Point Estimate Schemes to Solve the Probabilistic Power Flow

Juan. M. Morales, *Student Member, IEEE*, and Juan. Pérez-Ruiz, *Member, IEEE*

***Abstract*—This paper analyzes the behavior of Hong's point estimate methods to account for uncertainties on the probabilistic power flow problem. This uncertainty may arise from different sources as load demand or generation unit outages. Point estimate methods constitute a remarkable tool to handle stochastic power system problems because good results can be achieved by using the same routines as those corresponding to deterministic problems, while keeping low the computational burden. In previous works related to power systems, only the two-point estimate method has been considered. In this paper, four different Hong's point estimate schemes are presented and tested on the probabilistic power flow problem. Binomial and normal distributions are used to model input random variables. Results for two different case studies, based on the IEEE 14-bus and IEEE 118-bus test systems, respectively, are presented and compared against those obtained from the Monte Carlo simulation. Particularly, this paper shows that the use of the $2m + 1$ scheme provides the best performance when a high number of random variables, both continuous and discrete, are considered.**

***Index Terms*—Monte Carlo simulation, point estimate methods, power flow analysis, power system modeling, uncertainty.**

I. Introduction

MANY engineering problems are subject to uncertainty, due to the inherent randomness of natural phenomena or to the implicit and inaccurate assumptions related to the considered modeling approach. Computational methods which tackle uncertainty allow engineers to propose solutions less sensitive to environmental influences, while achieving simultaneously cost reduction or reliability improvement. In power system analysis, the use of techniques able to account for uncertainty is required to control and minimize the risks associated with design and operation [1].

Power flow studies [2], [3] help engineers to accomplish these tasks by revealing the steady state of the system under consideration for a specified set of power generations, loads, and network conditions. In an open-access environment, this information is not certain, and for this reason, it is desirable to assess system variables (bus voltages and line flows) for a range of load and generation conditions. Using deterministic power flow routines, it is necessary to run them many times so as to encompass all, or at least the majority of, possible system states. Hence, from a practical point of view, it turns out to be convenient to approach the power flow as a probabilistic problem. This leads to the problem known as probabilistic power flow [4], [5], where the state variables and output network quantities of a power system are obtained as random variables, and thus, it becomes easy to identify the possible ranges of the power flow results.

In the technical literature, there are several techniques to deal with problems under uncertainty. These techniques may be classified into the three main categories [6]: Monte Carlo simulation, analytical methods, and approximate methods.

Monte Carlo simulation [7] randomly generates values for uncertain input variables, and these values are taken into account to solve a deterministic problem. This technique has been widely used in power systems analysis to model uncertainty. The main drawback of the Monte Carlo method is the great number of simulations required to attain convergence. However, it uses deterministic routines to solve the problem in each simulation.

Analytical methods are computationally more effective, but they require some mathematical assumptions in order to simplify the problem. A multilinear model is used to handle network nonlinearities [8], [9]. Likewise, convolution techniques are used to obtain a mathematical description of the behavior of output random variables. The fast Fourier transform is employed in [10], and the cumulant method is used to solve the probabilistic power flow in [11]. The later is usually combined with the Gram–Charlier expansion in order to estimate the probability functions of output random variables [12]. Von Mises functions are introduced in [13] to handle discrete distributions. The fuzzy load flow presented in [14] and methods which combine analytical techniques and Monte Carlo simulation [15], [16] may be pointed out as well.

Approximate methods provide an approximate description of the statistical properties of output random variables. Within these techniques, first-order second-moment method (FOSMM) [17] and point estimate methods stand out.

In this paper, the point estimate method approach is used to solve the probabilistic power flow problem. The main advantages follow.

1) As Monte Carlo simulation, point estimate methods use deterministic routines for solving probabilistic problems; however, they require a much lower computational burden.
2) Furthermore, point estimate methods overcome the difficulties associated with the lack of perfect knowledge of the probability functions of stochastic variables, since these functions are approximated using only their first few statistical moments (i.e., mean, variance, skewness, and kurtosis). Therefore, a smaller level of data information is needed.

Manuscript received February 6, 2007; revised April 12, 2007. Paper no. TPWRS-00078-2007.

J. M. Morales is with the Departamento de Ingeniería Eléctrica, Electrónica, Automática y Comunicaciones, ETS Ingenieros Industriales, Universidad de Castilla-La Mancha, Ciudad Real E-13071, Spain (e-mail: JuanMiguel.Morales@uclm.es).

J. Pérez-Ruiz is with the Departamento de Ingeniería Eléctrica, ETS Ingenieros Industriales, Universidad de Málaga, Málaga E-29013, Spain (e-mail: jperez@uma.es).

Digital Object Identifier 10.1109/TPWRS.2007.907515

TABLE I
QUALITATIVE DESCRIPTION OF POINT ESTIMATE METHODS

Method's author [Ref.]	Number of simulations	Efficiency in Large Scale problems	Ability to handle correlated variables	Ability to handle asymmetric variables
Rosenblueth [19]	2^m	Very Low	Yes	Yes
Li [20]	$O(m^3)$	Low	Yes	Yes
Harr [24]	$2m$	High	Yes	No
Hong [25]	Km or $Km+1$	High	No	Yes

The aim of any point estimate method is to compute the moments of a random variable z that is a function of m random input variables p_l, i.e., $z = F(p_1, p_2, \ldots, p_m)$. The first point estimate method was developed by Rosenblueth in 1975 [18] for symmetric variables and was later revisited in 1981 [19] to consider asymmetric variables. Since then, several methods that improve the original Rosenblueth's method have been presented. They basically differ on the type of random variables they consider (symmetric or asymmetric, correlated or not) and on the number of evaluations to be performed. Table I presents a qualitative description of the main point estimate methods (K is a parameter depending on Hong's method used).

In realistic size power system problems, the number of input random variables involved is high. Hence, the Rosenblueth's original method, as well as recent and more accurate point estimate methods based on the Rosenblueth's approach [21]–[23], are not suitable because the number of simulations could be even greater than in the Monte Carlo simulation.

On the other hand, the number of simulations to be performed by using the point estimate methods developed by Harr [24] or Hong [25] grows linearly with the number of input random variables. However, although Harr's method is suitable for correlated variables, it is constrained to symmetric variables (skewness equals zero).

In this paper, Hong's point estimate methods are used to solve the probabilistic power flow problem. Four different concentration schemes are presented and tested over two different case studies of different size. Results are compared against those obtained using Monte Carlo simulation. These case studies take into account both continuous variables (with normal distribution) and discrete variables (with binomial distribution).

To the best of our knowledge, there is only one work [26], later refined in [27], which applies a point estimate method to solve the probabilistic power flow problem. The two-point estimate method used in [27] is equivalent to the Hong's $2m$ scheme. Besides this, a recent paper on probabilistic optimal power flow [28] also uses the two-point estimate method. The two-point estimate method, or $2m$ scheme, does not provide generally good results if the number of input random variables is high, and thus, it is not adequate for power system problems of realistic size. As it will be shown below, the $2m+1$ scheme is able to overcome this drawback by performing only one additional evaluation of the function F.

This paper is organized as follows. Hong's point-estimate methods are introduced in Section II, and two of the four different schemes considered in this paper (the $2m$ and $2m+1$ schemes) are developed, while the other two schemes (the $3m$ and $4m+1$ schemes) are included in the Appendix. Section III presents the application of these point estimate methods to solve the probabilistic power flow problem. Results for two case studies, based on the IEEE 14-bus and IEEE 118-bus test systems, respectively, are presented and compared against those obtained from the Monte Carlo simulation. Finally, in Section V, some relevant conclusions are summarized.

II. HONG'S POINT ESTIMATE METHODS

Point estimate methods concentrate the statistical information provided by the first few central moments of a problem input random variable on K points for each variable, named *concentrations*. By using these points and the function F, which relates input and output variables, information about the uncertainty associated with problem output random variables can be obtained.

The kth concentration $(p_{l,k}, w_{l,k})$ of a random variable p_l can be defined as a pair composed of a location $p_{l,k}$ and a weight $w_{l,k}$. The location $p_{l,k}$ is the kth value of variable p_l at which the function F is evaluated. The weight $w_{l,k}$ is a weighting factor which accounts for the relative importance of this evaluation in the output random variables.

By using Hong's point estimate methods, the function F has to be evaluated only K times for each input random variable p_l at the K points made up of the kth location $p_{l,k}$ of the input random variable p_l and the mean (μ) of the $m-1$ remaining input variables, i.e., at the K points $(\mu_{p_1}, \mu_{p_2}, \ldots, p_{l,k}, \ldots, \mu_{p_m})$. In other words, the deterministic problem has to be solved K times for each input random variable p_l, and the difference among these problems is the deterministic value $p_{l,k}$ assigned to p_l, while the remaining input random variables are fixed to their corresponding mean. The number K of evaluations to carry out depends on the *scheme* used. Therefore, the total number of evaluations of F is $K \times m$.

Specific variants, or schemes, of Hong's point estimate method take into account one more evaluation of function F at the point made up of the m input random variables means $(\mu_{p_1}, \mu_{p_2}, \ldots, \mu_{p_l}, \ldots, \mu_{p_m})$. Hence, for these schemes, the total number of evaluations of F is $K \times m + 1$.

Hong's point estimate method is described with the aid of Fig. 1. The K concentrations $(p_{l,k}, w_{l,k})$ of the m input random variables p_l are obtained from the statistical input data (e.g., the probability density function f_{p_l} in Fig. 1).

The location $p_{l,k}$ to be determined is

$$p_{l,k} = \mu_{p_l} + \xi_{l,k}\sigma_{p_l} \tag{1}$$

where $\xi_{l,k}$ is the standard location, and μ_{p_l} and σ_{p_l} (input data) are the mean and standard deviation of the input random variable p_l.

The standard location $\xi_{l,k}$ and the weight $w_{l,k}$ are obtained by solving the nonlinear system of equations [25]

$$\left.\begin{aligned} \sum_{k=1}^{K} w_{l,k} &= \tfrac{1}{m} \\ \sum_{k=1}^{K} w_{l,k}\,(\xi_{l,k})^j &= \lambda_{l,j} \quad j = 1, \ldots, 2K-1 \end{aligned}\right\} \tag{2}$$

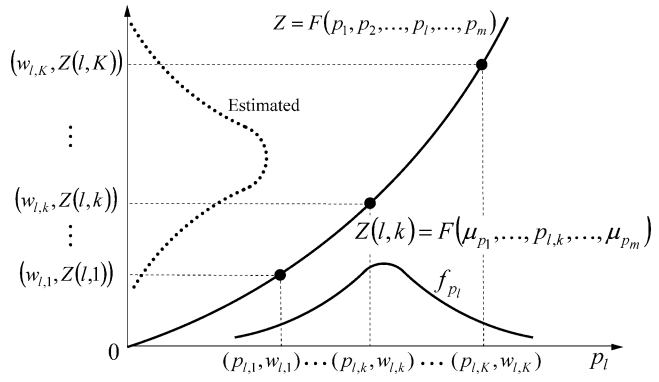

Fig. 1. Hong's point estimate methods.

where $\lambda_{l,j}$ denote the jth standard central moment of the random variable p_l with probability density function f_{p_l}, that is

$$\lambda_{l,j} = \frac{M_j(p_l)}{(\sigma_{p_l})^j} \tag{3}$$

$$M_j(p_l) = \int_{-\infty}^{\infty} (p_l - \mu_{p_l})^j f_{p_l} dp_l. \tag{4}$$

Note that $\lambda_{l,1}$ equals zero, $\lambda_{l,2}$ equals one, and $\lambda_{l,3}$ and $\lambda_{l,4}$ are, respectively, the skewness and kurtosis of p_l. The system of (2) can be efficiently solved by the procedure developed by Miller and Rice [29].

Once all the concentrations $(p_{l,k}, w_{l,k})$ are obtained, the function F is evaluated at the points $(\mu_{p_1}, \mu_{p_2}, \ldots, p_{l,k}, \ldots, \mu_{p_m})$ yielding $Z(l,k)$, where Z is the vector of output random variables. Finally, by using the weighting factors $w_{l,k}$ and the $Z(l,k)$ values, the jth raw moment of the output random variables can be estimated according to the expression (5)

$$\mu'_j = E\left[Z^j\right] \cong \sum_{l=1}^{m} \sum_{k=1}^{K} w_{l,k} (Z(l,k))^j. \tag{5}$$

Detailed information about the estimation error involved in (5) can be found in [30].

In this paper, four different concentration schemes are considered: two $K \times m$ type schemes ($2m$ and $3m$) and two $K \times m + 1$ type schemes ($2m + 1$ and $4m + 1$). The $3m + 1$ scheme is not considered because normally distributed input random variables yield non-real (complex) standard locations, and hence, the concentration parameters $(p_{l,k}, w_{l,k})$ are also useless non-real values.

Solving analytically (2) to obtain the standard locations and weights is only possible for the $2m$ and $2m+1$ schemes. These two schemes are described below. On the other hand, the procedure to obtain the standard locations and weights of the other two schemes ($3m$ and $4m + 1$), which requires a numerical non-analytical solution, is described in the Appendix.

A. $2m$ Scheme ($K = 2$)

This scheme only uses two concentrations for each input random variable. From (2) for $K = 2$, the statistical information supplied by the skewness $\lambda_{l,3}$ is enough to achieve an analytical solution for the standard locations and weights of random variable p_l

$$\xi_{l,1} = \frac{\lambda_{l,3}}{2} + \sqrt{m + \left(\frac{\lambda_{l,3}}{2}\right)^2} \quad \xi_{l,2} = \frac{\lambda_{l,3}}{2} - \sqrt{m + \left(\frac{\lambda_{l,3}}{2}\right)^2} \tag{6}$$

and

$$w_{l,1} = -\frac{1}{m} \frac{\xi_{l,2}}{\xi_{l,1} - \xi_{l,2}} \quad w_{l,2} = \frac{1}{m} \frac{\xi_{l,1}}{\xi_{l,1} - \xi_{l,2}} \tag{7}$$

and then, taking into account the mean and standard deviation of p_l, the locations $p_{l,1}$ and $p_{l,2}$ can be computed from (1).

It should be noted in (6) that the standard locations, $\xi_{l,1}$ and $\xi_{l,2}$, depend on the number m of input random variables. As m increases, the locations $p_{l,1}$ and $p_{l,2}$ move away from the mean μ_{p_l} according to $\sqrt{m}$. Hence, the locations may be at points where the probability distribution of p_l is not well known, or even fall outside its domain of definition. This drawback, shared by all $K \times m$ schemes, has been pointed out in [31].

On the other hand, the $2m$ scheme has significant advantages related to its simplicity, its low computational burden, and the fact that it always provides real value solutions for the concentrations.

B. $2m + 1$ Scheme ($K = 3$, $\xi_{l,3} = 0$)

This is the first $K \times m + 1$ type scheme considered in this paper. It requires only one additional evaluation of function F than the $2m$ scheme. From a mathematical point of view, this scheme arises from solving (2) for $K = 3$ (see the $3m$ scheme in the Appendix) with one of the three standard locations $\xi_{l,k}$ set to zero.

Let $\xi_{l,3} = 0$. Then, the standard locations and weights are:

$$\xi_{l,k} = \frac{\lambda_{l,3}}{2} + (-1)^{3-k} \sqrt{\lambda_{l,4} - \frac{3}{4}\lambda_{l,3}^2} \quad k = 1,2 \quad \xi_{l,3} = 0 \tag{8}$$

$$w_{l,k} = \frac{(-1)^{3-k}}{\xi_{l,k}(\xi_{l,1} - \xi_{l,2})} \quad k = 1,2$$

$$w_{l,3} = \frac{1}{m} - \frac{1}{\lambda_{l,4} - \lambda_{l,3}^2}. \tag{9}$$

It should be noted in (1) that setting $\xi_{l,3} = 0$ yields $p_{l,k} = \mu_{p_l}$ and so, m of the $3m$ locations are the same point $(\mu_{p_1}, \mu_{p_2}, \ldots, \mu_{p_l}, \ldots, \mu_{p_m})$. Hence, it is enough to run only one evaluation of function F at this location, provided that the corresponding weight is updated to the value w_0

$$w_0 = \sum_{l=1}^{m} w_{l,3} = 1 - \sum_{l=1}^{m} \frac{1}{\lambda_{l,4} - \lambda_{l,3}^2}. \tag{10}$$

From (8), it can be seen that the standard location values of the $2m + 1$ scheme do not depend on the number m of input random variables, as do the $K \times m$ type schemes. Moreover, this is a common feature of all the $K \times m + 1$ concentration schemes.

The $2m+1$ scheme is more accurate than the $2m$ scheme because it takes into account the kurtosis $\lambda_{l,4}$ of the input random variables while only one additional evaluation of the function F is needed.

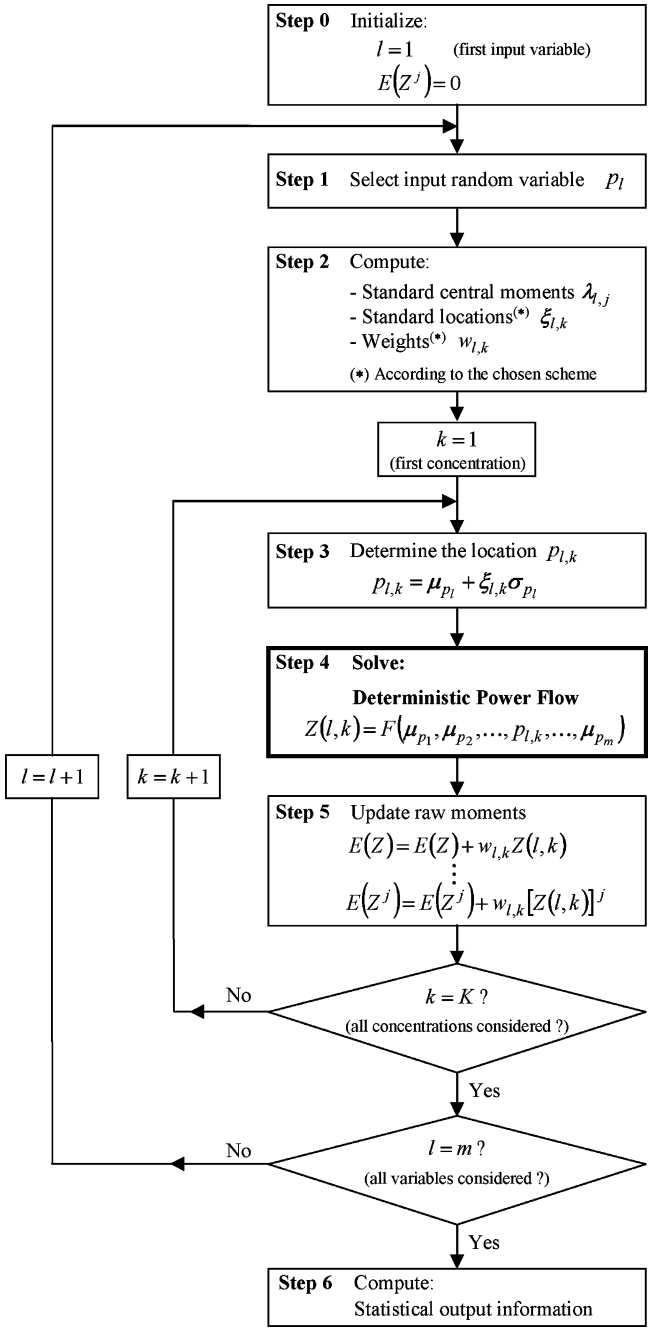

Fig. 2. Flow diagram of the algorithm.

On the other hand, (8) indicate that this scheme yields nonreal locations when $\lambda_{l,4}-(3/4)\lambda_{l,3}^2<0$. However, for the probability distributions usually employed to model the uncertainty in power system problems (normal, binomial, or uniform), the locations are always real values.

III. Application of Hong's Point Estimate Methods to the Probabilistic Power Flow

The algorithm to solve the probabilistic power flow problem by means of Hong's point estimate methods is shown in Fig. 2.

The power flow input data, variables and/or parameters are modeled as random variables with a known probability distribution and, depending on the concentration scheme used, the locations and weights have to be computed as described previously. It should be noted that the point estimate method allows modeling as random variables both input variables of the power flow problem (e.g., power injections) and input parameters (e.g., line reactance). Therefore, in the probabilistic power flow problem, the number of input random variables may be greater than the number of output random variables.

A deterministic power flow must be run for each point $(\mu_{p_1},\mu_{p_2},\ldots,p_{l,k},\ldots,\mu_{p_m})$. Note that an ac deterministic routine may be used to carry out the computations because only deterministic values are involved. The solution of a power flow problem is

$$Z(l,k)=F\left(\mu_{p_1},\mu_{p_2},\ldots,p_{l,k},\ldots,\mu_{p_m}\right) \tag{11}$$

where $Z(l,k)$ is the vector of output random variables associated to the kth concentration of random variable p_l, and F represents the nonlinear relation between input and output variables in the power flow problem. The total number of deterministic power flow problems to be run depends on the concentration scheme considered.

The vector $Z(l,k)$ is used to estimate the raw moments of the output random variables

$$\begin{aligned} E\left(Z\right) &\cong E\left(Z\right)+w_{l,k}Z\left(l,k\right) \\ E\left(Z^j\right) &\cong E\left(Z^j\right)+w_{l,k}\left(Z\left(l,k\right)\right)^j. \end{aligned} \tag{12}$$

The algorithm ends once all the concentrations of all input random variables are taken into account. Then, the estimated raw moments of the output random variables are used to compute the desired statistical information. In this paper, besides the mean and the standard deviation, the Gram–Charlier expansion [12], [32] is used to compute the probability density functions (pdf) and the cumulative density functions (cdf) of the output random variables.

In order to provide a general overview of the overall performance of the different point estimate schemes, the following error indices are defined for each output random variable x:

$$\varepsilon_{\mu}^{x}=\left|\frac{\mu_{MC}^{x}-\mu_{PEM}^{x}}{\mu_{MC}^{x}}\right|\times 100\ \ [\%] \tag{13}$$

$$\varepsilon_{\sigma}^{x}=\left|\frac{\sigma_{MC}^{x}-\sigma_{PEM}^{x}}{\sigma_{MC}^{x}}\right|\times 100\ \ [\%] \tag{14}$$

where μ_{MC}^{x} and σ_{MC}^{x} are the mean and standard deviation calculated with Monte Carlo simulation and, hence, taken as reference values. Similarly, μ_{PEM} and σ_{PEM} are the mean and standard deviation estimated by a given point estimate method. Likewise, the average error indices $\bar{\varepsilon}_{\mu}^{X}$ and $\bar{\varepsilon}_{\sigma}^{X}$ are defined as the mean values of ε_{μ}^{x} and ε_{σ}^{x}, respectively, for each set X of variables, where X may refer to voltages (V), angles (δ), active power injections (P_i), reactive power injections (Q_i), active power line flow (P_{ij}), or reactive power line flow (Q_{ij}).

IV. Case Studies

Two case studies are presented to illustrate the overall performance of the point estimate methods presented in this paper. The probabilistic power flow problem is solved for the IEEE

TABLE II
MEAN AND STANDARD DEVIATION RESULTS FOR THE IEEE 118-BUS TEST SYSTEM (SELECTED VALUES)

IEEE 118 μ and σ results		V_{96} [p.u.]	δ_{99} [degree]	P_{69} [p.u.]	Q_{27} [p.u.]	$P_{89,90}$ [p.u.]	$Q_{8,30}$ [p.u.]
2m	μ	0.9957	26.5593	5.1422	0.0262	1.1288	0.2488
	σ	0.0724	11.3285	1.8481	2.1875	0.3532	0.9626
2m+1	μ	0.9917	26.9302	5.2072	-0.0893	1.1084	0.2967
	σ	0.0015	4.8097	1.9308	0.0214	0.0638	0.0475
4m+1	μ	0.9917	26.9301	5.2072	-0.0893	1.1084	0.2967
	σ	0.0015	4.8110	1.9310	0.0214	0.0638	0.0478
Monte Carlo	μ	0.9917	26.9176	5.2086	-0.0890	1.1082	0.2967
	σ	0.0016	4.8143	1.9307	0.0214	0.0642	0.0494

14-bus and IEEE 118-bus test systems [33]. The deterministic data available in [33] are considered the *base case*, that is, the location point associated with the mean of the m input random variables.

The probabilistic data for the IEEE 14-bus test system are obtained from [11] and [34]. The probabilistic data for the IEEE 118-bus test system are determined as follows.

- Generation units. Each generation plant is divided into four units with the same power production and a forced outage rate of 0.09. A binomial distribution is used to model each generation plant, and the mean of this input random variable is set to the base case power production of the corresponding generation plant. This condition is used to compute the power production of the units.
- Load demand. The active and reactive power of the load buses are modeled as normal distributions, whose means equal the base case data, and whose standard deviations are set arbitrarily as follows: 7% from bus #1 to bus #33, 4% from bus #34 to bus #59, 9% from bus #60 to bus #79, and 5% from bus #80 to bus #118.

The accuracy and efficiency of the analyzed point estimate schemes are tested by comparing the results obtained with those derived from the Monte Carlo simulation with 10 000 samples. This number of simulations is high enough to guarantee the convergence of Monte Carlo method: the differences among the solutions obtained by running several executions of the Monte Carlo simulation with 10 000 samples are on the order of 10^{-4} per unit. For the purposes of this paper, these differences are small enough to consider the results provided by the Monte Carlo simulation with 10 000 samples as reference, or exact values.

The first result to stand out is the inappropriate behavior of the $3m$ scheme with binomial distributions. In both case studies, the injected active power in generation plants, modeled as binomial distributions, yields non-real (complex) concentration values. This complex concentrations are useless, and thus, the results over the $3m$ scheme have been omitted.

Table II shows the mean and standard deviation results of the 118-bus case study for a selected set of output random variables, which have been considered representative of the general results obtained from the different point estimate schemes analyzed in this paper. It can be observed that the two $K \times m + 1$ type schemes provide good results compared with the Monte Carlo values, both for the mean and the standard deviation. The $2m$ scheme mean results also yield acceptable values, except for Q_{27}; however, the differences grow up for the standard deviation values, where only the variable P_{69} presents a value similar to the Monte Carlo simulation result.

TABLE III
AVERAGE ERROR INDICES FOR IEEE 14-BUS TEST SYSTEM

IEEE 14		V	δ	P_i	Q_i	$P_{i,j}$	$Q_{i,j}$
2m	$\bar{\varepsilon}_{\mu}^{*}$	0.0018	0.0953	0.1664	0.5179	0.0806	0.3272
	$\bar{\varepsilon}_{\sigma}^{*}$	0.5120	0.5101	1.5090	0.8331	0.5361	0.6987
2m+1	$\bar{\varepsilon}_{\mu}^{*}$	0.0018	0.0954	0.1665	0.5192	0.0806	0.3281
	$\bar{\varepsilon}_{\sigma}^{*}$	0.2553	0.5933	1.5506	0.8273	0.5304	0.5312
4m+1	$\bar{\varepsilon}_{\mu}^{*}$	0.0018	0.0954	0.1665	0.5192	0.0806	0.3281
	$\bar{\varepsilon}_{\sigma}^{*}$	0.2552	0.5933	1.5506	0.8273	0.5304	0.5312

TABLE IV
AVERAGE ERROR INDICES FOR IEEE 118-BUS TEST SYSTEM

IEEE 118		V	δ	P_i	Q_i	$P_{i,j}$	$Q_{i,j}$
2m	$\bar{\varepsilon}_{\mu}^{*}$	0.1354	1.4565	0.9732	39.1229	18.6734	49.5651
	$\bar{\varepsilon}_{\sigma}^{*}$	4695.2	81.1050	3.4946	6174.6	451.02	3662.7
2m+1	$\bar{\varepsilon}_{\mu}^{*}$	0.0006	0.1483	0.2772	0.1408	0.9021	0.2123
	$\bar{\varepsilon}_{\sigma}^{*}$	3.6962	0.5617	0.8241	4.4711	0.7094	2.1324
4m+1	$\bar{\varepsilon}_{\mu}^{*}$	0.0006	0.1484	0.2775	0.1410	0.9019	0.2123
	$\bar{\varepsilon}_{\sigma}^{*}$	3.3109	0.5671	0.8364	4.1138	0.7101	2.0315

By comparing the results of the two case studies of different size, the effect of the number of input random variables m on the behavior of the different point estimate schemes analyzed can be pointed out. In this sense, note that the number of input random variables included in the 14-bus case study is 23, while the probabilistic power flow run over the 118-bus case study takes into account 170 input random variables.

Tables III and IV show the average error indices corresponding to the estimations of the different point estimate methods developed in this paper for the IEEE 14-bus and IEEE 118-bus test systems, respectively. The average error indices of the $2m$ scheme for the IEEE 14-bus test system (see Table III) are as good as those of the $2m + 1$ and $4m + 1$ schemes. This is due to the fact that the $2m$ scheme concentrations depend on the number m of input random variables. This dependence causes the estimation errors incurred by this scheme to increase as m does. Note, by comparing both tables, that the estimation errors of the $2m$ scheme increase when the total number of input random variables goes from 23 in the IEEE 14-bus system to 170 in the IEEE 118-bus test system. This effect is also pointed out by the two previous works [27], [28] applying point estimate methods to power systems reported in Section I.

However, the $2m+1$ and $4m+1$ schemes keep their good behavior when m increases because their concentrations do not depend on it. This stable behavior over the number of input random variables is the reason why these $K \times m + 1$ type schemes provide better results than $K \times m$ type schemes when a realistic power system is considered.

Results for the IEEE 118-bus test system (see Table IV) also show that the average errors of estimations incurred by schemes $2m + 1$ and $4m + 1$ are much smaller than those corresponding to the estimations of the $2m$ scheme. For instance, with regard to mean values calculated by schemes $2m + 1$ and $4m + 1$, the

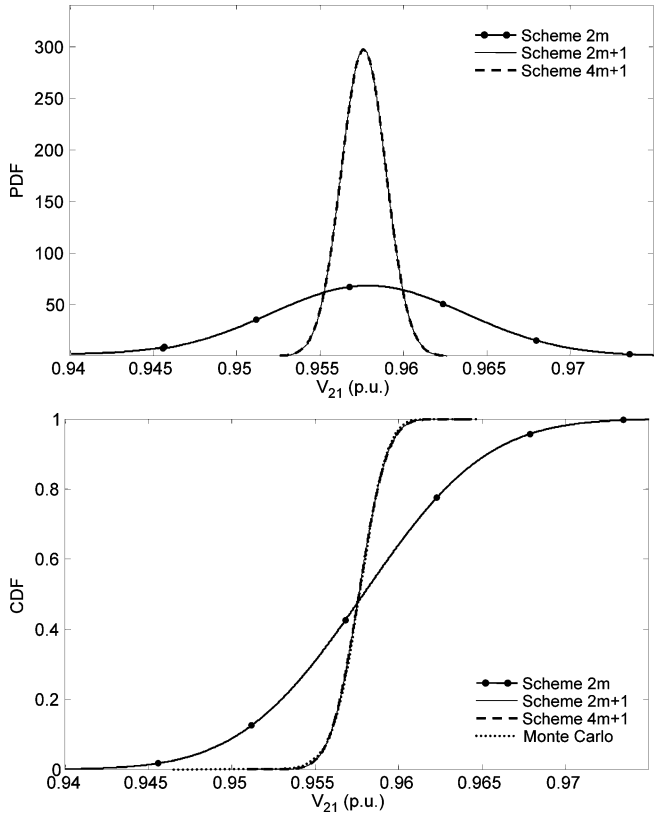

Fig. 3. Bus #21 voltage pdf and cdf.

largest average error corresponds to active power line flows and does not exceed 1%. However, in the case of scheme $2m$, the highest average error corresponds to reactive power line flows and is about 50%. As for the standard deviations, the differences between the estimations provided by the $2m$ scheme and those provided by $K \times m+1$ schemes ($2m+1$ y $4m+1$) are still more remarkable. Note, for example, that the average error associated with the estimations of scheme $2m$ for the standard deviations of reactive power injections is greater than 6000%, whereas this error does not exceed 5% in the case of schemes $2m+1$ and $4m+1$.

It can also be noted that the standard deviation average errors are greater than the mean ones in almost all the cases. This makes evident that the accuracy of the point estimate methods worsens as the order of the estimated statistical moment becomes higher. This difference is especially remarkable for the case of the $2m$ scheme applied to 118-bus test system. As an example, the average error incurred by this scheme for the mean of the reactive power injections is 39.1229%, while this error reaches a value of 6174.6% for the corresponding standard deviation. This very significant difference is even greater than what could be expected due to the effect of the large number of input random variables on the $2m$ scheme.

Fig. 3 shows the pdf and cdf of the voltage of bus #21 (V_{21}) for the IEEE-118 bus test system. To obtain these functions, the Gram–Charlier expansion is used. Cumulative density functions are also compared with those obtained using Monte Carlo simulation. In Fig. 3, the probability and cumulative density functions corresponding to the $2m+1$ scheme are almost not visible because they coincide with those of $4m+1$, and both provide

TABLE V
CPU TIME (IN SECONDS)

CPU Time	$2m$	$2m+1$	$4m+1$	Monte Carlo
IEEE-14	0.2700	0.2600	0.6610	77.8520
IEEE-118	6.9600	6.1190	12.5680	332.0370

TABLE VI
AVERAGE ERROR INDICES FOR THE MODIFIED IEEE 118-BUS TEST CASE (ONLY NORMAL DISTRIBUTIONS)

IEEE 118		V	δ	P_i	Q_i	$P_{i,j}$	$Q_{i,j}$
2m	$\bar{\varepsilon}_\mu^*$	0.0089	0.5213	0.9735	1.1038	1.3703	0.9183
	$\bar{\varepsilon}_\sigma^*$	176.1834	20.5514	17.6705	156.0192	6.6206	62.6869
3m	$\bar{\varepsilon}_\mu^*$	9.1608e-004	0.2107	0.4950	0.2192	1.1198	0.4264
	$\bar{\varepsilon}_\sigma^*$	2.9585	0.7762	0.7259	3.9953	0.6406	2.3273
2m+1	$\bar{\varepsilon}_\mu^*$	9.1608e-004	0.2107	0.4950	0.2192	1.1198	0.4264
	$\bar{\varepsilon}_\sigma^*$	2.9584	0.7762	0.7259	3.9953	0.6406	2.3273
4m+1	$\bar{\varepsilon}_\mu^*$	9.1593e-004	0.2108	0.4950	0.2192	1.1198	0.4264
	$\bar{\varepsilon}_\sigma^*$	2.9390	0.7770	0.7266	3.9762	0.6408	2.3159

good fitting to the cumulative density function obtained with the Monte Carlo method. Fig. 3 also shows that the cumulative density function estimated by the $2m$ scheme is different from the others. This is another symptom of the bad behavior of the $2m$ scheme if the total number of input random variables is high: a good fit for the mean, but a high estimation error for the standard deviation.

Table V presents the CPU time needed to compute the statistical moments of output variables for each scheme considered, as well as the Monte Carlo simulation with 10 000 evaluations. The code was developed in the MATLAB environment, and MATPOWER [35] was used to solve the deterministic power flows on a Pentium Intel Centrino 1.6-GHz PC with 512 MB of RAM.

The $4m+1$ scheme exhibits the highest computational burden due to the high number of deterministic power flow solutions performed. However, it should be noted that the $2m+1$ scheme is faster than the $2m$ scheme, even though one additional deterministic power flow has to be solved. As the total number of input random variables m increases, the $2m$ scheme locations move away from the mean, and the convergence of Newton–Rapshon method, used to solve the deterministic power flows, becomes slower. It should be taken into account that the output values obtained by solving the deterministic power flow for the mean values of all input variables are considered the initial point for the iterative procedure.

Finally, to check the $3m$ scheme performance, one more test is presented. This scheme was not previously considered due to the complex values obtained for the concentrations of the input random variables modeled with binomial distributions. To overcome this drawback, the generation plants probabilistic input data for the case study based on IEEE-118 bus test system are changed. Instead of binomial distributions, the generation of each unit is considered normally distributed, and thus, all input variables have the same distribution. Table VI shows the results obtained. It can be appreciated in this table that the $3m$ scheme average error indices match the $2m+1$ ones. In fact,

both schemes provide the same results because the $3m$ scheme, when only normal distributions are taken into account, chooses the mean as one of the three locations, and thus, it turns out to be the $2m+1$ scheme.

V. Conclusion

Point estimate methods are appropriate tools to handle stochastic power system problems because good results can be achieved by using the same routines as in the associated deterministic problems, while keeping low the computational burden involved. The use of Hong's point estimate methods provides an appropriate solution for the tradeoff between the accuracy of the results and the efficiency of the computational procedure for large-scale power system problems.

The four different Hong's point estimate schemes considered in this paper have been tested on the probabilistic power flow problem. Results over two different case studies have been presented and compared against those obtained from the Monte Carlo simulation. Binomial and normal distributions are used to model input random variables.

This paper shows that the use of the $2m+1$ scheme provides the best performance when a high number of input random variables, both continuous and discrete, is considered. Similar results are obtained by using the $4m+1$ scheme, but it implies a considerably higher computational burden. The $2m$ scheme accuracy decreases as the total number of input random variables increases, and thus, it seems not to be adequate for large-scale power system problems. Finally, the $3m$ scheme usually provides useless complex concentration values when the Binomial distribution is used to model input data uncertainty. On the other hand, if input data uncertainty modeling is reduced to normally distributed random variables, the $3m$ scheme becomes the $2m+1$ scheme, but it does not include an analytical solution for the concentrations, and $m-1$ additional deterministic power flows must be computed.

Although this paper focuses on solving the probabilistic power flow problem, the presented procedure can be applied to other power system problems affected by uncertainty, even those that include optimization algorithms, such as the probabilistic optimal power flow.

Appendix

In this appendix, the $3m$ and $4m+1$ schemes are described.

$3m$ Scheme ($K=3$): Solving analytically (2) for $K=3$ for the standard locations and weights is not possible. By applying the Miller and Rice's procedure [29], the standard locations $\xi_{l,k}$ are the roots of the following third-order polynomial $\pi(\xi) \equiv \xi^3 + C_2\xi^2 + C_1\xi + C_0$, whose coefficients are computed from

$$\begin{pmatrix} \frac{1}{m} & 0 & 1 \\ 0 & 1 & \lambda_{l,3} \\ 1 & \lambda_{l,3} & \lambda_{l,4} \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \\ C_2 \end{pmatrix} = - \begin{pmatrix} \lambda_{l,3} \\ \lambda_{l,4} \\ \lambda_{l,5} \end{pmatrix} \tag{15}$$

and then, the concentration weights are given by the expression

$$w_{l,k} = \frac{1 + \frac{1}{m} \prod\limits_{j \neq k} \xi_{l,j}}{\prod\limits_{j \neq k} (\xi_{l,j} - \xi_{l,k})}. \tag{16}$$

This concentration scheme provides theoretically a higher accuracy than the $2m$ scheme because it takes into account a more detailed statistical information of the random variable (the fourth and fifth central moments are also considered), and likewise, three points for each input random variable, instead of only two, are considered. However, as m additional evaluations of function F must be carried out, the $3m$ scheme requires a greater computational effort.

$4m+1$ Scheme ($K=5$, $\xi_{l,5}=0$): This scheme is derived from the $5m$ scheme by setting to zero one of the five standard locations (let $\xi_{l,5}=0$). As in the $3m$ scheme, there is not an analytical formula to compute the concentration values.

The other four standard locations $\xi_{l,k}$ are the roots of the fourth-order polynomial $\pi(\xi) \equiv \xi^4 + C_3\xi^3 + C_2\xi^2 + C_1\xi + C_0$, whose coefficients are the solution of the system of linear equations

$$\begin{pmatrix} 0 & 1 & \lambda_{l,3} & \lambda_{l,4} \\ 1 & \lambda_{l,3} & \lambda_{l,4} & \lambda_{l,5} \\ \lambda_{l,3} & \lambda_{l,4} & \lambda_{l,5} & \lambda_{l,6} \\ \lambda_{l,4} & \lambda_{l,5} & \lambda_{l,6} & \lambda_{l,7} \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \end{pmatrix} = - \begin{pmatrix} \lambda_{l,5} \\ \lambda_{l,6} \\ \lambda_{l,7} \\ \lambda_{l,8} \end{pmatrix}. \tag{17}$$

Once the standard locations are obtained, the weights are determined by solving the linear system

$$\begin{pmatrix} \xi_{l,1} & \xi_{l,2} & \xi_{l,3} & \xi_{l,4} \\ \xi_{l,1}^2 & \xi_{l,2}^2 & \xi_{l,3}^2 & \xi_{l,4}^2 \\ \xi_{l,1}^3 & \xi_{l,2}^3 & \xi_{l,3}^3 & \xi_{l,4}^3 \\ \xi_{l,1}^4 & \xi_{l,2}^4 & \xi_{l,3}^4 & \xi_{l,4}^4 \end{pmatrix} \begin{pmatrix} w_{l,1} \\ w_{l,2} \\ w_{l,3} \\ w_{l,4} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \lambda_{l,3} \\ \lambda_{l,4} \end{pmatrix} \tag{18}$$

and taking into account that the weight $w_{l,5}$ (which corresponds to $\xi_{l,5}=0$) is given by

$$w_{l,5} = \frac{1}{m} - \sum_{k=1}^{4} w_{l,k}. \tag{19}$$

Once again, it should be noted in (1) that setting $\xi_{l,5}=0$ yields $p_{l,k} = \mu_{p_l}$, and therefore, m of the $5m$ locations are the same $(\mu_{p_1}, \mu_{p_2}, \ldots, \mu_{p_l}, \ldots, \mu_{p_m})$. Hence, the weight of this concentration must be updated to the value w_0

$$w_0 = \sum_{l=1}^{m} w_{l,5} = 1 - \sum_{l=1}^{m} \sum_{k=1}^{4} w_{l,k}. \tag{20}$$

For the above reason, the $4m+1$ scheme can be viewed as a $5m$ scheme with a computational burden reduction of $m-1$ evaluations of F.

The higher the number of points used by a point estimate method, the more accurate estimation of the statistical moments of the output random variables is achieved. Nevertheless, if the number of points increases, more information about the input random variables is required, a greater computational burden is needed, and the probability of obtaining non-real solutions, or even an infinite set of solutions, raises.

Acknowledgment

The authors would like to thank Prof. A. J. Conejo for his assistance in the preparation of this paper.

REFERENCES

[1] G. J. Anders, *Probability Concepts in Electric Power Systems*. New York: Wiley, 1990.

[2] A. R. Bergen, *Power Systems Analysis*, ser. Electrical and Computer Engineering. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[3] A. J. Wood and B. F. Wollenberg, *Power Generation Operation and Control*. New York: Wiley, 1996.

[4] B. Borkowska, "Probabilistic load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, pp. 752–759, 1974.

[5] J. F. Dopazo, O. A. Klitin, and A. M. Sasson, "Stochastic load flows," *IEEE Trans. Power App. Syst.*, vol. PAS-94, pp. 299–309, 1975.

[6] Y.-K. Tung and B.-C. Yen, *Hydrosystems Engineering Uncertainty Analysis*. New York: McGraw-Hill, 2005.

[7] R. Y. Rubinstein, *Simulation and the Monte Carlo Method*. New York: Wiley, 1981.

[8] R. N. Allan and A. M. Leite da Silva, "Probabilistic load flow using multilinearizations," *Proc. Inst. Elect. Eng. C: Gen., Transm., Distrib.*, vol. 128, no. 5, pp. 280–287, 1981.

[9] A. P. S. Meliopoulos, G. J. Cokkinides, and X. Y. Chao, "A new probabilistic power flow analysis method," *IEEE Trans. Power Syst.*, vol. 5, no. 1, pp. 182–190, Feb. 1990.

[10] R. N. Allan, A. M. Leite da Silva, and R. C. Burchett, "Evaluations methods and accuracy in probabilistic load flow solutions," *IEEE Trans. Power App. Syst.*, vol. PAS-100, pp. 2539–2546, May 1981.

[11] R. N. Allan and M. R. G. Al-Shakarchi, "Probabilistic techniques in a.c. load flow analysis," *Proc. IEEE*, vol. 124, no. 2, pp. 154–160, Feb. 1977.

[12] 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.

[13] L. A. Sanabria and T. S. Dillon, "Stochastic power flow using cumulants and von Mises functions," *Int. J. Elect. Power Energy Syst.*, vol. 8, no. 1, pp. 47–60, 1986.

[14] V. Miranda, M. A. Matos, and J. T. Saraiva, "Fuzzy load flow new algorithms incorporating uncertain generation and load representation," in *Proc. 10th Power System Computation Conf.*, Graz, Austria, 1990, pp. 621–627.

[15] A. M. Leite da Silva, V. L. Arienti, and R. N. Allan, "Probabilistic load flow considering dependence between input nodal powers," *IEEE Trans. Power App. Syst.*, vol. PAS-103, pp. 1524–1530, Jun. 1984.

[16] A. M. Leite da Silva and V. L. Arienti, "Probabilistic load flow by a multilinear simulation algorithm," *Proc. Inst. Elect. Eng. C: Gen., Transm., Distrib.*, vol. 137, no. 4, pp. 276–282, Jul. 1990.

[17] M. Madrigal, K. Ponnambalam, and V. H. Quintana, "Probabilistic optimal power flow," in *Proc. 1998 IEEE Can. Conf. Electrical and Computer Engineering*, Waterloo, ON, Canada, May 1998, pp. 385–388.

[18] E. Rosenblueth, "Point estimates for probability moments," *Proc. Nat. Acad. Sci.*, vol. 72, pp. 3812–3814, Oct. 1975.

[19] E. Rosenblueth, "Two-point estimates in probability," *Appl. Math. Model.*, vol. 5, pp. 329–335, Oct. 1981.

[20] K. S. Li, "Point-estimate method for calculating statistical moments," *J. Eng. Mech.-ASCE*, vol. 118, no. 7, pp. 1506–1511, 1992.

[21] H. S. Seo and B. M. Kwak, "Efficient statistical tolerance analysis for general distributions using three-point information," *Int. J. Prod. Res.*, vol. 40, no. 4, pp. 931–944, 2002.

[22] C. W. Tsai and S. Franceschini, "Evaluation of probabilistic point estimate methods in uncertainty analysis for environmental engineering applications," *J. Environ. Eng.-ASCE*, vol. 131, no. 3, pp. 387–395, 2005.

[23] W. Liping, D. Beeson, and G. Wiggs, "Efficient and accurate point estimate method for moments and probability distribution estimation," in *Proc. 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conf.*, Albany, NY, Aug.-Sep. 30–1, 2004.

[24] M. E. Harr, "Probabilistic estimates for multivariate analysis," *Appl. Math. Model.*, vol. 13, no. 5, pp. 313–318, 1989.

[25] H. P. Hong, "An efficient point estimate method for probabilistic analysis," *Reliab. Eng. Syst. Saf.*, vol. 59, pp. 261–267, 1998.

[26] C.-L. Su, "A new probabilistic load flow method," in *Proc. IEEE Power Eng. Soc. General Meeting*, Jun. 2005, pp. 1795–1800.

[27] C.-L. Su, "Probabilistic load-flow computation using point estimate method," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1843–1851, Nov. 2005.

[28] G. Verbič and C. A. Cañizares, "Probabilistic optimal power flow in electricity markets based on a two-point estimate method," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1883–1893, Nov. 2006.

[29] A. C. Miller and T. R. Rice, "Discrete approximations of probability distributions," *Manage. Sci.*, vol. 29, no. 3, pp. 352–362, Mar. 1983.

[30] J. M. Morales, *Application of Point Estimate Methods to Probabilistic Power Flow Problem* (in Spanish). Málaga, Spain: E. T. S. Ingenieros Industriales. Univ. Málaga, 2006.

[31] J. T. Christian and G. B. Baecher, "The point-estimate method with large numbers of variables," *Int. J. Numer. Anal. Meth. Geomech.*, vol. 26, pp. 1515–1529, 2002.

[32] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics. Volume 1. Distribution Theory*. London, U.K.: Charles Griffin, 1963.

[33] *Power Systems Test Case Archive.* [Online]. Available: http://www.ee.washington.edu/research/pstca.

[34] L. A. Sanabria and T. S. Dillon, "Power system reliability assessment suitable for a deregulated system via the method of cumulants," *Int. J. Elect. Power Energy Syst.*, vol. 20, no. 3, pp. 203–211, 1998.

[35] R. Zimmerman, C. Murillo-Sánchez, and D. Gan, *MATPOWER, a MATLAB Power System Simulation Package*, 3.0.0. ed. Ithaca, NY: Cornell University: Power System Engineering Research Center (PSERC), School of Electrical Engineering, 2005.

Juan M. Morales (S'06) received the Ingeniero Industrial degree from the Universidad de Málaga, Málaga, Spain, in 2006. He is currently pursuing the Ph.D. degree at the Universidad de Castilla-La Mancha, Ciudad Real, Spain.

His research interests are in the fields of power systems economics, reliability, stochastic programming, and electricity markets.

Juan Pérez-Ruiz (S'96–M'99) received the Electrical Engineering degree from the Universidad Pontificia Comillas, Madrid, Spain, in 1992 and the Ph.D. degree from the Universidad de Málaga, Málaga, Spain, in 2001.

He is currently an Associate Professor at the Universidad de Málaga. His research interests include operations, planning, and economics of electric energy systems as well as statistics and optimization theory and its applications.