Open access • Journal Article • DOI:10.1109/TPWRS.2005.846184

Cumulant-based probabilistic optimal power flow (P-OPF) with Gaussian and gamma distributions — Source link

A. Schellenberg, William Rosehart, José A. Aguado

Institutions: University of Calgary, University of Málaga

Published on: 02 May 2005 - IEEE Transactions on Power Systems (IEEE)

Topics: Monte Carlo method, Gamma distribution, Gaussian, Hessian matrix and Higher-order statistics

Cumulant-Based Probabilistic Optimal Power Flow (P-OPF) With Gaussian and Gamma Distributions

Antony Schellenberg, William Rosehart, and José Aguado

***Abstract*—This paper introduces the cumulant method for the probabilistic optimal power flow (P-OPF) problem. By noting that the inverse of the Hessian used in the logarithmic barrier interior point can be used as a linear mapping, cumulants can be computed for unknown system variables.**

Results using the proposed cumulant method are compared against results from Monte Carlo simulations (MCSs) based on a small test system. The Numerical Results section is broken into two sections: The first uses Gaussian distributions to model system loading levels, and cumulant method results are compared against four MCSs. Three of the MCSs use 1500 samples, while the fourth uses 20 000 samples. The second section models the loads with a Gamma distribution. Results from the proposed technique are compared against a 1000-point MCS.

The cumulant method agrees very closely with the MCS results when the mean value for variables is considered. In addition, the proposed method has significantly reduced computational expense while maintaining accuracy.

***Index Terms*—Cumulants, optimal power flow (OPF), probabilistic optimization.**

I. Introduction

OPTIMAL POWER FLOW (OPF) is a tool that has been commonly used within the power systems industry for many years [1] and has generally been addressed as a deterministic optimization problem. However, it is becoming increasingly important that solution methods to the optimal power flow problem be developed to address probabilistic quantities and, thus, transform the optimal power flow problem into the *probabilistic optimal power flow (P-OPF)* problem [2].

Probabilistic programming, or probabilistic optimization, is concerned with the introduction of probabilistic randomness or uncertainty into conventional linear and nonlinear programs [3]. However, the randomness introduced tends to have some structure to it, and this structure is generally represented with a *probability density function (PDF)* [4]. The goal of the P-OPF problem is to determine the PDFs for all variables in the problem. These PDFs are the distributions of the optimal solutions. A typical example of an uncertain or probabilistic parameter in a P-OPF problem is bus loading.

The cumulant method for *probabilistic power flow* was discussed in [5] and [6]. The present paper briefly outlines the fundamentals of the cumulant method and presents the adaptation of the method in [5] to the P-OPF problem using a logarithmic barrier interior point method (LBIPM) [7]-type solution.

This paper is structured in the following manner. Section II presents information related to the Edgeworth form of the Gram–Charlier A series. It includes the A series itself, in addition to some background information on Tchebycheff–Hermite polynomials and computation of A series coefficients. Next, in Section III, an overview of the pure Newton step in the LBIPM for numerical programming is provided. In Sections IV and V, the *cumulant method* is presented, in addition to the proposed application to the P-OPF problem. Numerical results from a system based on the Matpower 9-bus system [8] using normally (Gaussian) and Gamma distributed independent random loads with the proposed cumulant method for P-OPF are detailed in Section VI. Finally, conclusions are presented in Section VII.

Two appendixes are included to provide background information in probability and statistics, focusing on moments and cumulants, as well as information on Gaussian and Gamma distributions.

II. Gram–Charlier A Series

The Gram–Charlier A Series allows many PDFs, including Gaussian and Gamma distributions, to be expressed as a series composed of a standard normal distribution and its derivatives. As a part of the proposed P-OPF method, distributions are reconstructed with the use of the Gram–Charlier A Series. Additional information can be found in [9]. The series can be stated as follows:

$$f(x) = \sum_{j=0}^{\infty} c_j He_j(x)\alpha(x) \tag{1}$$

where $f(x)$ is the PDF for the random variable $\mathbf{x}, c_j$ is the jth series coefficient, $He_j(x)$ is the jth Tchebycheff–Hermite, or Hermite, polynomial, and $\alpha(x)$ is the standard normal distribution function (see Appendix I-A).

The Gram–Charlier form uses moments to compute series coefficients, while the Edgeworth form uses cumulants. Since the work presented in this paper is based on cumulants, only Edgeworth's form of the A series is discussed.

Throughout this section, the operator D is defined as the derivative with respect to x to simplify notation.

The remainder of this section is devoted to discussion of the Hermite polynomials and the computations of A series coefficients in (1).

Manuscript received June 10, 2004; revised November 17, 2004. Paper no. TPWRS-00300-2004.

A. Schellenberg and W. Rosehart are with the University of Calgary, Calgary, AB T2N 1N4, Canada (e-mail: schellen@enel.ucalgary.ca; rosehart@enel.ucalgary.ca).

J. Aguado is with the University of Malaga, 29071 Malaga, Spain (e-mail: jaguado@uma.es).

Digital Object Identifier 10.1109/TPWRS.2005.846184

A. Tchebycheff–Hermite Polynomials

There are two different forms for Hermite polynomials [10]. The first is based on e^{-x^2} and the second on $e^{-(1/2)x^2}$. The second form is the same as the PDF for a standard normal distribution and is more convenient for this application. To avoid confusion, the notation He is used in this paper to denote the use of the second type of Hermite polynomial.

Since the PDF for a normal distribution is an exponential term, taking derivatives successively returns the original function with a polynomial coefficient multiplier. These coefficients are referred to as Tchebycheff–Hermite, or Hermite, polynomials.

To illustrate how the Hermite polynomials are generated, the first four derivatives of the standard unit normal distribution are taken as follows:

$$D^0\alpha(x) = D^0 e^{-\frac{1}{2}x^2} = e^{-\frac{1}{2}x^2} \tag{2}$$
$$D^1\alpha(x) = D^1 e^{-\frac{1}{2}x^2} = -xe^{-\frac{1}{2}x^2} \tag{3}$$
$$D^2\alpha(x) = D^2 e^{-\frac{1}{2}x^2} = (x^2-1)e^{-\frac{1}{2}x^2} \tag{4}$$
$$D^3\alpha(x) = D^3 e^{-\frac{1}{2}x^2} = (3x-x^3)e^{-\frac{1}{2}x^2} \tag{5}$$
$$D^4\alpha(x) = D^4 e^{-\frac{1}{2}x^2} = (x^4-6x^2+3)e^{-\frac{1}{2}x^2} \tag{6}$$

where D^n is the nth derivative.

The Tchebycheff–Hermite polynomials are the polynomial coefficients in the derivatives. Using the results of the first four derivatives in (2)–(6), the first five Tchebycheff–Hermite polynomials are written as follows:

$$He_0(x) = 1 \tag{7}$$
$$He_1(x) = x \tag{8}$$
$$He_2(x) = x^2 - 1 \tag{9}$$
$$He_3(x) = x^3 - 3x \tag{10}$$
$$He_4(x) = x^4 - 6x^2 + 3. \tag{11}$$

Because of the structure of (2)–(6), the highest power coefficient of the odd derivatives, i.e., the third, fifth, seventh, etc., are negative. Equations (7)–(11) have been formed following the convention that the equations relating to the odd derivatives are multiplied by negative one, such that the coefficient of the highest power is positive [9].

Therefore, the nth Tchebycheff–Hermite polynomial can be symbolically written as

$$He_n(x)\alpha(x) = (-D)^n\alpha(x). \tag{12}$$

In addition, a recursive relationship is available to determine third-order and higher polynomials

$$He_n(x) = xHe_{n-1}(x) - (n-1)He_{n-2}. \tag{13}$$

B. Edgeworth A-Series Coefficients

Given the cumulants for a distribution in standard form, i.e., zero mean and unit variance, the coefficients for the Edgeworth form of the A series can be computed. In order to find the equations for the A series coefficients, an exponential representation of the PDF is broken into its series representation and equated with the Gram–Charlier A series (1).

The PDF, as an exponential, is written in the following form using cumulants [9]:

$$f(x) = e^{(-\frac{K_3}{3!}D^3 + \frac{K_4}{4!}D^4 - \frac{K_5}{5!}D^5 + \cdots)}\alpha(x) \tag{14}$$

where D^n is the nth derivative of the unit normal distribution, K_n is the nth cumulant, and $\alpha(x)$ is the standard unit normal PDF. A complete derivation of (14) can be found in [9].

Expanding (14) as an exponential series yields

$$f(x) = \left[1 + \frac{\left(-\frac{K_3}{3!}D^3 + \frac{K_4}{4!}D^4 - \frac{K_5}{5!}D^5 + \cdots\right)}{1!} + \frac{\left(-\frac{K_3}{3!}D^3 + \frac{K_4}{4!}D^4 - \frac{K_5}{5!}D^5 + \cdots\right)^2}{2!} + \frac{\left(-\frac{K_3}{3!}D^3 + \frac{K_4}{4!}D^4 - \frac{K_5}{5!}D^5 + \cdots\right)^3}{3!} + \cdots\right]\alpha(x). \tag{15}$$

If each of the terms are expanded individually and grouped based on powers of D, the following result is obtained:

$$f(x) = \left[1 - \frac{K_3}{3!}D^3 + \frac{K_4}{4!}D^4 - \frac{K_5}{5!}D^5 + \left(\frac{K_6}{6!} + \frac{K_3^2}{2!3!^2}\right)D^6 - \left(\frac{K_7}{7!} + \frac{2K_3K_4}{2!3!4!}\right)D^7 + \cdots\right]\alpha(x). \tag{16}$$

Using the relationship for Hermite polynomials in (12) to replace the powers of D in (16) gives

$$f(x) = 1 + \frac{K_3}{3!}He_3\alpha(x) + \frac{K_4}{4!}He_4\alpha(x) + \frac{K_5}{5!}He_5\alpha(x) + \frac{(K_6+10K_3^2)}{6!}He_6\alpha(x) + \frac{(K_7+35K_3K_4)}{7!}He_7\alpha(x) + \cdots \tag{17}$$

Returning to the definition for the Gram–Charlier A series in (1) and expanding the summation yields

$$f(x) = c_0He_0\alpha(x) + c_1He_1\alpha(x) + c_2He_2\alpha(x) + \cdots \tag{18}$$

Comparing (17) and (18), the values for the coefficients can be determined. Based on the equations presented, the first seven terms of the Edgeworth form of the A series are presented in Table I.

III. Introduction to the LBIPM

The LBIPM has been used in a wide variety of applications. It is a method for the solution of constrained optimization problems and makes use of a variety of techniques from many contributors. Detailed information for the LBIPM can be found in [7], [11], and [12].

TABLE I
A SERIES COEFFICIENT EQUATIONS

Coefficient	Equation
0	1
1	0
2	0
3	$\frac{1}{6}K_3$
4	$\frac{1}{24}K_4$
5	$\frac{1}{120}K_5$
6	$\frac{1}{720}(K_6+10K_3^2)$
7	$\frac{1}{5040}(K_7+35K_3K_4)$

The algorithm proposed in this paper uses some matrices and components computed as part of the Newton step in the LBIPM. Therefore, a very brief outline of the Newton step and its computation in an LBIPM framework is included.

A. Gradient, Hessian, and Newton Direction

Any optimal solution must satisfy the condition that the gradient of the Lagrangian is zero. This condition is known as the Karush–Kuhn–Tucker (KKT) first-order condition for optimality.

A Newton–Raphson-based solver can be used to solve this system of nonlinear equations. In this case, the Newton–Raphson method attempts to solve the system using the following relationship:

$$H(y)\Delta y = -\nabla_{\mathcal{L}}(y) \tag{19}$$

where y is a vector of primal and dual variables in the optimization problem, and $H(y)$ and $\nabla_{\mathcal{L}}(y)$ are the Hessian and the gradient of the Lagrangian, respectively, evaluated for the variable values at the current iteration. The update step Δy is computed from the Newton–Raphson method and is known as the *Newton step* or, alternatively, the *Newton direction* if magnitude is normalized. System variables are updated via the following equation:

$$y_{k+1} = y_k + \alpha\Delta y. \tag{20}$$

Step lengths are chosen to ensure that the resulting point remains within the feasible space.

IV. CUMULANT METHOD

The cumulant method relies on the behavior of random variables and their associated cumulants when they are combined in a linear fashion. This section discusses the formation of random variables from a linear combination of others and the role cumulants play in this combination.

Given a new random variable $\mathbf{z}$, which is the linear combination of independent random variables, $\mathbf{c_1}, \mathbf{c_2}, \ldots, \mathbf{c_n}$

$$\mathbf{z} = a_1\mathbf{c_1} + a_2\mathbf{c_2} + \cdots + a_n\mathbf{c_n} \tag{21}$$

the moment generating function $\Phi_{\mathbf{z}}(s)$ for the random variable $\mathbf{z}$ can be written as follows:

$$\Phi_{\mathbf{z}}(s) = E[e^{sz}] \tag{22a}$$
$$= E\left[e^{s(a_1\mathbf{c_1}+a_2\mathbf{c_2}+\cdots+a_n\mathbf{c_n})}\right] \tag{22b}$$
$$= E\left[e^{s(a_1\mathbf{c_1})}e^{s(a_2\mathbf{c_2})}\ldots e^{s(a_n\mathbf{c_n})}\right]. \tag{22c}$$

Since $\mathbf{c_1}, \mathbf{c_2}, \ldots, \mathbf{c_n}$ are independent

$$\Phi_{\mathbf{z}}(s) = E\left[e^{s(a_1\mathbf{c_1})}\right]E\left[e^{s(a_2\mathbf{c_2})}\right]\ldots E\left[e^{s(a_n\mathbf{c_n})}\right] \tag{23a}$$
$$= \Phi_{\mathbf{c_1}}(a_1s)\Phi_{\mathbf{c_2}}(a_2s)\ldots\Phi_{\mathbf{c_n}}(a_ns). \tag{23b}$$

The cumulants for the variable $\mathbf{z}$ can be computed using the cumulant generating function, as defined in Appendix II, in terms of the component variables as follows:

$$\Psi_{\mathbf{z}}(s) = \ln(\Phi_{\mathbf{z}}(s)) \tag{24a}$$
$$= \ln(\Phi_{\mathbf{c_1}}(a_1s)\Phi_{\mathbf{c_2}}(a_2s)\ldots\Phi_{\mathbf{c_n}}(a_ns)) \tag{24b}$$
$$= \ln(\Phi_{\mathbf{c_1}}(a_1s)) + \ln(\Phi_{\mathbf{c_2}}(a_2s)) + \cdots + \ln(\Phi_{\mathbf{c_n}}(a_ns)) \tag{24c}$$
$$= \Psi_{\mathbf{c_1}}(a_1s) + \Psi_{\mathbf{c_2}}(a_2s) + \cdots + \Psi_{\mathbf{c_n}}(a_ns). \tag{24d}$$

To compute the second-order cumulant, the zero-, first-, and second-order derivatives of the cumulant generating function for the random variable $\mathbf{z}$ are computed as

$$\Psi_{\mathbf{z}}(s) = \Psi_{\mathbf{c_1}}(a_1s) + \Psi_{\mathbf{c_2}}(a_2s) + \cdots + \Psi_{\mathbf{c_n}}(a_ns) \tag{25}$$
$$\Psi'_{\mathbf{z}}(s) = a_1\Psi'_{\mathbf{c_1}}(a_1s) + a_2\Psi'_{\mathbf{c_2}}(a_2s) + \cdots + a_n\Psi'_{\mathbf{c_n}}(a_ns) \tag{26}$$
$$\Psi''_{\mathbf{z}}(s) = a_1^2\Psi''_{\mathbf{c_1}}(a_1s) + a_2^2\Psi''_{\mathbf{c_2}}(a_2s) + \cdots + a_n^2\Psi''_{\mathbf{c_n}}(a_ns). \tag{27}$$

Evaluating (27) at $s = 0$ gives

$$\Psi''_{\mathbf{z}}(0) = a_1^2\Psi''_{\mathbf{c_1}}(0) + a_2^2\Psi''_{\mathbf{c_2}}(0) + \cdots + a_n^2\Psi''_{\mathbf{c_n}}(0). \tag{28}$$

Third- and higher order cumulants can be computed following the same procedure. In general, the nth-order cumulant for $\mathbf{z}$, a linear combination of independent random variables, can be determined with the following equation:

$$\lambda_n = \Psi_{\mathbf{z}}^{(n)}(0) \tag{29a}$$
$$= a_1^n\Psi_{\mathbf{c_1}}^{(n)}(0) + a_2^n\Psi_{\mathbf{c_2}}^{(n)}(0) + \ldots + a_n^n\Psi_{\mathbf{c_n}}^{(n)}(0) \tag{29b}$$

where the exponent (n) denotes the nth derivative with respect to s.

V. ADAPTATION OF THE CUMULANT METHOD TO P-OPF PROBLEM

The cumulant method is adapted from the basic derivation above to accommodate the P-OPF problem when an LBIPM-type solution is used. The Hessian of the Lagrangian is necessary for the computation of the Newton step in the LBIPM. The

inverse of the Hessian, however, can be used as the coefficients for the linear combination of random bus loading variables.

A. Inverse Hessian as Linear MAP

The pure Newton step is computed at iteration k of the LBIPM using the following equation:

$$H(y_k)\Delta y_k = -F(y_k) \tag{30}$$

where y is the vector of variables, $H(y_k)$ and $F(y_k)$ are Hessian and gradient of the Lagrangian, respectively, evaluated at y_k, and Δy_k is the pure Newton step. Replacing Δy_k with $y_{k+1}-y_k$ in (30) and rearranging gives

$$y_{k+1} = -H^{-1}(y_k)F(y_k) + y_k. \tag{31}$$

A general linear equation can be written as

$$l = mn + b \tag{32}$$

where m is the slope, n is the variable, and b is the y-intercept. Noting the similarities in the form of (32) and (31), the matrix $-H^{-1}(y_k)$, the inverse Hessian, in (31) contains the multipliers for a linear combination of PDFs for random bus loads. Alternatively stated, the negative inverse Hessian is a linear map from one variable to another.

B. Including Random Loads

It is necessary to introduce the cumulants related to the random loads into the system in such a way that the cumulants for all other system variables can be computed. Some characteristics of the gradient of the Lagrangian are used to accomplish this.

When the gradient of the Lagrangian is taken, the power flow equations appear unmodified in this vector. Therefore, cumulant models in the bus loads map directly into the gradient of the Lagrangian. For the purposes of mapping, the mismatch vector, $F(y_k)$ in (31), is replaced by a new vector containing the cumulants of the random loads in the rows corresponding to their associated power flow equations.

C. Generalized Results

The linear mapping information contained in the inverse Hessian can be used to determine cumulants for other variables when bus loading is treated as a random variable. If $-H^{-1}(y_k)$ is written in the following form

$$-H^{-1} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,n} \end{bmatrix} \tag{33}$$

then the nth cumulant for the ith variable in y is computed using the following equation:

$$\lambda_{yi,n} = a^n_{i,1}\lambda_{x_1,n} + a^n_{i,2}\lambda_{x_2,n} + \cdots + a^n_{i,n}\lambda_{x_n,n} \tag{34}$$

where yi is the ith element in y, and $\lambda_{x_j,n}$ is the nth cumulant for the jth component variable.

In the proposed cumulant method for P-OPF, the cumulants for unknown random variables are computed from known random variables, and PDFs are reconstructed using the Gram–Charlier/Edgeworth Expansion theory [5].

D. Computation and Use of the Statistical Step

The proposed algorithm makes use of statistical information during every iteration of the LBIPM.

The statistical step is computed at each iteration in the following fashion.

1) Distributions for system variables are reconstructed based on computed cumulants and the Edgeworth form of the A series.
2) The difference between the variable values for the current iteration and the peak values of the distributions is computed.
3) A step, known as the statistical step, is set equal to this difference.

The statistical step is combined with the pure Newton step in a linear fashion to produce a step that is applied to system variables. The emphasis on the Newton step increases as the LBIPM progresses toward a solution to ensure good convergence behavior. The original variable update (20) can be rewritten in the following fashion to include the linear combination of the pure Newton and statistical steps:

$$y_{k+1} = y_k + \alpha(\gamma\Delta y_{\text{Newton}} + \eta\Delta y_{\text{Statistical}}) \tag{35}$$

where γ and η are the scalar weighting for the linear combination used to vary the emphasis between the pure Newton step and the statistical step.

It is particularly noteworthy that when Gaussian distributions are used, there is no obvious statistical step available since the mean of the distribution corresponds to the peak in all cases. Therefore, the statistical step in this case is always zero.

In the case of non-Gaussian distributions, including Gamma distributions, the peak of the PDF does not generally correspond to the mean. Consequently, a statistical step is available for use in the procedure described.

The statistical step is introduced to place greater emphasis on optimizing around parameter settings that are more likely to occur. In this paper, the weighting factors for the statistical step have been arbitrarily chosen, and a range of different weightings have been tested. However, convergence problems developed when the statistical step was heavily weighted compared to the Newton step. It was found, in the simulations, that when the initial weighting of the statistical step was greater than $\eta = 0.85$, where $\gamma = 1 - \eta$, the solution tended to fail. This is expected since, when using a primal-dual interior point approach, the Newton step is required to converge to the optimal solution. For all systems tested in this paper, the program converged to the same point regardless of values of η and γ, subject to them being below the threshold discussed above. Therefore, final results using the statistical step were identical to the results if the distribution mapping was applied only once after the optimization was completed.

VI. Numerical Results

The proposed cumulant method was tested using random bus loads with the mean value set at the nominal bus loading

TABLE II
MEAN VALUE COMPARISON TABLE

Parameter	Bus	Cumulant	MC 20,000		MC 1,500 #1		MC 1,500 #2		MC 1,500 #3	
			Value	Difference	Value	Difference	Value	Difference	Value	Difference
Voltage	1	1.1000	1.1000	0.0025	1.1000	0.0027	1.1000	0.0027	1.1000	0.0032
	2	1.0974	1.0974	0.0024	1.0974	0.0043	1.0974	0.0045	1.0974	0.0060
	3	1.0866	1.0866	0.0013	1.0866	0.0013	1.0866	0.0013	1.0867	0.0034
	4	1.0942	1.0942	0.0027	1.0942	0.0046	1.0942	0.0045	1.0942	0.0045
	5	1.0844	1.0844	0.0016	1.0844	0.0029	1.0844	0.0023	1.0844	0.0030
	6	1.1000	1.1000	0.0000	1.1000	0.0000	1.1000	0.0000	1.1000	0.0000
	7	1.0895	1.0895	0.0009	1.0895	0.0009	1.0895	0.0012	1.0895	0.0035
	8	1.1000	1.1000	0.0000	1.1000	0.0000	1.1000	0.0000	1.1000	0.0000
	9	1.0718	1.0717	0.0035	1.0717	0.0090	1.0717	0.0090	1.0717	0.0070
Angle	1	0	0	0	0	0	0	0	0	0
	2	0.0854	0.0854	0.0120	0.0855	0.0943	0.0855	0.0835	0.0854	0.0642
	3	0.0567	0.0567	0.0048	0.0568	0.1576	0.0568	0.1538	0.0567	0.0335
	4	-0.0430	-0.0430	0.0478	-0.0431	0.1648	-0.0430	0.1446	-0.0430	0.0424
	5	-0.0695	-0.0695	0.0057	-0.0695	0.0704	-0.0695	0.0452	-0.0695	0.0382
	6	0.0105	0.0105	0.1126	0.0106	0.3118	0.0106	0.3639	0.0105	0.2849
	7	-0.0209	-0.0209	0.1977	-0.0209	0.0803	-0.0209	0.0247	-0.0210	0.3786
	8	0.0158	0.0158	0.2007	0.0158	0.0235	0.0158	0.0073	0.0157	0.4437
	9	-0.0806	-0.0806	0.0863	-0.0808	0.2493	-0.0807	0.2326	-0.0806	0.1112
Active Power	1	0.8980	0.8984	0.0420	0.8994	0.1569	0.8992	0.1367	0.8983	0.0339
	2	1.3432	1.3437	0.0335	1.3449	0.1254	1.3447	0.1088	1.3436	0.0281
	3	0.9419	0.9422	0.0330	0.9430	0.1238	0.9429	0.1073	0.9421	0.0272
Reactive Power	1	0.1297	0.1297	0.0490	0.1301	0.3650	0.1301	0.3378	0.1299	0.2179
	2	0.0003	0.0008	62.3015	0.0013	75.2012	0.0013	75.7508	0.0015	79.1732
	3	-0.2263	-0.2261	0.1243	-0.2260	0.1476	-0.2260	0.1452	-0.2256	0.3067

level. Problems based on the Matpower nine- and 118-bus systems [8] are used to show general trends and characteristics under random loading conditions using the proposed cumulant method. Of particular interest are the optimal distributions for the decision variables.

Two different and independent sets of results are included. In the first set, loads are modeled in the nine-bus and 118-bus systems using Gaussian distributions with variances such that the 99% confidence interval is equal to $\pm 10\%$ of the nominal loading value. The second set models loads using Gamma distributions such that the variance is 15% of the nominal loading value, and only results for the nine-bus system are presented. For all problems, the problem converges when the barrier parameter μ is less than 10^{-8} as computed using the complementary gap.

A. Gaussian Distributions

The results for the Gaussian distributions are divided into three sections. The first section presents the results for the mean value of the distributions and includes discussion about these results. The second section presents and discusses the results for the variance of the distributions. The first and second sections use the nine-bus problem, while the third section presents results using the 118-bus system to illustrate the cumulant method's performance as the system size increases.

In all cases, the mean value for bus loading was taken at the nominal loading value from the Matpower problems, and the variance is such that the 99% confidence interval is 10% of the nominal value.

For the nine-bus system, a total of four Monte Carlo simulations (MCSs) are included. Three were run using 1500 samples and one with 20 000 samples. The raw results are included in both sections in addition to the comparison between the cumulant method results and the MCS results.

1) Mean Values: Table II contains all results related to the mean values for system variables. Values for the system variables are included, in per unit, as well as a comparison between the cumulant method and each of the four MCSs presented as an absolute percent difference. Columns labeled "value" are the actual value of the variable in p.u., while columns titled "difference" are the absolute percent difference between the cumulant method and the MCS results.

The results for the mean value of the distributions using the cumulant method are, in general, well within 1% of the values found using MCS. With the exception of the reactive power generation at bus 2, the maximum percent difference between the mean from any of the MCSs and the cumulant method is 0.4437% and occurred for the angle at bus 8 in the third 1500 sample MCS. The results for reactive power generation at bus 2, however, had a very high percent difference between MCS and cumulant method. It is noteworthy that the minimum absolute percent difference of 62.3% for reactive power generation at this bus resulted from a difference of 0.0005 p.u. Similarly, the maximum difference of 79.17% was from a difference of only 0.0012 p.u. Although the percent difference is for reactive power generation at this bus is high, the actual error in per unit is small. Therefore, the percent error is a somewhat misleading measure for these situations.

In general, the cumulant method approximation matches very well with the MCS results with respect to the mean values.

2) Variance Values: The covariances are calculated numerically from the results in the MCSs. The results from the MCSs are presented in Table III along with the results using the proposed cumulant method to allow for a direct comparison. Again, columns labeled "value" are the actual value of the variable in p.u., while columns titled "difference" are the absolute percent difference between the cumulant method and the MCS results.

TABLE III
VARIANCE VALUE COMPARISON TABLE

Parameter	Bus	Cumulant	MC 20,000		MC 1,500 #1		MC 1,500 #2		MC 1,500 #3	
			Value	Difference	Value	Difference	Value	Difference	Value	Difference
Voltage	1	6.26e-23	2.71e-08	100.0000	2.99e-08	100.0000	3.06e-08	100.0000	3.90e-08	100.0000
	2	5.94e-07	5.86e-07	1.4168	5.94e-07	2.63e-03	5.99e-07	0.9146	6.02e-07	1.3554
	3	2.23e-07	2.20e-07	1.1913	2.12e-07	5.1661	2.16e-07	3.3384	2.15e-07	3.4815
	4	4.68e-07	4.24e-07	10.2268	4.12e-07	13.4101	4.12e-07	13.4639	4.34e-07	7.7562
	5	1.31e-06	1.27e-06	3.5698	1.24e-06	6.1883	1.24e-06	5.7892	1.27e-06	3.2559
	6	3.44e-21	2.16e-24	1.59e+05	1.49e-24	2.30e+05	1.49e-24	2.31e+05	8.62e-25	3.98e+05
	7	3.35e-07	3.37e-07	0.6052	3.17e-07	5.5082	3.19e-07	5.0381	3.35e-07	0.0487
	8	3.41e-24	2.14e-11	100.0000	3.27e-11	100.0000	3.27e-11	100.0000	1.79e-23	80.9238
	9	2.64e-06	2.58e-06	2.2999	2.54e-06	3.9937	2.54e-06	4.1125	2.72e-06	2.7800
Angle	1	0	0	0	0	0	0	0	0	0
	2	7.03e-06	7.01e-06	0.2976	6.75e-06	4.2307	6.75e-06	4.1912	7.10e-06	0.9347
	3	5.75e-06	5.79e-06	0.6624	5.44e-06	5.6448	5.45e-06	5.5091	5.95e-06	3.4661
	4	1.29e-06	1.27e-06	1.8115	1.24e-06	4.0709	1.26e-06	2.6671	1.34e-06	3.5568
	5	5.34e-06	5.27e-06	1.3313	5.44e-06	1.8603	5.49e-06	2.7997	5.77e-06	7.5224
	6	4.97e-06	4.97e-06	0.0331	4.74e-06	4.9220	4.74e-06	4.9040	5.24e-06	5.0701
	7	1.18e-05	1.17e-05	1.1821	1.12e-05	4.9569	1.12e-05	5.1159	1.22e-05	3.4951
	8	5.21e-06	5.15e-06	1.2895	4.98e-06	4.6728	4.96e-06	5.1172	5.29e-06	1.4650
	9	8.08e-06	7.98e-06	1.2664	7.70e-06	4.9402	7.77e-06	4.0183	8.35e-06	3.1585
Active Power	1	5.53e-04	5.44e-04	1.5730	5.32e-04	3.7969	5.40e-04	2.3757	5.74e-04	3.7552
	2	8.21e-04	8.08e-04	1.6856	7.91e-04	3.7606	8.03e-04	2.3306	8.54e-04	3.8708
	3	3.97e-04	3.91e-04	1.6853	3.83e-04	3.6990	3.89e-04	2.2669	4.14e-04	3.9061
Reactive Power	1	1.82e-04	1.94e-04	5.8578	1.93e-04	5.2883	1.93e-04	5.3287	2.10e-04	13.2946
	2	2.06e-04	2.03e-04	1.2402	2.06e-04	0.0834	2.08e-04	0.9149	2.11e-04	2.3349
	3	8.04e-05	7.94e-05	1.2341	7.66e-05	4.8607	7.80e-05	3.0573	7.84e-05	2.5273

Of particular interest is the fact that the angle at bus 1 has zero variance for all cases. This phenomenon results from the fact that bus 1 is used as the angle reference bus and is fixed at precisely zero.

In most cases, the percent difference between the MCSs and the cumulant method results is less than 6%. The notable exception to this statement is the percent difference in variances for voltage variables, which is substantially higher, in general, than other system variables. In particular, the variance in bus voltages at buses 1, 4, 6, and 8 is much higher than 6%. Although the percent difference is much higher for voltage, the difference in the actual variance value is not. The worst absolute percent difference for all variance results occurred for the voltage magnitude at bus 6. However, the actual value of the variance for this bus voltage is extremely small, i.e., less than 10^{-20}. Therefore, this variable is almost deterministic, compared to others in the problem, and can be treated as such without any significant loss in statistical information.

3) 118-Bus System: For the 118-bus system, the results are aggregated since the number of buses and variables is too high to present them individually. The results presented in Table IV provide an illustration of how the proposed cumulant performs when applied to larger systems. Results have been tabulated in terms of mean and variance values since the number of variables in the system exceeds 300. The cumulant method is compared against a MCS consisting of 1500 samples.

The column in Table IV labeled *MPE* denotes the *mean percent error*, that is, the average error with the sign considered. In contrast, the column labeled *MAPE (mean absolute percent error)* takes the absolute value of the individual percent errors prior to computing the average. As in the nine-bus system and discussed in Section VI-A2, several variables are effectively deterministic and are, therefore, not analyzed as probabilistic. In general, the difference in the absolute mean values for system variables are less than 1.5% compared to the MCSs. The difference in the variances is between 2% and 8.01%. Fig. 1 shows the PDF for the objective function in the 118-bus problem. In this problem, the objective is a linear active power generation cost function.

TABLE IV
118-BUS SYSTEM GAUSSIAN DISTRIBUTION RESULTS SUMMARY: MPE—MEAN PERCENT ERROR, MAPE—MEAN ABSOLUTE PERCENT ERROR

	Mean		Variance	
Variable	MPE	MAPE	MPE	MAPE
Voltage	0.0003	0.0016	-0.4833	2.5620
Angle	-0.3298	0.5723	-2.0478	3.5601
Active Power	0.4564	0.4836	-4.2630	4.2630
Reactive Power	-0.7698	1.3610	-1.3617	8.0105

B. Gamma Distributions

Results included in this section are based on Gamma-distributed random loads with the mean at the nominal load and the variance 15% of the nominal value. MCSs are performed with 1000 samples, and these results are taken as the reference solution.

Reconstructions using the Gram–Charlier A series can be performed with any number of cumulants. Results are presented here for solutions using up to fifth- and ninth-order cumulants. As the cumulant order increases, the computational expense for the reconstruction also increases.

One of the benefits of the proposed algorithm is a substantial reduction in computational expense while maintaining a high level of accuracy. Table V demonstrates the difference in computation time for several different algorithms. Based on the re-

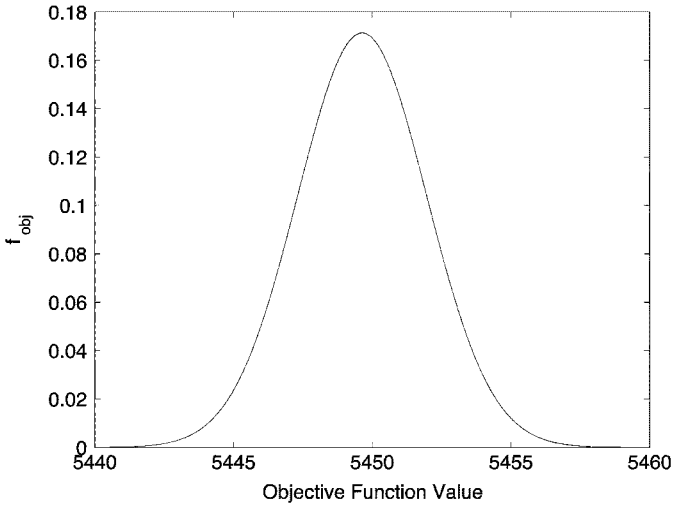

Fig. 1. PDF for 118-bus system.

TABLE V
COMPUTATIONAL EXPENSE WITH GAMMA DISTRIBUTIONS

Algorithm	Time (s)	
	Fifth Order	Ninth Order
Monte Carlo	674.9	704.5
Proposed Method	76.4	92.5

sults presented, the proposed algorithm is several times faster than an MCS.

MCS run times are highly dependent on the number of points used. In this case, only 1000 samples are taken, and in general, a higher number of points results in better solutions. Comparatively, the proposed methodology depends only on the cumulant order used. Generally, a higher number of cumulants should result in better approximations. However, additional complications can occur as order increases including negative probabilities in the tail sections and multimodal distributions [9].

Figs. 2 and 3 show the reconstructed distribution compared with the Monte Carlo results. In these figures, the solid black line is the result using the proposed cumulant method, and the histogram is the raw MCS results. Non-Gaussian behavior is evident in both of the figures, and alternative methodologies that only consider Gaussian distributions would be unable to accurately model the lopsided shape of these distributions.

Of note is the fact that the approximation in Fig. 3 extends far to the right of the MCS results. This occurs because the proposed cumulant method returns a continuous result that does not consider operational limitations in the distributions, although operational limitations are enforced during the optimization process. In this case, the PDF should be truncated on the right-hand side to limit the maximum value for the voltage magnitude.

The assumptions made in the proposed system do not change based on system size. Therefore, larger systems can be analyzed in the same fashion. It is expected that accuracy would remain high, and there will be increasing computational performance gain of the proposed algorithm compared with Monte Carlo. The computational performance will increase because the individual optimizations in the MCS will take longer, in addition to the fact that more simulations will be required to address uncertainty in the sampling process.

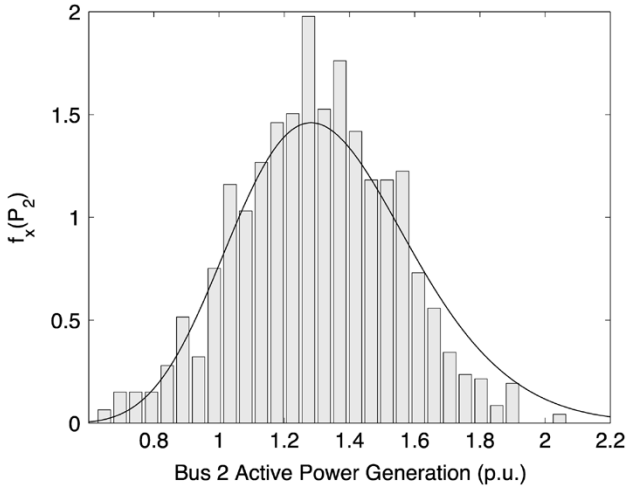

Fig. 2. Bus 2 Active Power Generation—Fifth-Order Reconstruction. Histogram—Monte Carlo, Line—Cumulant Method.

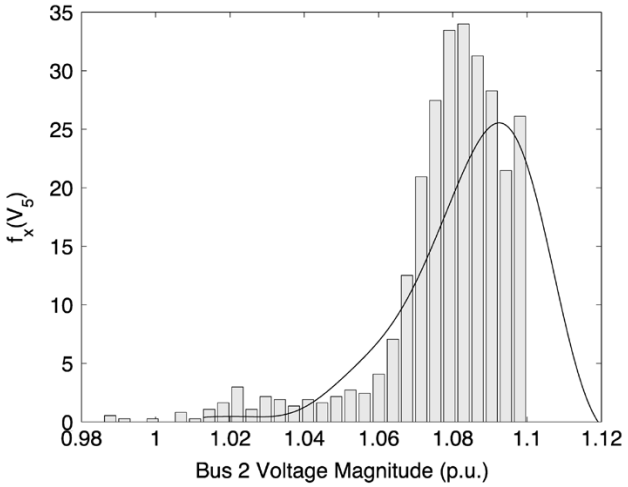

Fig. 3. Bus 5 Voltage Magnitude—Fifth-Order Reconstruction. Histogram—Monte Carlo, Line—Cumulant Method.

VII. CONCLUSION

This paper introduces the cumulant method for the P-OPF problem. By noting that the inverse of the Hessian used in the logarithmic barrier interior point can be used to perform linear mapping, cumulants can be computed for unknown system variables.

Results using the cumulant method are compared against the results from MCS for nine- and 118-bus systems. The proposed algorithm has a substantial reduction in computational expense while maintaining a high level of accuracy.

APPENDIX I
RANDOM VARIABLE DISTRIBUTIONS

The material presented in this paper makes use of two different types of PDFs: Gaussian and Gamma. This appendix

briefly introduces these distributions and some important properties of each. Further information is available in [4] and [13].

A. Gaussian Distributions

The Gaussian Distribution, also known as the Normal Distribution [13], is commonly used in a variety of different areas. It is a simple distribution and is characterized by its mean μ and variance σ^2, according to the following relationship:

$$f_{\mathbf{x}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{36}$$

where $\mathbf{x}$ is the random variable.

The *standard unit normal distribution* is defined as the Gaussian distribution with zero mean and unit variance. In a Gaussian distribution, the peak of the PDF always occurs at the mean value.

B. Gamma Distributions

Another frequently used distribution is known as the Gamma distribution. This distribution is characterized by three variables: the random variable in addition to two non-negative shape parameters.

The general formula for the Gamma distribution is as follows [4]:

$$f_{\mathbf{x}}(x) = \begin{cases} \dfrac{x^{\alpha-1}}{\Gamma(\alpha)\beta^{\alpha}} e^{-x/\beta}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{37}$$

where $\mathbf{x}$ is the random variable, and α and β are the shape parameters. The notation $\Gamma(\alpha)$ stands for the complete Gamma function, which can be written in the following manner [4]:

$$\Gamma(\alpha) = \int_0^{\inf} x^{\alpha-1} e^{-x} dx. \tag{38}$$

In the case that α is an integer, the complete Gamma function can be written in the following simplified form [4]:

$$\Gamma(\alpha) = (\alpha - 1)! \tag{39}$$

The Gamma distribution has several important properties. First, the shape is controlled by two independent non-negative parameters. The values for these two parameters strongly affect the shape of the resulting PDF. Second, the PDF can only be nonzero for positive values of the random variable. In other words, the PDF is bounded on one side. Third, the mean value and the peak value of the PDF are generally different.

APPENDIX II
PROBABILITY AND STATISTICS BACKGROUND

This appendix includes an overview of information related to moments and cumulants for PDFs. More detailed information is available in [4].

A. Moments and Cumulants

Moments and cumulants are both measures of a PDF. Since equations related to cumulants are often developed in terms of moments, this section begins with an introduction to moments and then uses these results to develop the necessary cumulant relationships.

The *expected value* of a random variable $\mathbf{x}$ is defined as

$$E[\mathbf{x}] = \int_{-\infty}^{\infty} x f_{\mathbf{x}}(x)\, dx \tag{40}$$

where $f_{\mathbf{x}}(x)$ is the *PDF* of $\mathbf{x}$.

The nth-order raw moment m_n is defined in the following manner:

$$m_n = E[\mathbf{x}^n] = \int_{-\infty}^{\infty} x^n f_{\mathbf{x}}(x)\, dx. \tag{41}$$

It is possible to compute the raw moments through the use of the *moment generating function* $\Phi_{\mathbf{x}}(s)$. Mathematically, this function is stated as [4]

$$\Phi_{\mathbf{x}}(s) = E[e^{s\mathbf{x}}]. \tag{42}$$

The nth raw moment is computed from the moment generating function by taking the nth derivative with respect to s and evaluating at $s = 0$. For example, the third raw moment can be computed as follows:

$$m_3 = \left. \frac{d^3 \Phi_{\mathbf{x}}(s)}{ds^3} \right|_{s=0} \tag{43a}$$

$$= \left. \frac{d^3 E[e^{s\mathbf{x}}]}{ds^3} \right|_{s=0} \tag{43b}$$

$$= E\left[\left. \frac{d^3 e^{s\mathbf{x}}}{ds^3} \right|_{s=0} \right] \tag{43c}$$

$$= E\left[\mathbf{x}^3 e^{s\mathbf{x}} \Big|_{s=0} \right] \tag{43d}$$

$$= E[\mathbf{x}^3]. \tag{43e}$$

The *cumulant generating function*, denoted by $\Psi_{\mathbf{x}}(s)$, is often written in terms of the moment generating function $\Phi_{\mathbf{x}}(s)$, as follows [4]:

$$\Psi_{\mathbf{x}}(s) = \ln \Phi_{\mathbf{x}}(s). \tag{44}$$

The cumulant generating function is employed in the same manner as the moment generating function; successive derivatives are taken with respect to s and evaluated at $s = 0$. The nth cumulant is denoted as λ_n.

B. Illustrative Example

Consider the standard normal distribution based on (36) with zero mean and unit variance. The moment generating function in (42) can be developed using the definition of expected value from (40) as follows:

$$\Phi_{\mathbf{x}}(s) = E[e^{s\mathbf{x}}] \tag{45a}$$

$$= \int_{-\infty}^{\infty} f_{\mathbf{x}}(x) e^{sx} dx \tag{45b}$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} e^{sx} dx \tag{45c}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} e^{sx} dx \tag{45d}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{sx - \frac{1}{2}x^2} dx \tag{45e}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-s)^2} e^{\frac{1}{2}s^2} dx \tag{45f}$$

$$= e^{\frac{1}{2}s^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-s)^2} dx \tag{45g}$$

$$= e^{\frac{1}{2}s^2}. \tag{45h}$$

Taking the result in (45h) and computing the cumulant generating function based on (44) yields

$$\Psi_{\mathbf{x}}(s) = \ln\left[\Phi_{\mathbf{x}}(s)\right] = \ln\left[e^{\frac{1}{2}s^2}\right] = \frac{1}{2}s^2. \tag{46}$$

To compute the nth cumulant, n successive derivatives of the cumulant generating function are taken and evaluated at $s = 0$. Suppose the third cumulant for a standard normal distribution is of interest and needs to be computed. The first three derivatives of (46) with respect to s are taken as follows:

$$\Psi'_{\mathbf{x}}(s) = s \tag{47}$$

$$\Psi''_{\mathbf{x}}(s) = 1 \tag{48}$$

$$\Psi'''_{\mathbf{x}}(s) = 0. \tag{49}$$

The third cumulant λ_3 can be found by evaluating (49) at $s = 0$

$$\lambda_3 = \Psi'''_{\mathbf{x}}(0) = 0. \tag{50}$$

For this example, the third cumulant is zero. The same process can be repeated for any cumulant of interest.

All cumulants of order three and higher are zero in this example since the derivative of zero is zero. In fact, this result is true for any general Gaussian distribution, and all higher order cumulants, third or greater, are zero. Consequently, cumulants can, in some sense, be considered as a measure of the departure from normality.

REFERENCES

[1] M. Huneault and F. Galiana, "A survey of the optimal power flow literature," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 762–770, May 1991.

[2] G. Viviani and G. Heydt, "Stocahstic optimal energy dispatch," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 7, pp. 3221–3228, Jul. 1981.

[3] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. New York: Springer-Verlag, 1997.

[4] A. Papoulis and S. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. New York: McGraw-Hill, 2002.

[5] 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.

[6] W. Tian, D. Sutanto, Y. Lee, and H. Outhred, "Cumulant based probabilistic power system simulation using Laguerre polynomials," *IEEE Trans. Energy Convers.*, vol. 4, no. 4, pp. 567–574, Dec. 1989.

[7] G. Torres and V. Quintana, "An interior-point method for nonlinear optimal power flow using voltage rectangular coordinates," *IEEE Trans. Power Syst.*, vol. 13, no. 4, pp. 1211–1218, Nov. 1998.

[8] R. Zimmerman and D. Gan. (1997) Matpower. Power Systems Engineering Research Center (PSERC). [Online]. Available: http://blackbird.pserc.cornell.edu/matpower/

[9] M. G. Kendall and A. Stuart, *The Advanced Thoery of Statistics*, 4th ed. New York: Macmillan, 1977.

[10] J. B. Calvert, "Hermite polynomials,", http://www.du.edu/~jcalvert/math/hermite.htm, 2004.

[11] G. Astfalk, I. Lustig, R. Marsten, and D. Shanno, "The interior-point method for linear programming," *IEEE Softw.*, vol. 9, no. 4, pp. 61–68, Jul. 1992.

[12] J. D. Weber, "Implementation of a Newton-based optimal power flow into a power system simulation environment," Univ. Illinois at Urbana-Champaign, [Online]. Available: http://www.powerworld.com/Document%20Library/Newton-based%20Optimal%20Power%20Flow.PDF, 1997.

[13] E. W. Weisstein. *Eric Wesstein's World of Mathematics* [Online]. Available: http://mathworld.wolfram.com/

Antony Schellenberg received the B.Sc. degree in electrical engineering from the University of Calgary, Calgary, AB, Canada, in 2002. He is currently a graduate student pursuing the Ph.D. degree at the University of Calgary.

William Rosehart received the Ph.D. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 2001.

He is currently an Assistant Professor in the Department of Electrical and Computer Engineering, University of Calgary, Calgary, AB, Canada.

José Aguado received the Ph.D. degree from the University of Malaga, Malaga, Spain, in 2001.

Currently, he is an Associate Professor in the Department of Electrical Engineering at the University of Malaga.