristics of the original random variable is constructed using non-parametric estimation. The smaller variances of the new random variables contribute to improving the sampling efficiency. Then, modified LHS are introduced to improve the accuracy under low sampling times and the speed of correlation processing.

The advantage of LHISM is that it, not only keeps the advantages of LHS, such as accuracy with low sampling times, but also further narrows the sampling range. In addition, a simpler method for correlation processing is also applied. Our simulations verified that LHISM has a higher accuracy and computational efficiency compared with the MCM algorithm, especially with short sampling times. In practical applications, LHISM can evaluate the power flow limit after DG access, with a shorter computation time than the LHS method.

Due to the limitations of data, this paper only considered a grid model with fan access. The principle of this algorithm is also applicable to other distributed generation grids, but it needs some samples to verify its performance. Further research is needed to extend the algorithm to hybrid generation models, such as power systems with wind, light, and biomass energy.

Author Contributions: Conceptualization, Q.L. and X.W.; methodology, Q.L.; software, Q.L.; validation, X.W. and S.R.; formal analysis, Q.L.; data curation, X.W.; writing—original draft preparation, Q.L.; writing—review and editing, X.W. and S.R.

Funding: This research was funded by the National Natural Science Foundation of China (No. 61673268) and the Key Project of NSFC (No. 61533012).

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Vita, V.; Alimardan, T.; Ekonomou, L. The Impact of Distributed Generation in the Distribution Networks' Voltage Profile and Energy Losses. In Proceedings of the 2015 IEEE European Modelling Symposium (EMS), Madrid, Spain, 6–8 October 2015; pp. 260–265.
- 2. Lazarou, S.; Vita, V.; Christodoulou, C.; Ekonomou, L. Calculating operational patterns for electric vehicle charging on a real distribution network based on renewables' production. *Energies* **2018**, *11*, 400. [CrossRef]
- 3. Gupta, N.; Pant, V.; Das, B. Probabilistic load flow incorporating generator reactive power limit violations with spline based reconstruction method. *Electr. Power Syst. Res.* **2014**, *106*, 203–213. [CrossRef]
- 4. Borkowska, B. Probabilistic load flow. IEEE Trans. Power App. Syst. 1974, 93, 752–755. [CrossRef]
- 5. Bordbari, M.J.; Seifi, A.R.; Rastegar, M. Probabilistic energy consumption analysis in buildings using point estimate method. *Energy* **2018**, *142*, 716–722. [CrossRef]
- 6. Chen, C.; Wu, W.; Zhang, B.; Sun, H. Correlated probabilistic load flow using a point estimate method with nataf transformation. *Int. J. Electr. Power Energy Syst.* **2015**, *65*, 325–333. [CrossRef]
- Delgado, C.; Domínguez-Navarro, J.A. Point estimate method for probabilistic load flow of an unbalanced power distribution system with correlated wind and solar sources. *Int. J. Electr. Power Energy Syst.* 2014, 61, 267–278. [CrossRef]
- 8. Sakamoto, J.; Mori, Y.; Sekioka, T. Probability analysis method using fast Fourier transform and its application. *Struct. Saf.* **1997**, *19*, 21–36. [CrossRef]
- 9. Pathak, A.K. A simple probabilistic proof for the alternating convolution of the central binomial coefficients. *Am. Stat.* **2018**, *72*, 287–288. [CrossRef]
- 10. Conti, S.; Raiti, S. Probabilistic load flow using Monte Carlo techniques for distribution networks with photovoltaic generators. *Sol. Energy* **2007**, *81*, 1473–1481. [CrossRef]
- Huang, H.; Chung, C.Y.; Chan, K.W.; Chen, H. Quasi-Monte Carlo based probabilistic small signal stability analysis for power systems with plug-in electric vehicle and wind power integration. *IEEE Trans. Power Syst.* 2013, 28, 3335–3343. [CrossRef]
- 12. Valverde, G.; Saric, A.T.; Terzija, V. Probabilistic load flow with non-gaussian correlated random variables using gaussian mixture models. *IET Gener. Transm. Distrib.* **2012**, *6*, 701–709. [CrossRef]
- Oshnoei, A.; Khezri, R.; Tarafdar Hagh, M.; Techato, K.; Muyeen, S.M.; Sadeghian, O. Direct probabilistic load flow in radial distribution systems including wind farms: An approach based on data clustering. *Energies* 2018, 11, 310. [CrossRef]
- 14. Huynh, K.V.; Ngo, D.V.; Le, D.D.; Nguyen, T.N. Probabilistic power flow methodology for large-scale power systems incorporating renewable energy sources. *Energies* **2018**, *11*, 2624. [CrossRef]
- 15. Zou, B.; Xiao, Q. Solving probabilistic optimal power flow problem using quasi Monte Carlo method and ninth-order polynomial normal transformation. *IEEE Trans. Power Syst.* **2014**, *29*, 300–306. [CrossRef]
- Liu, H.; Tang, C.; Han, J.; Li, T.; Li, J.; Zhang, K. Probabilistic load flow analysis of active distribution network adopting improved sequence operation methodology. *IET Gener. Transm. Distrib.* 2017, 11, 2147–2153. [CrossRef]
- 17. Cai, D.; Shi, D.; Chen, J. Probabilistic load flow with correlated input random variables using uniform design sampling. *Int. J. Electr. Power Energy Syst.* **2014**, *63*, 105–112. [CrossRef]
- 18. Shields, M.D.; Zhang, J. The generalization of Latin hypercube sampling. *Reliab. Eng. Syst. Saf.* **2016**, *148*, 96–108. [CrossRef]
- 19. Cai, D.; Shi, D.; Chen, J. Probabilistic load flow computation using copula and latin hypercube sampling. *IET Gener. Transm. Distrib.* **2014**, *8*, 1539–1549. [CrossRef]
- 20. Chen, Y.; Wen, J.; Cheng, S. Probabilistic load flow method based on nataf transformation and Latin hypercube sampling. *IEEE Trans. Sustain. Energy* **2013**, *4*, 294–301. [CrossRef]
- 21. Hajian, M.; Rosehart, W.D.; Zareipour, H. Probabilistic power flow by Monte Carlo simulation with Latin supercube sampling. *IEEE Trans. Power Syst.* **2013**, *28*, 1550–1559. [CrossRef]
- 22. Kabir, M.N.; Mishra, Y.; Bansal, R.C. Probabilistic load flow for distribution systems with uncertain PV generation. *Appl. Energy* **2016**, *163*, 343–351. [CrossRef]

- 23. Groeneboom, P.; Jongbloed, G. *Nonparametric Estimation under Shape Constraints*; Cambridge University Press: Cambridge, UK, 2014; Volume 38.
- 24. Liu, P.-L.; Der Kiureghian, A. Multivariate distribution models with prescribed marginals and covariances. *Probab. Eng. Mech.* **1986**, *1*, 105–112. [CrossRef]
- 25. Morales, J.M.; Baringo, L.; Conejo, A.J.; Minguez, R. Probabilistic power flow with correlated wind sources. *IET Gener. Transm. Distrib.* **2010**, *4*, 641–651. [CrossRef]
- 26. Sadanandan, S.; Radman, G. Optimizing Power System Losses Using Quadrature Boosters on the IEEE 30-Bus Test System. In Proceedings of the SoutheastCon 2016, Norfolk, VA, USA, 30 March–3 April 2016; pp. 1–8.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).