Probabilistic Load Flow: A Review

P. Chen, Student Member, IEEE, Z. Chen, Senior Member, IEEE, and B. Bak-Jensen, Member, IEEE

Abstract—This paper reviews the development of the probabilistic load flow (PLF) techniques. Applications of the PLF techniques in different areas of power system steady-state analysis are also discussed. The purpose of the review is to identify different available PLF techniques and their corresponding suitable applications so that a relatively accurate and efficient PLF algorithm can be determined for the concerned system, e.g. a distribution system with large integration of renewable energy based dispersed generations.

Index Terms—Analytical approach, Monte Carlo, Probabilistic load flow, Stochastic load flow

I. NOMENCLATURE
BSR : Bulk system reliability
CHP : Combined heat and power
CDF : Cumulative distribution function
CRE : Composite reliability evaluation
DG : Dispersed generation
DLF : Deterministic load flow
LF : Load flow
MC : Monte Carlo
PDF : Probabilistic density function
PLF : Probabilistic load flow
SLF : Stochastic load flow
WT: : Wind turbine

II. INTRODUCTION

The DLF is used to analyze and assess the planning and operating of power systems on a daily routine. DLF uses specific values of power generations and load demands of a selected network configuration to calculate system states and power flows. Therefore, DLF ignores uncertainties in the power systems, e.g. the outage rate of generators, the change of network configurations and the variation of load demands. Furthermore, modern power systems with integration of DG units, such as WTs and photovoltaic systems, introduce additional power fluctuations into the system due to their uncontrollable prime sources. Therefore, the deterministic approach is not sufficient for the analysis of modern power systems and the results from DLF may give an unrealistic assessment of the system performance. In order to take the uncertainties into consideration, different mathematical approaches for uncertainty analysis can be used, such as the probabilistic approach, fuzzy sets and interval analysis [1]. The probabilistic approach has a solid mathematical background and has been applied to power systems in different areas [2]. This paper provides a review on the PLF techniques, which are used to analyze the system steady state performance.

The PLF was first proposed in 1974 and has been further developed and applied into power system normal operation, short-term/long-term planning as well as other areas [3][4][5]. The PLF requires inputs with PDF or CDF to obtain system states and power flows in terms of PDF or CDF, so that the system uncertainties can be included and reflected in the outcome. The PLF can be solved numerically, i.e. using a MC method, or analytically, e.g. using a convolution method, or a combination of them [6][7][8]. The main concern about the MC method is the need of large number of simulations, which is very time-consuming; whereas the main concerns about the analytical approach are the complicated mathematical computation and the accuracy due to different approximations. In parallel, a similar technique called SLF has also been developed to deal with the same problems [2]. It is based on the assumption that the probabilistic distributions of the system states and power flow outputs are normal distributions. This assumption, although it simplifies the calculation, is demonstrated to be unreliable by other researches [4]. Therefore, the application of the SLF is very limited and will not be further discussed in this paper.

Reference [9] provides an extensive bibliography on PLF published before 1988. Reference [2] also summarizes main techniques of PLF published before 1987. The main focuses of these literatures are on the linearization of LF equations, network outages and the interdependence among nodal power injections. However, there are also numerous literatures on PLF published from 1989 up to now, regarding issues such as the efficiency of algorithms, power system planning and the inclusion of voltage control devices [10][11][11]. A review on the traditional and newly developed PLF algorithms will provide a clearer indication on the different available techniques and corresponding application areas. A suitable PLF technique can thereafter be selected to cope with the concerned issues associated with modern power systems, such as distribution systems integrated with a large amount of stochastic DG. This paper is organized as follows. First, the basic techniques and related assumptions of the PLF technique are analyzed. Then miscellaneous techniques, developed to improve the accuracy and efficiency of the PLF algorithm, are discussed. Finally, the application and extension of the PLF technique in different areas of power systems are presented.

This work was supported by the Danish Agency for Science Technology and Innovation, under the project of 2104-05-0043.

P. Chen, Z. Chen and B. Bak-Jensen are all with the Institute of Energy Technology, Aalborg University, Aalborg, 9220 Denmark (e-mail: pch@iet.au.dk, zch@iet.au.dk, bbj@iet.au.dk).
III. Basic Techniques of PLF

The PLF can be performed by using either a numerical approach or an analytical approach. The numerical approach, e.g. a MC method, substitutes a chosen number of values for the stochastic variables and parameters of the system model and performs a deterministic analysis for each value so that the same number of values are obtained in the results; whereas the analytical approach analyzes a system and its inputs using mathematical expressions, e.g. PDFs, and obtains results also in terms of mathematical expressions.

A. Numerical Approach

The numerical approach is to adopt a MC method for the PLF analysis. The two main features of MC simulation are random number generation and random sampling. Software such as MATLAB provides algorithms for pseudorandom number generation. Refer to [12] for different techniques of random sampling, e.g. simple random sampling, stratified random sampling, etc. Although sampling techniques can be rather sophisticated, the PLF using MC is in principle doing DLF for a large number of times with inputs of different combinations of nodal power values. Therefore, the exact non-linear form of LF equations as shown in (1)-(5) can be used in the PLF analysis.

\[ P_i = U_i \sum_{k=1}^{n} U_k \left( G_{ik} \cos \theta_{ik} + B_{ik} \sin \theta_{ik} \right) \]  
(1)

\[ Q_i = U_i \sum_{k=1}^{n} U_k \left( G_{ik} \sin \theta_{ik} - B_{ik} \cos \theta_{ik} \right) \]  
(2)

\[ P_{ik} = -t_{ik} G_{ik} U_i^2 + U_i U_k \left( G_{ik} \cos \theta_{ik} + B_{ik} \sin \theta_{ik} \right) \]  
(3)

\[ Q_{ik} = t_{ik} B_{ik} U_i^2 - B_{ik} U_i^2 + U_i U_k \left( G_{ik} \sin \theta_{ik} - B_{ik} \cos \theta_{ik} \right) \]  
(4)

\[ Q_{i(ih)} = U_i^2 B_{i(ih)} \]  
(5)

where \( P_i \) and \( Q_i \) are the net active and reactive power injection at bus \( i \); \( P_{ik} \) and \( Q_{ik} \) are the active and reactive power flows in line \( ik \) at the bus \( i \) side; \( U_i \) and \( U_k \) are the voltage magnitude at bus \( i \) and \( k \); \( \theta_{ik} \) is the angle difference between the voltages at bus \( i \) and \( k \); \( G_{ik} \) and \( B_{ik} \) are the real and imaginary part of the corresponding admittance matrix. The capability to use the exact non-linear LF equations is the reason why results obtained from the PLF using MC are usually taken as a reference to the results obtained from other PLF algorithms with simplified LF equations, so as to check the accuracy of the algorithms [13]. In spite of its relatively high accuracy, the MC method requires large amount of computation time due to the large number of LF calculations.

B. Analytical Approach

The basic idea of the analytical approach is to do arithmetic, i.e. using convolution techniques, with PDFs of stochastic variables of power inputs so that PDFs of stochastic variables of system states and line flows can be obtained. However, the difficulties of solving PLF equations by the convolution of PDFs of input power variables are mainly twofold [2][14]:

a) LF equations (1)-(5) are non-linear
b) input power variables at different buses are usually not completely independent of or linear-correlated

d) total independent or linear-correlated power variables
e) normal distribution and discrete distribution are usually assumed for the load and generation, respectively
f) network configuration and parameter are constant

As a result, the LF equations are linearized around the estimated mean of the system states \( \hat{X} \) with the first-order Taylor expansion [13]. If (1) and (2) are represented by a more general form as:

\[ Y = f(X) \]  
(6)

then the linearized form can be expressed as:

\[ X = \bar{X} + A(X - \bar{X}) \]  
(7)

where

\[ A = \left( \frac{\partial f}{\partial X} \right)_{X=\bar{X}}^{-1} \]  
(8)

A is also referred to as sensitivity coefficient matrix in the PLF formulation. Similar expressions can be derived for (3), (4) and (5). In the DLF solved by using Newton-Raphson method, the Jacobian matrix \( A \) is also computed for each iteration until errors of the results are less than specified values. However, in the PLF here, the Jacobian matrix is only computed once for the computation of each LF. Therefore, errors caused by the linearization of LF equations should be noted and taken care of. Equation (7) shows that the system states are expressed by a linear combination of input power variables. With the assumption of independence, a convolution technique can then be applied to derive the PDFs of system states \( X \), which is:

\[ f(X) = f(Y_1 - \bar{Y}_1) * f(Y_2 - \bar{Y}_2) * \ldots * f(Y_n - \bar{Y}_n) \]  
(9)

Refer to [7][13][15] for detailed convolution techniques of mixed continuous and discrete variables.

IV. Improvement of PLF Techniques

A. Non-linear LF Equations

As the non-linear LF equations are linearized around the expected value region, the accuracy of the results become worse when values of the input power variables are far from their corresponding mean values. The errors are usually reflected in the tail regions of the results, e.g. the two ends of a distribution curve of a bus voltage. This may greatly impact the decision-making judged by adequacy indices such as the probability of a bus voltage outside its operational limits. Therefore, different methods have been proposed to mitigate the error caused by the linearization of the LF equations. Two typical solutions are PLF using multi-linearization [16][8] and the quadratic PLF [17][18].

Multilinearization of the PLF is to linearize the LF
equations around several other points besides the mean value. Around each linearization point, a similar convolution technique can be used to obtain the probabilistic distributions of the results, and these are properly combined to give the final probabilistic representation of the results. The key point is to find the different linearization points that are of interests, e.g. the maximum or minimum of the output stochastic variables. This can be achieved by a so-called boundary LF algorithm, which maximizes (or minimizes) e.g. the linear equation shown in (7). Theoretically, the maximum (or minimum) value of a normal distributed variable $Y_j$ is $\infty$ (or $-\infty$), which cannot be used to obtain the maximum or minimum value of $X$. This can be solved by truncating the distribution of each (normal distributed) input variable so as to obtain a final interval of $\pm 3\sigma_{eq}$, where $\sigma_{eq}$ is the standard deviation of the equivalent normal for a given linearization. Other linearization points between the mean value and the maximum (or minimum) value of the state or output variables can be found in a similar way by specifying the truncating factor $\beta$ on the normal distributed input variables [16]. It is shown in [16] that the resemblance between the results obtained from the multilinearized PLF algorithm and those obtained from the MC method are much improved as compared to the use of the traditional linearization algorithm. Another method to determine linearization points is proposed in [8], which is to use a criterion based on the total active system load. It combines the MC simulation and the multilinearized LF equations to achieve a relatively simple and efficient algorithm for the PLF analysis.

The quadratic PLF is to include the second-order term of the Taylor series expansion of the LF equations. However, the non-linear part in the second-order term, $\left( X - \hat{X} \right)^T \left( X - \hat{X} \right)$, is replaced by the first-order Taylor series expansion as shown in (7). As a result, the quadratic expressions of the LF equations are obtained. It is discussed in [17] that the contribution of including the quadratic terms is generally small but becomes substantial with heavily loaded operating points and large variation of loads.

Another approach to mitigate the error caused by the linearization is proposed in [19], which accounts the non-linearity of the LF equations by partitioning system loads into several segments and performing the linearized PLF algorithm at each load segment. The improvement of the results by using several load segments is also shown with reference to the results obtained from the MC simulation.

B. Network Outage Rates

The change of network configurations inevitably leads to the change of the set of LF equations and relating system outputs. The algorithm of the PLF considering network outage rates are discussed in [20]. The algorithm considers the network configuration as a discrete stochastic variable with specified probability of each network component. The final PDFs or CDFs of system states and line power flows are obtained from a weighted sum of the PDFs or CDFs obtained under each network configuration, respectively. Due to the inclusion of limited number of network configurations, the probability of the considered configurations needs to be modified so as to obtain a probability of 1 for the total considered network configurations. It is also shown in [20] that the consideration of network component outage rates in the PLF analysis is especially important when the load uncertainties are small, e.g. in the power system operational planning. When the load uncertainties are large and dominate the system uncertainties, the effect of network component outage rates are not significant.

Another method of dealing with network outage rates is proposed in [21]. The line outage is simulated by modifying the injected powers at both ends of the line so that the total power leaving the line is the same as the case of the actual line outage. However, the system states need to be modified according to a defined threshold. There are mainly two errors introduced in the solution by this method. One error is due to the accuracy of the estimated values for the power injection changes at both ends of the line. The other one is due to the value of the chosen threshold. Although the complete LF equations instead of linearized ones are carried out when the change of the system states is larger than the chosen threshold, the computation time is increased.

Another issue related to the network random phenomenon is the variation of network parameters due to e.g. the variation of temperature. The parameter variation should be considered as a continuous stochastic variable. This is not treated in the conventional PLF. A point estimate method is proposed in [10] to include the parameter variation in the PLF analysis. Distributions of the resistance and reactance of the series line impedance are assumed uniform distributions with different mean values. The distribution of the susceptance of a line is assumed a binary distribution. The two-point estimate method uses two weighted factor $w_{l,1}$ and $w_{l,2}$ (instead of a PDF of a stochastic variable) and corresponding two concentration values $p_{l,1}$ and $p_{l,2}$ of the stochastic variable (two values expressed by a linear combination of the mean value and standard deviation of the variable) in the non-linear LF equations to calculate the statistical moments of the output line flow variables. This approximation of the method is to consider independently the contribution of the random effect of each input stochastic variables to the output line flow variables. Each input stochastic variable is further approximated by two concentration values. The method can be applied to discrete stochastic variables in general. Different degrees of variations of the line parameter are simulated and the results are compared with those obtained from the corresponding MC simulations. It is also shown that the error of the two-point estimate method slowly increases as the uncertainty level of line parameters increases. However, the accuracy of the method is more sensitive to the bus power variation than the line parameter variation. It is also shown that the results obtained from the two-point estimate method is more accurate and faster in execution time than the multilinear simulation algorithm proposed in [8].

C. Interdependence of Stochastic Variables

Usually, the input power variables are not completely independent of or linear-correlated with each other. This is
due to the reason that a certain type of loads such as residential loads has similar behaviors. Furthermore, central generations and loads are also correlated in a way due to the operator action or economic dispatch. DG such as CHP units and WTs is also correlated with loads due to the weather. However, it is difficult to model the interdependence of multivariable problems if only marginal probability distributions of stochastic variables are known but not the joint probability distributions. If synchronized statistical data of input power variables of the PLF are available, the interdependence can be readily modeled by sampling different power variables at the same hour. However, if this is not the case, different interdependence relations have to be considered individually.

The dispatching law can be approximated by a linear equation which includes a component \( R \) that represents the change of generation due to the change of load and another component \( \Delta P_c \) that represents the redistribution of the power generation due to generation outages [2]. This method can be easily implemented if \( R \) is known. Another method is also discussed in [2], which assumes that the balancing of the power is only a function of the sum of the power inputs and the outputs, instead of the balancing at each individual bus. The latter method is more practical than the former one due to the reason that \( R \) is usually difficult to estimate.

For the modelling of the correlation between the active and reactive power of loads as well as loads connected at different buses, a linear dependence was assumed in [22]. However, it is not appropriate to assume linear dependence when generation is involved as this ignores the generator outage rates. The generation is then divided into two types, the independent generation system and the dependent generation system [22]. The former one corresponds to base loads and is independent of load variations. Therefore, it can be modeled by independent binomial variables. The latter one corresponds to intermittent and peaking loads. Therefore a dispatching function based on the economic operating criteria is used to model the interdependence between the generations and the loads. Another method of modelling the short-term load demand was suggested in [23] to consider a partial correlation instead of the totally correlated or totally independent assumption. The partial correlation is to consider that load demands have mean values rising and falling together in step with a small independent random variation about the mean. The small random variation can be described by a normal distribution.

The foregoing generations mentioned are dispatchable. The inclusion of non-dispatchable generations, e.g. DG, in the modelling of interdependence is treated in [24]. The interdependence between the load demands and the non-dispatchable generations are modeled through two levels, i.e. time of day or season and weather. These two levels are to account for the interdependence due to the cyclic phenomenon (time of day, day of the week, season) and the random phenomenon (temperature, cloud cover, wind speed) related to the load demands and the non-dispatchable generations. However, this is of course based on the availability of the statistical data of the input power variables. The modelling of the interdependence structure of DG, e.g. power generations from wind farms, is of great importance due to the reason that generations from WTs are strongly correlated among adjacent wind farms due to the similar wind speed at the area. The significance of interdependence in modeling the stochastic generation is also demonstrated in [25]. Theoretical expressions need to be further developed in order to account for the interdependence among stochastic generations and loads so that it can be included in the analytical PLF algorithm.

### D. Efficiency of PLF algorithm

The way to perform an efficient convolution using Laplace transformation of discrete stochastic variables and continuous stochastic variables is shown in [7]. Another method using FFT to perform convolution is proposed in [15], which shows a better efficiency. However, convolution of a variable with \( r \) impulses and a variable with \( s \) impulses will always have \( r \times s \) impulses, which inevitably require a large storage and computation time.

As discussed before, the multilinear simulation algorithm is faster than a pure MC simulation [8]. The point estimation method is also claimed to be faster than the multilinear simulation algorithm [10]. Another method to improve the efficiency of a PLF algorithm by combining cumulants and Gram-Charlier expansion theory is proposed in [26]. The main idea of using Gram-Charlier series is that it can express the CDF and PDF of a standardized variable of any type by a sum of corresponding weighted standard normal distributions at different orders of derivatives. The weights can be calculated through different orders of cumulants. In addition, the property of cumulants is superior to that of the moments in the sense that the cumulants at any order of a stochastic variable \( Z = B+C \) (\( B \) and \( C \) are independent stochastic variables) is the sum of the cumulants of \( B \) and \( C \) at the corresponding order; whereas this is not valid for the moments higher than the second order. In this way, CDFs can also be obtained directly without by integrating PDFs. It is claimed in [26] that the proposed method is 20-30 times faster than the MC simulation.

### V. Application and Extension of PLF

#### A. Power System Planning

For power system operational planning [23][27][28] and expansion planning [5][27], it is important to obtain accurate values of adequacy indices for a composite generation and transmission system or so-called bulk system. There are basically two probabilistic techniques to assess BSR indices: PLF and CRE [5]. These two techniques are very similar to each other but the difference of them is well explained in [5]. The basic difference of the PLF and the CRE is that CRE also provides adequacy indices (BSRr) after resolving the network problems through remedial actions; whereas the PLF only calculates the adequacy indices (BSRp) based on the current system condition before any remedial actions are taken. However, both techniques can be extended to include each other. The BSRr indices include probability of system
problems, loss of load probability, expected energy not supplied, etc. The BSRp indices include probability of power flow being greater than the corresponding equipment or line thermal rating, probability of a bus voltage being outside the required limits, etc.

According to [27], the load modelling should be divided into short term load modelling, which accounts for the uncertainties of environmental and social factors in the operational planning, and long-term load modelling, which accounts for the uncertainties of demographic and economic factors in the long term planning. The short term load modelling collects at a substation the daily peak load values for about two months. The long term load modelling collects at a substation observed annual peak load values for a certain number of years. Then a PLF is performed to obtain the system states using the linearized LF equations. The line flows are obtained from the system states using the original nonlinear LF equations. In addition to the foregoing adequacy indices, the simulation results from the PLF provide more insights than that from a conventional deterministic study, e.g. an alternative of increasing the reactive generation support instead of construction of a power line is too risky to be adopted [5].

In addition, reference [28] discusses the short term network planning of a distribution system to take into account the stochastic behavior of DG units. The simulation results show the capability of the statistical planning method to increase the network transfer capability as compared to the traditional worst case planning principle. Reference [29] discusses 5 different non-deterministic approaches for the transmission expansion planning, including PLF, probabilistic reliability criteria, scenario techniques, decision analysis and fuzzy decision making. An extensive bibliography for each approach is also included.

B. Systems with voltage control devices

Power systems usually contain a number of voltage control devices, such as tap-changing transformers, switched capacitor banks, and static var compensator. The dimensioning and setting of these control devices for a better voltage control in the network should be determined in a more realistic way than the traditional deterministic approach. The issue of the voltage control with the system control devices through the PLF analysis has been investigated [30]-[35]. The basic idea is to include a control variable, such as transformer taps, shunt compensation devices and voltages at PV buses, in a constrained LF analysis, so that some or all the elements of system states and line flow outputs are within operating limits.

In addition, operating constraint violations are obtained together with the probability of each violation. The procedure can be considered as an extension of the PLF in the system analysis after the remedial actions are taken as discussed in the previous section. The algorithm of the probabilistic constrained LF are discussed in detail in [30][31].

C. Systems Integration with DG

Due to the stochastic behavior of the prime sources of some DG, such as wind speed and temperature, the steady state analysis of the systems with integration of such DG units requires a probabilistic approach. The complexity of including such DG units in the PLF analysis is mainly two fold. The first one is due to the continuous and usually large variation of the power generation from DG units. This may result in a large error if linearized LF equations around the mean values are used. Instead, the multilinear or the quadratic PLF may be needed. In addition, the stochastic wind power generation is usually not a normal distribution but, e.g. a Weibull distribution. The second one is due to the interdependence of the power generation from different DG units, e.g. correlated wind power generation from two adjacent wind farms as illustrated before, as well as the interdependence of the power generation and load demand, e.g. the power generation from CHP units usually reflects the human work and rest behavior and may strongly correlated with the residential load demand. A detailed analysis of PLF in systems integrated with DG can be found in [6], [36]-[39].

D. Other Aspects

The PLF analysis has also been extended to include the three-phase unbalance [40] and applied to the harmonic LF [41]. Assessment of the voltage instability using the PLF is presented in [42]. The formulation of the probabilistic optimal power flow and its application to the electricity markets are also demonstrated in [43] and [44], respectively.

VI. CONCLUSION

This paper has first of all discussed the necessity of using the probabilistic approach for the analysis of the power system steady state performance followed by a brief history of the PLF technique. Secondly, the basic PLF technique with its assumptions has been demonstrated. Thirdly, different approaches have been discussed to improve the accuracy and efficiency of the PLF algorithm. Finally, the application of the PLF algorithm in the power system planning and the extension of the PLF algorithm to include voltage control devices and systems integrated with DG have been presented.

VII. REFERENCES


