Multivariate Modeling of Uncharacteristic Harmonics Using Archimedean Copulas

H. Valizadeh Haghi^{*}, *Student Member, IEEE*, M. Tavakoli Bina, *Senior Member, IEEE*, M. A. Golkar Faculty of Electrical Engineering, K. N. Toosi University of Technology Sayyed Khandan, P.O. Box: 16315-1355, Tehran 16314, Iran *E-mail: valizadeh@ieee.org

Abstract— This paper introduces "Copulas" analytical tool for multivariate modeling of stochastic harmonic generation mechanism. Additional stochastic harmonics can be modulated through the power inverters at the point of common coupling (PCC) under unbalanced non-linear loads. The proposed multivariate unbalance modeling via copulas is applied to evaluate aggregate harmonics injection at the PCC. Copulas have become a popular analytical tool in multivariate modeling, where recently has been applied in many fields. Here, the contributions of copulas to Monte Carlo method are described. It is first come up with modeling unbalance of three-phase active and reactive powers at a distribution substation. To introduce a firmer basis for the suggested procedure, the investigation is carried out based on the measured data for pursuing further analysis that is associated with simulating statistical correlation between stochastic harmonics and realistic unbalanced conditions for a static compensator (STATCOM) at the point of common coupling (PCC).

Keywords- Copula; correlation; voltage unbalance; stochastic harmonics.

I. INTRODUCTION

When the studied system is complex and the effects of certain sequences of events are of a particular interest, statistical/stochastic simulation methods can often be the only means of obtaining the solution to the system model. A popular stochastic simulation is the Monte Carlo method in which the simulation process is repeated for different sets of system parameters. The key activity in the Monte Carlo simulation process is the selection of system parameters to obtain sample solutions, which is applied to problems involving random variables with available probability distributions. A sample from a Monte Carlo simulation is similar to a sample of experimental observations. Therefore, the results of these studies can then be used as modeled samples to study mathematical models of realworld systems or statistical studies.

However, one of the main difficulties associated with the application of the analytical methods in probabilistic power system studies is that the random variables are often dependent. This needs using joint probability distribution functions, imposing an additional difficulty in the already complex problems. Therefore, the majority of analytical approaches assume either independence of the random variables or somehow inaccurate dependencies only through the correlations. Using copulas is suitable for applications with multivariate dependency structure such as in Monte Carlo studies [1]. The use of copulas fits well the stochastic modeling of dependent chaotic variables as well as time series in power systems. They can efficiently be used to produce unconventional multivariate distributions for Monte Carlo studies. On the other hand, the use of copulas for modeling purposes includes two straightforward steps; first, modeling the marginal distributions along with their correlation matrix. The second step consists of fitting a proper copula. It should be noted that finding a multivariate distribution and fitting it to the available data could be a hard task, where copula enables this in principle. The use of copulas is practical as some good software packages have already provided its complete implementation (such as [2]-[3]).

Relevant methods to the stochastic uncertainty analysis combine deterministic simulation techniques with stochastic analysis [4]-[14]. Furthermore, uncertainties in parameter values are only considered with major approximations, where it is also neglected the modeling of dependence structures between parameters of an integral system. The importance of these studies can be better appeared when adequate accuracy is also introduced in addressing the parametric uncharacteristic uncertainties related to the operation of power electronic devices in interaction with each other and the network [11].

This paper introduces the use of an integrated deterministic and probabilistic simulation algorithm in which the stochastic dependence structures are modeled, considering the deterministic dependence and interaction of available devices in a system. Therefore, the key point of such an analysis is the modeling of stochastic dependence that can be suitably done using copulas. The main procedures for using copulas are presented. Meanwhile, the statistical modeling of three-phase active and reactive powers containing practical measurement noises and unexpected chaos is considered using a copula function. The obtained modeling using a copula is then used to produce a realistic voltage unbalance condition at the PCC. Using copulas in a Monte Carlo simulation, it is estimated the correlation between uncharacteristic current THD produced by a STATCOM and the levels of exchanged reactive power.

II. EMPLOYED COPULA ALGORITHM

Copulas are introduced in [1] as "functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions".



Figure 1. Bivariate representation for some of Archimedean and Elliptical copulas. The chosen shape parameters are identical.

The construction of a copula based on a set of arbitrary marginal distribution functions can be assumed and therefore, the *C* defines a multivariate distribution function evaluated at $x_1, x_2, ..., x_n$ as [15]:

$$C[F_1(x_1), F_2(x_2), \dots, F_n(x_n)] = F(x_1, x_2, \dots, x_n)$$
(1)

It is shown in [16] that any multivariate distribution function F can be introduced in the form of Copula representation (1). Further, if the marginal distributions are continuous, then the copula representation will be unique. The aforementioned copulas representation along with their uniqueness is referred to as Sklar's theorem, clarifying the relations between the dependence and the copula for a distribution function. It should be noted that constructing multivariate distributions without the concept of copula has some drawbacks such as:

- Different families are necessarily needed for different marginal distributions.
- Extension of multivariate distributions to thrivariate cases and above is unclear.
- Measures of dependence often appear in the marginal distributions.

A. Modeling the Stochastic Dependence

There are various applications of power systems in which dependent random vectors and arrangements are simulated as evidenced by Monte Carlo algorithms. Examples of such applications are the noise modeling, reliability studies, materials and natural phenomena uncertainty, risk assessment, complex modeling and etc. Behavior of random variables in such circumstances may be assumed completely dependent, linearly correlated, superposed or completely independent. Choosing the most appropriate behavior is influenced by several factors such as the characteristics of the system under study and the required accuracy. Nevertheless, many cases in power system applications are highly dependent. Hence, copula approach is generally managed as follows:

- Estimate matrix of pairwise rank correlations,
- Estimate marginal distributions,
- Combine this information using a copula.

To perform a simulation, therefore, the following information should be specified from the measured or calculated data:

- The copula family and any required shape parameters,
- The rank correlations among random variables,
- The marginal distributions for each random variable.

The most commonly used copulas are the Gaussian copula for linear correlation, Gumbel copula for extreme distributions, and the Archimedean copula and the t-copula for dependence in tail [1], [17]. Contour plots of some commonly used Copulas are depicted in Fig. 1 for the bivariate case, including Frank, Clayton, Gumbel and Guassian.

Furthermore, a realistic correlation matrix must be positive semi-definite, real-valued and symmetric. An important stage of the algorithm is the proper modeling of such a matrix based on realistic data; specifically when there is some data misplacement or the values are noisy, unavailable or unreliable. While there are practical methods to deal with the problems associated with the improper correlation matrices [18], these methods are generally applied to abnormal data recordings and become redundant in Gaussian copula construction algorithms.

Building up the marginal distributions is another key point in a reliable dependency modeling. One could fit a parametric model separately to each dataset, and use those estimates as the marginal distributions. Since a parametric model may not be sufficiently flexible, it might be appropriate to link the marginal distributions by means of a nonparametric model. Meanwhile, using empirical cumulative distributions results in a discrete representation that may be undesirable for a continuous distribution. Therefore, it is common to apply a smoothing technique such as kernel smoothing or to interpolate with a piecewise linear function [3].

B. Simulation Algorithm and the Copula Function

It is preferred to simulate experiments with different copulas and correlations. Two main simulation strategies are the Archimedean and compounding methods [19]. Both methods can be easily implemented for more than two dimensions (multivariate case). Nonetheless, the compounding algorithms are computationally more straightforward than the conditional distribution approach used in Archimedean methods. Meanwhile, it requires generation of an additional variable which sharply increases the needed computations. Because the power system problems typically require extensive calculations, addition of extra variables may accentuate the complexity of the analysis. Therefore, like most statistical packages (e.g. [2]), the Archimedean construction is used in this paper. One method is briefly described as follows [19]:

- 1) Generate independent uniform random numbers U_1 , U_2 , ..., U_n .
- 2) Set $X_1 = F_1^{-1}(U_1)$ and $c_0 = 0$.

3) Calculate recursively X_k using

$$U_{k} = F_{k}(X_{k} | x_{1}, \dots, x_{k-1}) = \frac{\Phi^{-1(k-1)} \{c_{k-1} + \Phi[F_{k}(x_{k})]\}}{\Phi^{-1(k-1)}(c_{k-1})}, \quad k = 2, 3, \dots, n \quad (2)$$

This algorithm generates $X_1, X_2, ..., X_n$, that can be represented by distribution functions like that of (1), the copula is

$$\begin{cases} C(u_1, u_2, \dots, u_n) = \Phi^{-1} [\Phi(u_1) + \dots + \Phi(u_n)] \\ c_k = \Phi [F_1(x_1)] + \dots + \Phi [F_k(x_k)] \end{cases}$$
(3)

Equation (3) defines a class of copulas known as Archimedean, which allows turning a multivariate copula into a single univariate function. The function Φ uniquely generates the copula [20].

Meanwhile, choosing a copula to fit the available data is an important but difficult task [21]. Since the real data generation mechanism is unknown, it is possible that several candidate copulas either fit the data reasonably well or not. When the maximum likelihood method is used, the general practice will be fitting the data with all candidate copulas and choose the ones with the highest likelihood [22]. The Kendall's process is a graphical-based tool for choosing a function among Archimedean copulas [23]. However, selection of a proper copula is an ongoing research area.

Considering the maximum likelihood, the Frank copula is chosen in this paper because it fits the studied data well (section III). Therefore, this copula is explained in detail. Other examples of possible types of copulas and their characteristics are presented in [19,21]. The Frank copula is a symmetric Archimedean copula with the following generator

$$\Phi(t) = \ln \frac{e^{\alpha t} - 1}{e^{\alpha} - 1} \tag{4}$$

Where α is the dependence parameter ($\alpha \in \mathbb{R}$). There is a one-to-one correspondence between each correlation measure and the dependence parameter α . This relationship for the bivariate case is

$$\begin{cases} \rho_s = 1 - \frac{12}{\alpha} [D_2(-\alpha) - D_1(-\alpha)] \\ \tau = 1 - \frac{4}{\alpha} [D_1(-\alpha) - 1] \end{cases}$$
(5)

Where the $D_k(.)$ is the so-called Debye functions defined as

$$D_k(x) = \frac{k}{x^k} \int_0^x \frac{t^k}{e^t - 1} dt, \quad k = 1, 2.$$
 (6)

It should be mentioned that Frank copula permits negative as well as positive dependence. Nonetheless, other types of Archimedean copulas permit only non-negative correlations because of the limited dependence parameter space. This is another reason for using Frank copula in the following section.

The above technique for random vectors can be applied to time series as well [24]. A moving window with a certain number of vectors is taken as a sample vector for a stationary time series. The marginal distributions and the copula are then estimated with this sample according to the described Archimedean algorithm.

In the following section two novel applications are demonstrated. First, the recorded three-phase active and reactive powers at a distribution substation in Tehran and their stochastic dependence are realistically modeled using a copula. Second, a Monte Carlo simulation is applied to a STATCOM for estimating the correlation between uncharacteristic harmonic distortion, levels of unbalance, and the exchanged reactive power at the PCC.

III. STOCHASTIC MULTIVARITE MODELING OF HARMONICS OF STATCOM

A. Modeling Three-phase Active and Reactive Powers

This study considers statistical simulation of three-phase active and reactive powers. Both advantages and applicability of the copula approach is demonstrated, modeling the dependence structures in power system context. To validate the proposed approach, a data logger is installed at the distribution substation (namely Alestom 20 kV/400 V, 1 MVA) that is located in north-west of Tehran.

A 15000-sample Monte Carlo simulation is arranged to get approximations over the measured data. The marginal distributions related to three-phase active and reactive power should be separately modeled in the proposed copula algorithm (section II.B). The next step is to find out the dependence structure between active and reactive powers, between the phases and between the marginal distributions. This is necessary to implement the proposed method using copulas. This dependence is interpreted in the form of a 6×6 rank correlation matrix that is obtained from a positive semi-definite Spearman's rank correlation. Considering the simulation procedure, proposed in section II.B, the three-phase active and reactive powers are simulated. The Frank copula is selected, as mentioned in section II.B, and adjusted to fit the data using the maximum likelihood method. Simulations are shown in Fig. 2(a) using scatter plots in a form that properly demonstrates the correlations too. Figure 2(b) shows the scatter plots for the measured powers. To verify the copula approach, Fig. 2(a) can be compared with Fig. 2(b) that the copula approach successfully models both the marginal distributions and the dependence structure of the real data. A goodness-of-fit test could also be used to quantitatively verify the adequacy of the presented modeling; however, the visual verification is completely adequate for our objective.

Dealing with the exact data under unbalanced conditions can impose an uncharacteristic and stochastic behavior to the system, devices, or operators. For example, stochastic harmonics can be imposed to power systems by power electronic switching devices under unbalanced applied voltages [9], [11]. The following section investigates the uncharacteristic harmonics that is generated by a voltage source converter-based STATCOM under realistic unbalanced applied voltages. It should be noted that copula analysis of the unbalanced condition takes into account the dependence structure between the uncharacteristic uncertainties in the system. This study enables the designer to evaluate the performance of the device with realistic uncertainties.



Figure 2. Scatter plots showing the dependence structures of the three-phase active and reactive powers; (a) Simulated samples from the proposed copula approach; (b) Measured data shown for comparison.

B. Stochastic Dependency Between Uncharacteristic Harmonics of a VSC and the Exchanged Reactive Power

In practice, VSC-based applications (e.g. STATCOM) impose harmonics to power systems. Interaction of harmonic distortions as well as the unbalance of the grid network could also produce additional frequency components. Deterministic analysis and full assessment of these interactions would be complex in terms of both steady-states and dynamic behavior of harmonic levels [9]. Therefore, it is required to pursue evaluation of these harmonic interactions through a proper combination of both measurements and statistical simulation studies. This paper presents a primary methodology for analyzing such interactions. To demonstrate the method using a practical application. a STATCOM is considered in which selective harmonic elimination (SHE) modulation technique is used for reactive power compensation. The SHE techniques use pre-calculated switching instants under ideal switching conditions along with fixed DC-link voltage, which give several advantages compared to the conventional carrier-based PWM schemes [25].

Considering the SHE, the pre-calculated chopping angles will not be optimal under distorted or unbalanced load conditions in which both DC and AC sides experiencing additional uncharacteristic harmonics emerged on the VSC. Hence, the amount of uncharacteristic harmonics that is injected to the distribution system stochastically depends on several factors such as operating conditions of the VSC [9].

The evaluation of this stochastic dependency should embrace all realistic uncertainties in the network. It should also be taken into consideration the high sensitivity of the produced uncharacteristic harmonics toward the error in the optimal chopping angles. In this study, a Monte Carlo simulation is proposed based on the described Copula approach. The emphasis is on the estimation of the level of dependence among a realistic voltage unbalance of the grid network, the produced uncharacteristic harmonics, and the reactive power operating point.

The harmonic-domain model and topology of [9] is used

here in order to provide the required calculation efficiency and accuracy. It should be noted that the impedance of the converter became a two-dimensional tensor in the harmonic-domain; when we try to model the non-linearity caused by the switching activity of the converter. The AC system on the high voltage side of the transformer is rated at 20 kV. Five chopping angles are used to modulate the output voltage of the VSC in steady state. The objective function of the SHE is arranged in a way that the fundamental component is regulated, eliminating the fifth, seventh, eleventh, and thirteenth harmonics. The proposed method is described as follows.

First, the three-phase active and reactive powers data, as simulated in Fig. 2 (a), are used to estimate the statistical behavior of the voltages with realistic unbalance by a three-phase load flow analysis. The estimated result is shown in Fig. 3 which demonstrates a good agreement with the measured data. The obtained accuracy is due to the fact that the dependence structure of the three-phase active and reactive powers is taken into account using a suitable copula. Furthermore, background harmonics are represented by harmonic voltage source obtained from the field measurements. Assume the capacitance of the DC-link is 2 mF, and then the resultant uncharacteristic current THD% is calculated for all data as shown in Figs. 4 and 5. Fig. 4 is a surface plot in which variation of the THD over various reactive power loadings is depicted for each observed voltage unbalance factor (VUF) defined by the IEC.



Figure 3. Recorded vs. simulated VUF% considering stochastic dependency.



Figure 4. Statistical simulation of the penetrated uncharacteristic current THD based on the measured unbalanced powers by a surface plot.



Figure 5. The information of Fig. 4 by a contour plot.

It is noticeable that the diversity of various probable VUF (causing different uncharacteristic THD) is also realistically modeled by the proposed procedure. Fig. 5 shows the same data by a contour plot in which the magnitudes of the THD are depicted by different colors. This can provide a sense that in what and how much area the THD is higher than the normal levels. It can be seen from Fig. 4 and 5 that a realistic voltage unbalance would not dramatically modifies the uncharacteristic THD of AC current coming out of VSC except for situations in which the compensator absorbs relatively small amounts of reactive power. Nonetheless, there are rather few observed VUF for which the uncharacteristic THD are significant.

To clarify the estimated stochastic dependency further, a pattern of variation for the correlation coefficient between the uncharacteristic current THD and the VUF is shown in Fig. 6. This correlation is about zero for very small reactive power loadings; conversely, it arises for higher amounts of reactive power absorption or generation.



Figure 6. Variation pattern of the correlation between uncharacteristic current THD and the exchanged reactive power of the VSC.



Figure 7. Distributions of the current THD corresponding to the reactive power loadings of -0.8 P.U., 0.9 P.U., 0.04 P.U. and -0.1 P.U., respectively.

To provide another representation of the obtained results, Figs. 7 (a)-(d) compares the probability distributions of the current THD for realistic VUF% when four different reactive powers are supplied by STATCOM. It can be seen from Figs. 7 (a)-(d) that reactive powers of -0.8 P.U. and 0.9 P.U. along with realistic VUF% give a relatively small range of the THD variations less than 1%. However, reactive powers equal to 0.04 and -0.1 P.U. along with the same realistic VUF% result in a considerable range of the THD variations around 55% and 25%, respectively.

This analysis can be easily extended to include other realistic conditions. In this manner, a purely deterministic analysis will become more complex and the proposed stochastic simulation would be more useful. All of the studied cases show that correlating dependent power system variables under copula modeling can be combined with conventional Monte Carlo simulations to get more realistic results. Other applications of the presented method are also possible where a stochastic multivariate uncertainty exists. Considering the rapid movement of passive power systems towards highly active topologies forming smarter grids, it seems that the need for modeling of such uncertainties will increase. In fact, both the technical problems such as interactions between converters and the grid network in the new interconnected active system, and the financial risks, requires an adequate representation and modeling of uncertainty in a multivariate context. This could be efficiently engaged by using copulas, as it is shown for a case study in this paper.

IV. CONCLUSIONS

This paper suggests a novel analysis related to the deterministic-stochastic dependencies in power system devices, emphasizing on power electronic switching interaction with the grid network. In order to model stochastic dependencies in a multivariate Monte Carlo simulation, the copula theory has been proposed and briefly introduced. Two case studies are arranged based on the recorded measured data from a 20 kV/400 V distribution substation located in Tehran for a one week period. First, the complete dependence structure of the three-phase active and reactive powers is modeled using a copula. Verifying the modeled structure with the measured data as well as comparing with a jointly normal distribution, the proposed method demonstrated itself having useful characteristics with a sufficient accuracy. Second, a Monte Carlo simulation is implemented based on the described copula approach to estimate the level of dependency between a realistic voltage unbalance of the grid network, the produced uncharacteristic harmonics, and the reactive power loadings. The presented results are used to evaluate harmonic performance of a VSCbased compensator. Future work can focus on other possible applications, extending to include other realistic conditions.

REFERENCES

- [1] R. B. Nelsen, An Introduction to Copulas, 2nd ed., Springer, New York, 2006.
- [2] R Development Core Team, 2006a. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (Online) http://www.R-project.org/, 2009.
- [3] The Math Works, MATLAB, http://www.mathworks.com/, 2009.
- [4] R.N. Allan, A.M.L. da Silva, R.C. Burchett, "Evaluation methods and accuracy in probabilistic load flow solutions," *IEEE Trans. Power App. Syst.*, vol. 100, no. 5, 1981, pp. 2539–2546.
- [5] G. Papaefthymiou, P.H. Schavemaker, L. van der Sluis, W.L. Kling, D. Kurowicka, R.M. Cooke, "Integration of stochastic generation in power systems," *Int. J. Electr. Power Energy Syst.*, vol. 28, 2006, pp. 655–667.
- [6] N. Locci, C. Muscas, L. Peretto, and R. Sasdelli, "A numerical approach to the evaluation of uncertainty in nonconventional measurements on power systems," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 4, 2002, pp. 734–739.

- [7] P. Pinson, "Estimation of the uncertainty in wind power forecasting," Ph.D. dissertation, Ecole des Mines de Paris, Paris, France, 2006. http://pastel.paristech.org/. 2009.
- [8] A.M. Gonzalez, A.M. San Roque, J. Garcia-Gonzalez, "Modeling and forecasting electricity prices with input/output hidden Markov models," *IEEE Trans. on Power Syst.*, vol. 20, no. 1, 2005, pp. 13–24.
- [9] H.V. Haghi, M.T. Bina, "Complete harmonic-domain modeling and performance evaluation of an optimal-PWM-modulated STATCOM in a realistic distribution network," *Prz. Elektrotech.*, vol. 85, no. 1, 2009, pp. 156–161.
- [10] M.T. Au, J.V. Milanović, "Development of stochastic aggregate harmonic load model based on field measurements," *IEEE Trans. Power Deliv.*, vol. 22, no. 1, 2007, pp. 323–330.
- [11] H.V. Haghi, M.T. Bina, "A study on probabilistic evaluation of harmonic levels produced by static compensators," in: *Proceedings of the 2008 AUPEC*, 2008, pp. 1–6.
- [12] M.T. Schilling, A.M.L. da Silva, R. Billinton, M.A. El-Kady, "Bibliography on power system probabilistic analysis (1962–1988)," *IEEE Trans. Power Syst.*, vol. 5, 1990, pp. 1–11.
- [13] H.R. Enslin, P.J.M. Heskes, "Harmonic interaction between a large number of distributed power inverters and the distribution network," *IEEE Trans. Power Electron.*, vol. 19, 2004, pp. 1586–1593.
- [14] Y.-J. Wang, L. Pierrat, "A method integrating deterministic and stochastic approaches for the simulation of voltage unbalance in electric power distribution networks," *IEEE Trans. Power Syst.*, vol. 16, no. 2, 2001, pp. 241–245.
- [15] D. Pfeifer, J. Neslehova, "Modeling dependence in finance and insurance: the copula approach," *Blatter der DGVFM*, vol. 26, no. 2, 2003, pp. 177–191, http://www.staff.uni-oldenburg.de/dietmar.pfeifer/ Copulas.pdf.
- [16] A. Sklar, "Functions de repartitiona n dimensions et leurs marges," Pub. Inst. Statist. Univ, Paris, vol. 8, 1959, pp. 229–230.
- [17] P. Embrechts, A. McNeil, D. Straumann, "Correlation and dependence in risk management: properties and pitfalls," *Risk Manag.*, 2002, pp. 176–223.
- [18] R. Rebonato, P. Jackel, "The most general methodology to create a valid correlation matrix for risk management and option pricing purposes," *J. Risk*, vol. 2, 2002, pp. 17–27.
- [19] E.W. Frees, E.A. Valdez, "Understanding relationships using copulas," in: Proceedings of 32nd Actuarial Research Conference, 1997, pp. 76– 88.
- [20] C. Genest, J. MacKay, "The joy of copulas: Bivariate distributions with uniform marginals," *Amer. Statistician*, vol. 40, 1986, pp. 280–283.
- [21] J. Yan, "Enjoy the joy of copulas," J. Statistical Softw., vol. 21, no. 4, 2007, pp. 1–2.
- [22] E. W. Frees, P. Wang, "Credibility using copulas," N. Amer. Actuarial J., vol. 9, no. 2, 2005, pp. 31–48.
- [23] C. Genest, L. P. Rivest, "Statistical inference procedures for bivariate Archimedean copulas," J. Amer. Statistical Assoc., vol. 88, 1993, pp. 1034–1043.
- [24] J.C. Strelen, F. Nassaj, "Ananlysis and generation of random vectors with copulas," in: *Proceedings of the 2007 Winter Simulation Conference*, 2007, pp. 488–496.
- [25] P.N. Enjeti, P.D. Ziogas, J.F. Lindsay, "Programmed PWM techniques to eliminate harmonics: a critical evaluation," *IEEE Trans. Ind. Appl.*, vol. 26, 1990, pp. 302–316.