International Conference on Smart Infrastructure and Construction 2019 (ICSIC):
Driving data-informed decision-making

ICE Publishing

DeJong, Schooling and Viggiani
ISBN 978-0-7277-6466-9
https://doi.org/10.1680/icsic.64669.387


CSIC Cambridge Centre for Smart Infrastructure & Construction

# EFFICIENT STOCHASTIC ANALYSIS OF POWER DISTRIBUTION SYSTEMS USING POLYNOMIAL MODELS

*N. Alemazkoor, H. Meidani**

*Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, USA*
** Corresponding author*

**ABSTRACT** With ever increasing penetration of distributed generation, voltage control has become a new challenge for distribution networks. Conventionally, to control the system, perturb-and-observe or Jacobian matrix inversion approaches are used to evaluate the impact of active and reactive power of distributed generators on voltage magnitudes. However, these approaches require constantly updating the Jacobian matrix or running several power flow simulations for different system states, which can be computationally cumbersome. In this work, we propose developing polynomial surrogates that approximate the voltage level as a function of power consumption, distributed power generation and power factor of distributed generators. Polynomial surrogates can be efficiently used for probabilistic analysis and control of the distribution network. The effectiveness of the proposed surrogate-based method is demonstrated on the IEEE 33-bus system, in the presence of a large number of correlated random inputs.

## 1. Introduction

Future power distribution networks will involve large penetration of renewable distributed generation (DG) units (Alanne and Saari, 2006). The penetration of DG units can improve power system efficiency and reliability and reduce it's vulnerability as a result of increase in power supply capacity and reduction in transmission loss (Hung and Mithulananthan, 2013). However, along with all advantages of DG integration, it also causes major challenges. Voltage control is one of the main challenges that high penetration level of DGs imposes (Coster et al., 2011). To reliably control voltage, it is necessary to understand how changes in active and reactive power of DGs impact the voltage profile of the network. This can be viewed in the context of sensitivity analysis.

Conventionally, local sensitivity analysis such as inverse of Jacobian matrix (Valverde and Van Cutsem, 2013), and perturb-and-observe approach (Tamp and Ciufo, 2014) are used to evaluate the change in voltage at a specific bus as a result of a change in active and/or reactive power at another bus. The main issue about these approaches is that they do not provide analytical insights on the impact of change in power on the voltage profile. Therefore, one has to constantly monitor the system and update the Jacobian matrix or run several power flow simulations for different system states during the operation. Recently, some studies have investigated the analytical approximation of voltage change. For instance, Weckx et al. (2015) used data to approximate voltage change as a linear function. Simplifying voltage profile as a linear function of system inputs, however, may not be accurate. Zhang et al. (2018) used a surface fitting approach to approximate voltage, which required a very large number of simulations. Additionally, power consumption was considered to be deterministic and not random. However, it is necessary to account for random behavior of consumers and renewable energy sources as it causes random fluctuations in voltage profile of distribution network. Jhala et al. (2017) derived an upperbound for change in voltage at a specific bus as a result of change in active and reactive power at other buses. Although the upperbound was verified using simulations, it may be substantially greater than the actual voltage change and result in misleading voltage analysis.

In this work, we consider power consumption and distributed power generation at all buses to be correlated random variables. We propose a novel approach to approximate voltage as a polynomial function of power consumption, power generation from DGs, and power factor of DGs at each bus. We also consider a scenario, where not all the DGs can be controlled. In this scenario the percentages of generated active power from the DGs that can be controlled at each bus are also considered as explanatory variables in voltage approximation. This approximated function can then be considered as a computationally cheap surrogate for power flow model. Although recently polynomial surrogate has been successfully used in probabilistic power flow analysis (Ni et al., 2017; Ren et al., 2016), it has not yet been used for probabilistic analysis of distribution networks with DGs. We use compressive sampling to approximate the surrogate and show that, with a relatively small number of simulations, a polynomial surrogate can be accurately constructed. Not only does this surrogate give analytical insights on how changes in power at a given bus impact voltages at other buses, but it can also be used for global sensitivity analysis and calculating probability distributions of voltage. Therefore, it can be applied for model predictive control of a distribution network with DGs or for planning and evaluating effect of integrating DGs into a distribution network. We will demonstrate the validity and efficiency of our

proposed surrogate-based probabilistic and sensitivity analysis approach on the IEEE 33-bus system.
This paper is organized as follows. Section 2 presents general concepts and theoretical background in polynomial approximation. In Section 3, we introduce our proposed surrogate-based probabilistic voltage sensitivity approach. Finally, Section 4 includes the tests results and discussion.

## 2. Polynomial Surrogate

This section introduces polynomial chaos expansion (PCE) and explains how compressive sampling can be applied to estimate the coefficients of a PCE.

### *2.1 Polynomial chaos expansion*

To expedite the stochastic computation in the analysis and design of complex systems, analytical surrogates that approximate and replace full-scale simulation models have been extensively used. Polynomial chaos expansion is one of the most widely adopted surrogates, which approximates the system response by a polynomial function in the space of random parameters (Doostan and Owhadi, 2011; Xiu, 2010) .

PC expansion is a spectral technique to approximate stochastic models using orthogonal polynomials of random inputs. Let $I_{\boldsymbol{X}}$ be a tensor-product domain that is the support of $\boldsymbol{X}$, where $\boldsymbol{X} = (X_1, \cdots, X_d)$ is the vector of independent random variables, i.e. $X_i \in I_{\boldsymbol{X_i}}$ and $I_{\boldsymbol{X}} = \times_{i=1}^{d} I_{\boldsymbol{X_i}}$. Also, let $\rho_i$: $I_{\boldsymbol{X_i}} \rightarrow \mathbb{R}^+$ be the probability measure for variable $X_i$ and let $\rho(\boldsymbol{X}) = \prod_{i=1}^{d} \rho_i(X_i)$. Given this setting, the set of univariate orthonormal polynomials, $\{\psi_{\alpha,i}\}_{\alpha\in\mathbb{N}_0}$, satisfies

$$\int_{I_{X_i}} \psi_{\alpha,i}(x_i)\psi_{\beta,i}(x_i)\rho_i(x_i)\, dx_i = \delta_{\alpha\beta}, \ \ \alpha, \beta \in \mathbb{N}_0, \tag{1}$$

where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, and $\delta_{\alpha\beta}$ is the Kronecker delta function. Therefore, the density function of $X_i, \rho_i(X_i)$, determines the type of polynomial. For example, Gaussian and uniform probability distributions enforce Hermite and Legendre polynomials, respectively. The $d$-dimensional orthonormal polynomials are then derived from the multiplication of one dimensional polynomials in all dimensions. For example,

$$\psi_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \psi_{\alpha_1,1}(x_1)\psi_{\alpha_2,2}(x_2) \cdots \psi_{\alpha_d,d}(x_d), \tag{2}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_d)$. Consequently, we have

$$\int_{I_{\boldsymbol{X}}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{x})\psi_{\boldsymbol{\beta}}(\boldsymbol{x})\rho(x)\, d\boldsymbol{x} = \delta_{\boldsymbol{\alpha\beta}}, \quad \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_0^d. \tag{3}$$

Using this construction, any function $u(\boldsymbol{X})$: $I_{\boldsymbol{X}} \rightarrow \mathbb{R}$ that is square-integrable can be represented as

$$u(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha}\in\mathbb{N}_0^d} c_{\boldsymbol{\alpha}}\, \psi_{\boldsymbol{\alpha}}(\boldsymbol{X}), \tag{4}$$

where $\{\psi_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}_0^d}$ is the set of orthonormal basis functions satisfying Equation (3) (Xiu, 2010).

However, for computation's sake, $u(\boldsymbol{X})$ is approximated by a finite order truncation of PC expansion given by

$$u^k(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha}\in\Lambda^{d,k}} c_{\boldsymbol{\alpha}}\, \psi_{\alpha}(\boldsymbol{X}), \tag{5}$$

where $k$ is the total order of the polynomial expansion and $\Lambda^{d,k}$ is the set of multi-indices defined as

$$\Lambda^{d,k} \coloneqq \{\boldsymbol{\alpha} \in \mathbb{N}_0^d \colon \|\boldsymbol{\alpha}\|_1 \leq k\}. \tag{6}$$

The cardinality of $\Lambda^{d,k}$, i.e. the number of expansion terms, here denoted by $K$, is a function of $d$ and $k$ according to

$$K \coloneqq \left|\Lambda^{d,k}\right| = \frac{(k+d)!}{k!d!}. \tag{7}$$

Given this setting, $u^k(\mathbf{X})$ approximates $u(\mathbf{X})$ in a proper sense and is referred to as the $k$-th degree PC approximation of $u(\mathbf{X})$. Regression is widely used to estimate PC coefficients (Shin and Xiu, 2016). For regression, the fundamentals of linear algebra require the number of samples (simulations) to be equal or greater than the number of unknown coefficients, $K$. Motivated by the fact that system response is sparse with respect to PC basis in many high-dimensional problems, compressive sampling has been employed to estimate PC coefficients with significantly smaller number of samples (simulations) than the coefficients (Doostan and Owhadi, 2011). In what follows, a brief background on compressive sampling for PC estimation is provided.

### *2.2 PC estimation using compressive sampling*

Compressive sampling was first introduced in the field of signal processing where conventionally the Shannon-Nyquist sampling rate (Candes and Wakin, 2008) determined the number of samples required to recover a signal. Compressive sampling allows for successful signal recovery using significantly fewer samples when the signal of interest is sparse. Recently, compressive sampling has gained substantial attention in estimating PC coefficients (Doostan and Owhadi, 2011). It is aimed to calculate the vector of unknown expansion coefficients c = $(c_{\boldsymbol{\alpha}^1};\cdots; c_{\boldsymbol{\alpha}^K})$ in Equation (5), given $M$ samples of systems response, denoted by the data vector $\boldsymbol{u} = (u(x^{(1)}), \cdots, u(x^{(M)}))^{\mathrm{T}}$. These sampled outputs are evaluated at the $M$ random realizations of the system input $\boldsymbol{X}$. Requiring $u^k(\boldsymbol{X})$ in Equation (5) to approximate $\mathrm{u}(\boldsymbol{X})$ results in following system of question,

$$\boldsymbol{\psi c} = \boldsymbol{u}, \tag{9}$$

where $\boldsymbol{\psi}$ is the model matrix, constructed according to

$$\boldsymbol{\psi} = [\psi_{ij}], \ \ \psi_{ij} = \psi_{\boldsymbol{\alpha}^j}(x^{(i)}), \ \ 1 \leq i \leq M, 1 \leq j \leq K. \tag{10}$$

We are interested in the underdetermined cases, where $M \ll K$, and infinitely many solutions for $\boldsymbol{c}$ may exist. Compressive sampling approach can be readily borrowed to find the sparsest solution, by formulating the sparse recovery problem as

$$\min_{\boldsymbol{c}} \|\boldsymbol{c}\|_0 \text{ subject to } \boldsymbol{\psi c} = \boldsymbol{u}; \tag{11}$$

where $||\boldsymbol{c}||_0$ indicates the $\ell_0$-norm, i.e. the number of non zero terms. However, since $\ell_0$-norm is non-convex and discontinuous, the above problem is NP-hard. Therefore, $\ell_0$ minimization is usually replaced by its convex relaxation where $\ell_1$-norm of the solution $\boldsymbol{c}$ is minimized instead, i.e.

$$\min_{\boldsymbol{c}} ||\boldsymbol{c}||_1 \text{ subject to } \boldsymbol{\psi c} = \boldsymbol{u}; \tag{12}$$

When $\boldsymbol{\psi}$ is sufficiently incoherent and $\boldsymbol{c}$ is sufficiently sparse, the solution of $\ell_0$ minimization is unique and is equal to the solution of $\ell_1$ minimization (Bruckstein et al, 2009). Several packages are available to solve minimization (12). In this work, we use the 'ConstrainedL1.m' code available by Lou (2017) for solving (12). In next section, we explain how compressive sampling can be applied to develop a surrogate for power flow model in a distribution network with DGs and facilitate probabilistic voltage sensitivity analysis.

## 3. Surrogate-based Probabilistic Voltage Sensitivity Analysis Approach

In this section, we discuss how to consider power factors of distributed generators as explanatory variables in the voltage estimation using PCE and how to address the correlation between random inputs. We also explain how to perform probabilistic voltage sensitivity analysis using the estimated polynomial surrogate.

### *3.1 Power factor of distributed generators and correlated random inputs*

Power factors of distributed generators are one of the main control factors in distribution networks. Therefore, they appear in the optimization or control problems as decision variables whose optimal values are identified. We aim the developed polynomial surrogates to analytically evaluate the probabilistic system conditions given various choices of power factors. To this end, we include power factors as input variables of PC surrogates, in addition to random inputs. Specifically, we treat power factors as input variables uniformly distributed between a minimum and maximum allowable values. In other words, the samples used to construct the polynomial surrogates are uniformly distributed in the space of allowable power factors. It should be noted that once the PCE is trained, for probabilistic analysis, we treat these power factors as control variables, and perform probabilistic analysis at fixed values for power factors.

We consider active power consumption at all buses to be normally distributed and positively correlated. Similarly, active power generation from distributed generators at all buses are considered to be normally distributed and positively correlated. To construct the PCE surrogate for correlated random inputs, common practice is to convert correlated variables to independent random variables using Cholesky decomposition or Copula (Ni et al, (2017), Ren et al, (2016)). However, we found this step in constructing PCE to be unnecessary. It should be noted that the reason for converting correlated input variables to uncorrelated variables is to maintain the orthogonality of PC basis functions. The orthogonality of PC basis is essential to have incoherent model matrix and consequently accurate PCE approximation. To achieve an incoherent model matrix, instead of converting input variables to uncorrelated variables, we generate input samples used in model matrix under the assumption that input variables are uncorrelated. Once the PCE surrogate is constructed, in order to calculate voltage probability distributions and perform sensitivity analysis we evaluate the constructed surrogate at correlated input samples obtained based on the correlation matrix.

### *3.2 Sensitivity analysis*

Sensitivity analysis aims at determining how variability (or uncertainty) in each input impacts the variability (or uncertainty) in system's response. One of the most widely used global sensitivity analysis approaches is the Sobol' method (Sobol, 2001) in which the variance of system's response is decomposed as summation of variances of different terms in the model and sensitivity is evaluated by Sobol' indices. Traditionally, Sobol' indices are calculated using Monte Carlo simulations. However, having constructed a PCE surrogate to replace expensive simulations, Sobol' indices can be calculated with minimal computation cost (Sudret, 2008). To evaluate Sobol' indices using PCE surrogate, we first need to decompose the PCE expansion based on the indices of its terms. Let us define $\nu$ to be a generic index set, $\nu \subset \{1, \cdots, d\}$. We also define $\Lambda_\nu^{d,k}$ to be a set that contains all the multi-indices within $\Lambda^{d,k}$ that have non-zero terms $\alpha_p \neq 0$ only and only if p $\in \nu$:

$$\Lambda_\nu^{d,k} = \{ \alpha \in \Lambda^{d,k} : p \in \nu \leftrightarrow \alpha_p \neq 0, p = 1; \cdots; \text{d} \}. \tag{13}$$

We can now rewrite PCE as the summation of terms that only depend on the input variables $\boldsymbol{X}_\nu$, where $\boldsymbol{X}_\nu$ is the subvector of $\boldsymbol{X}$ and includes the components in $\boldsymbol{X}$ that are labeled by the indices in $\nu$:

$$u^k(\text{X}) = c_0 + \sum_{\nu\subset\{1,\cdots,d\}} u_\nu^k(\boldsymbol{X}_\nu), \tag{14}$$

where,

$$u_\nu^k(\boldsymbol{X}_\nu) = \sum_{\boldsymbol{\alpha}\in\Lambda_\nu^{d,k}} c_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{X}_\nu). \tag{15}$$

Even though the PCE in this work is built under the assumption that input variables are uncorrelated, one must account for correlation when Sobol' indices are calculated. When $\boldsymbol{X}$ is a vector of correlated variables, the variance of $u^k(\boldsymbol{X})$ can be calculated as

$$\text{Var}(u^k(\text{X})) = \sum_{\nu\subset\{1,\cdots,d\}} Cov\,[u_\nu^k(\boldsymbol{X}_\nu), u^k(\boldsymbol{X})], \tag{16}$$

where,

$$\sum_{\nu\subset\{1,\cdots,d\}} Cov\,[u_\nu^k(\boldsymbol{X}_\nu), u^k(\boldsymbol{X})] = \text{Var}\,(u_\nu^k(\boldsymbol{X}_\nu)) + \sum_{\nu\neq\eta} Cov\,[u_\nu^k(\boldsymbol{X}_\nu), u_\eta^k(\boldsymbol{X}_\eta)]. \tag{17}$$

To calculate total covariance-based sensitivity indices $S_\nu^{(cov)}$, structural (physical) sensitivity indices, $S_\nu^{(S)}$, and correlative sensitivity indices, $S_v^{(C)}$, Equations (16) and (17) are normalized as follows (Sudret and Caniou, 2013):

$$S_v^{(cov)} = \frac{Cov\,[u_v^k(\boldsymbol{X}_v), u^k(\boldsymbol{X})]}{\mathrm{Var}(u^k(\boldsymbol{X}))}, \tag{18}$$

$$S_v^{(S)} = \frac{\mathrm{Var}\,(u_v^k(\boldsymbol{X}_v))}{\mathrm{Var}(u^k(\boldsymbol{X}))}, \tag{19}$$

$$S_v^{(C)} = S_v^{(cov)} - S_v^{(S)}\,. \tag{20}$$

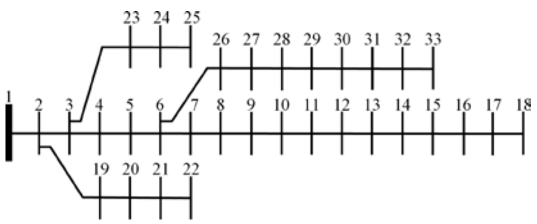
It should be noted that these indices are calculated by merely evaluating the PCE surrogate at sample inputs, a task that incurs minimal computation cost.

Limited research has been focused on global sensitivity analysis for power systems. In (Ni et al, 2018), active power at different buses are considered to be uncertain and global sensitivity analysis are used to rank buses at which active power most influentially impacts QoIs, e.g., voltage and branch currents. However, in distribution systems, voltage control often involves adjusting power factor of distributed generators (Zad et al, 2015). Therefore, it is vital to identify buses at which modifying the power factor has most influential impact on voltage. In this paper, we estimate voltage as a polynomial function of active power consumption and generation, and also power factors of distributed generators. We show that sensitivity analysis can be successfully used to determine at which buses the power factor of distributed generators must be modified to reduce the chance of voltage violation.

## 4. Test Results and Discussion

This section demonstrates the efficiency and validity of our proposed approach on the IEEE 33 bus test system (shown in Figure 1), and shows the surrogate-based analysis results compare with Monte Carlo simulation-based results. The nominal voltage of the system is 12.66 kV and the generator connected to node 1 is set to a voltage of 1.04 per unit.

Fig. 1 IEEE 33 bus distribution system



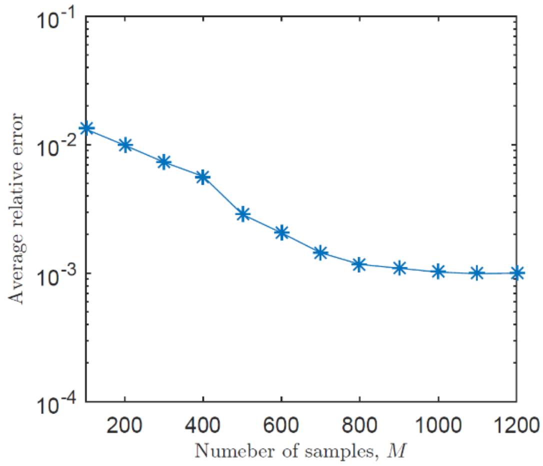

We consider distributed generation to exist at all the buses, with mean active power generation of 0.15 MW and a random power generation level following $N$(0.15, 0.04). We also set the correlation between active power generation at all buses to be 0.8. This is because distributed generators use renewable energy sources such as rooftop photovoltaic and wind generation, which are random and highly spatially correlated. We also consider power consumption at all buses to follow a normal distribution, with their means set to be the nominal values provided with the benchmark IEEE 33 bus system and their standard deviations set to be 10% of their mean values. Similar to the correlation in generation levels, we assume the correlation between active power consumption at all buses to be 0.8. We also consider the allowable range of power factor to be between 0.7 and 1.0 for all buses. It should be noted that these are illustrative assumptions that allows us to generate results and validate our approach without loss of generality. Actual power consumption and generation distributions together with their correlation characteristics can be estimated when data on generation and consumption, weather condition, proximity of buses, etc. are available.

We show the results for two different scenarios. In the first scenario, all DGs can be fully controlled. In the second scenario, we consider a proportion of DGs to be uncontrollable.

### *4.1 Scenario I: All DGs can be controlled*

To solve (12), the model matrix, $\boldsymbol{\psi}$, and the observation vector $\boldsymbol{u}$, must be generated. To this end, polynomial order, $\boldsymbol{k}$, and number of simulation samples, $M$, must be specified. We set the order of polynomial to be $\boldsymbol{k} = \mathbf{2}$, following studies showing that a second-order polynomial surrogate provides sufficient accuracy for estimating different quantity of interests in power systems (Ren et al, 2016). In general, a cross-validation approach can be always used to determine the number of simulation samples and polynomial order. Here, having specified $\boldsymbol{k} = \mathbf{2}$, we seek to determine how many simulation samples are needed using cross-validation. In particular, we calculate the PCE coefficients using different sample sizes and compare the resulting average relative validation error. This relative validation error, given an exact (test) data and a PCE result, is calculated as $\frac{1}{n}\sum_{i=1}^{n}\{\|\boldsymbol{u}_i - \boldsymbol{u}_i^{PCE}\|_2^2/\|\boldsymbol{u}_i\|_2^2\}$, where $\boldsymbol{u}_i$ and $\boldsymbol{u}_i^{PCE}$ are the vectors for exact and predicted voltage value at bus $i$, respectively, and $n$ is the number of buses. For each sample size, we calculate the average of this relative validation error over 100 exact (test) data points, obtained independently. Figure 2 shows this error measure versus the number of simulation samples used to construct the PCE surrogate. It can be seen that using even one hundred simulation samples, i.e., deterministic power flow analysis, in the compressive sampling approach, a relatively accurate PCE surrogate can built with an average error of about 0.01. This in fact confirms that voltage response is sparse or nearly sparse with respect to PCE basis function and that the second-order PCE provides sufficient accuracy.

Figure 2 Average relative validation error vs. the number of simulation samples used to construct PCE surrogate




In this work, we used 800 simulation samples to construct PCE surrogates as the rates of decay in the error measure gets noticeably smaller beyond this sample size. The 2nd-order PCE surrogate that is constructed using 800 simulation samples can then be used for voltage prediction and sensitivity analysis. Specifically, Figure 3 compares the probability distributions of voltage at bus 33, as an example of model output, or quantity of interest. In these calculations, the power factors of all distributed generators are set to 0.95. As the reference case, 50,000 Monte Carlo (MC) simulation samples are used to generate the "exact" distribution. The PCE-based "approximate" voltage distribution is obtained by evaluating the surrogate's 2nd-order polynomial function at the same 50,000 input samples that was used in the MC calculation. It can be seen the two distributions are in very close agreement.

Furthermore, Figure 4 compares the voltage violation probabilities, i.e., the probability of voltage levels exceeding 1.05 p.u. or being smaller than 0.95 p.u., for all buses. In this calculation, we still assume the power factor of all distributed generators are fixed at 0.95. Violation probabilities are evaluated based on voltage probability distributions generated using 500,000 MC simulation samples and PCE approximation at those sample points. It can be seen that PCE surrogates provide accurate estimation and the largest deviation from the exact violation probability is less than 1\%.

Figure 3 Probability distribution of voltage at node 33 generated by Monte Carlo simulations vs. approximated probability distribution using PCE surrogate

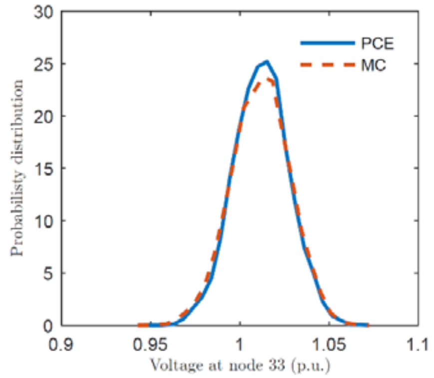
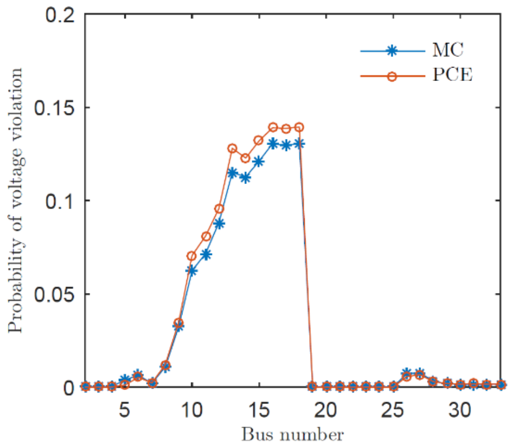
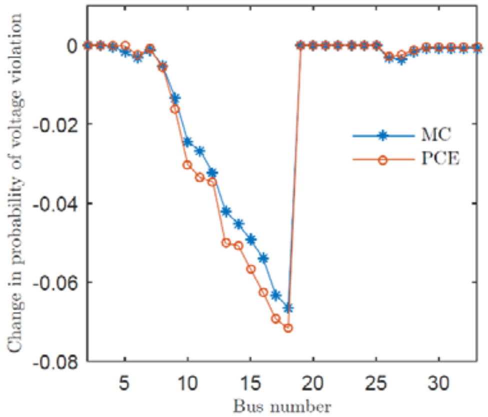



Figure 4 Monte Carlo estimation of voltage violation probability vs. approximated probability of voltage violation using PCE surrogate




Preventing voltage violation is usually done, in practice, by controlling the reactive power of DGs, i.e., by modifying the power factor of DGs. Therefore, it is critical to understand how such power factor modification can impact the probability of voltage violation. Figure 5 compares the change in probabilities of voltage violation calculated using MC simulations and PCE approximation when the power factor of DGs at bus 18 is arbitrarily reduced to 0.75. It can be seen that PCE surrogate results in accurate estimation of change in probability of voltage violation. It should be noted that while only one set of 800 simulation samples is used to construct the PCE surrogate, each time that the power factors are changed a new set of MC simulations is used to analyze the system and evaluate voltage probability distributions. This highlights the computational benefit achieved by constructing a surrogate for distribution system.

Figure 5 Estimated change in probability of voltage violation using Monte Carlo simulations vs. approximated change in probability of voltage violation using PCE surrogate



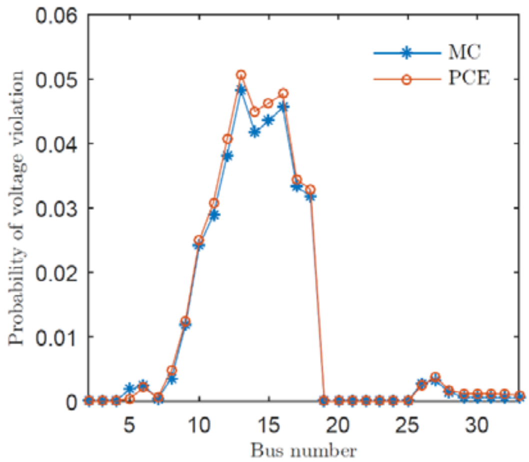

In the case of high probability of voltage violation, it is critical to identify the consumption or generation buses that have the highest impact on voltage conditions. To this end, we propose to use a variance-based sensitivity analysis, discussed earlier. In this work, we calculate Sobol' indices in order to determine which DG power factor(s) (at which bus) should be subject to control. In order to conduct the sensitivity analysis, we first identify the output of interest, that is the critical voltage condition, by revisiting Figure 4. It can be seen that Bus 18 has the highest risk of voltage violation. Therefore, we seek to rank all the buses based on the impact that their DG power factor modifications can have on reducing the voltage violation probability at Bus 18. The second column in Table 1 shows the ranking of top five buses obtained solely by calculating Sobol' indices for the constructed PCE surrogate for Bus 18. Once the ranking is obtained, the PCE surrogates can be used with minimal computational cost to evaluate new distribution of voltage as a result of potential control actions suggested by the ranking. To evaluate the ranking obtained by PCE surrogate, MC simulations and PCE surrogates are used to calculate the change in violation probability at Bus 18 as a result of reducing the power factor of DGs at the target bus, listed in second column of Table 1, to 0.7. The result are shown in third and fourth column of Table 1. It can be seen that the ranking suggested by PCE surrogate is in agreement with the ranking obtained by MC simulation.

Next, we consider the goal of always keeping voltage violation probability less than 0.05 at all buses. In order to do so, we focus on the bus with highest probability of voltage violation and its corresponding list of "influential" buses whose power factor modifications have highest impacts on the voltage violation condition. In order to ensure that voltage violation probability is smaller than 0.05 at all buses, the PC surrogate can be used to approximate new voltage distributions given different modifications. In this case, violation probability smaller than 0.05 at all buses can be achieved by reducing the power factor of DGs at Buses 18 and 17 to 0.7 and 0.85, respectively. These power factor modification levels are obtained using a trial and error scheme. It should be, however, noted that this search is systematic and starts with reducing the power factor of most influential nodes. Moreover, it is computationally cheap as it only involves evaluating an analytical function. On the other hand, the search using MC sampling is not systematic as influential buses are not known and also a large set of samples are required for each trial. Figure 6 shows the voltage violation probabilities (both "approximate" and "exact"), once power factors are reduced to 0.7 and 0.85 at Buses 18 and 17, respectively. It can be seen that PCE surrogate results in a sufficiently accurate approximation of violation probabilities and can be efficiently used to inform decisions on required modification in power factors of DGs.

Table 1 Sensitivity analysis results for Bus 18 when all DGs can be controlled

| | Bus number | Change in violation probability by reducing PF to 0.7 (MC) | Change in violation probability by reducing PF to 0.7 (PCE) |
|---|---|---|---|
| 1st | 18 | -0.076 | -0.081 |
| 2nd | 17 | -0.074 | -0.080 |
| 3rd | 16 | -0.063 | -0.066 |
| 4th | 15 | -0.058 | -0.063 |
| 5th | 14 | -0.053 | -0.057 |

Figure 6 Monte Carlo estimation of voltage violation probability after reducing the power factors of DGs at buses 18 and 17 vs. approximated voltage violation probability using PCE surrogate




*4.2 Scenario II: A subset of DGs can be controlled*

In this scenario we consider a more general case, where a proportion of active power generated at each bus comes from the DGs that are not controllable. We consider the power factor of DGs that are uncontrollable to be one and to be not adjustable. Since all DGs use renewable energy sources, we need to consider the fraction of active power that is generated from controllable DGs to be uncertain. Here as an illustrative assumption, we assign a uniform distribution between 0.5 and 0.7 and a uniform distribution between 0.8 and 1 to the fraction of active power that is generated from controllable DGs at buses with odd and even numbers, respectively.

Similar to Scenario I, as explained in Section 3.1, we generate input samples under the assumption that all variables are uncorrelated to construct the PCE surrogate. Cross-validation is used to select the size of simulation samples, and showed that the PCE surrogates constructed using 1000 samples provide sufficient prediction accuracy. Figure 7 compares the estimated violation probability using MC simulations and PCE surrogate, when power factor of controllable DGs are set to be 0.95. As expected, it's observed that violation probabilities have increased compared to the case when all DGs are controllable and their power factor are adjustable and set to be 0.95 (Figure 4).

To reduce violation probabilities, we first target bus 18, which has the highest violation probability. We then use Sobol' indices to rank the buses based on the impact of the power factor of their controllable DGs on voltage distribution at bus 18. Table 2 shows this ranking. To verify it, the change in violation probability at bus 18 as a result of reducing the power factor of controllable DGs at each bus in the second column to 0.7 is calculated using MC simulations and PCE surrogate. It can be seen that the change in violation probability follows the ranking in the first column. This ranking shows that controllable DGs at which buses must be prioritized for reducing their power factor. PCE surrogate can be used to estimate the new violation probabilities as results of potential control actions. A voltage violation probability smaller than 0.05 at bus 18 can be achieved by reducing power factor of controllable DGs at buses 18, 16, and 17 to 0.7. Figure 8 shows the voltage violation probabilities when power factor of controllable DGs at buses 18, 16, and 17 are reduced to 0.7. It can be seen that the PCE surrogate can accurately estimate violation probabilities.

Figure 7 Monte Carlo estimation of voltage violation probability when a proportion of DGs are controllable vs. approximated probability of voltage violation using PCE surrogate

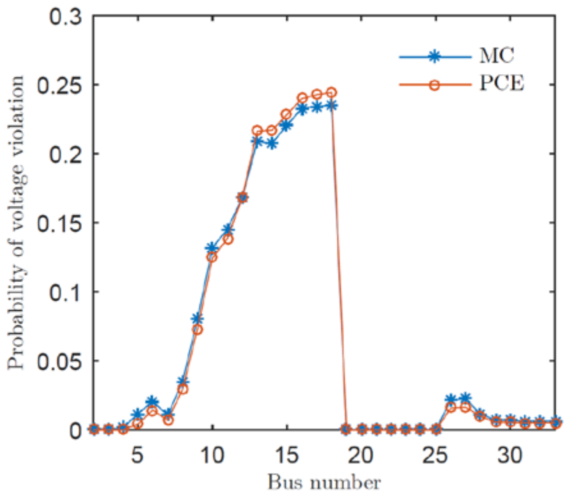
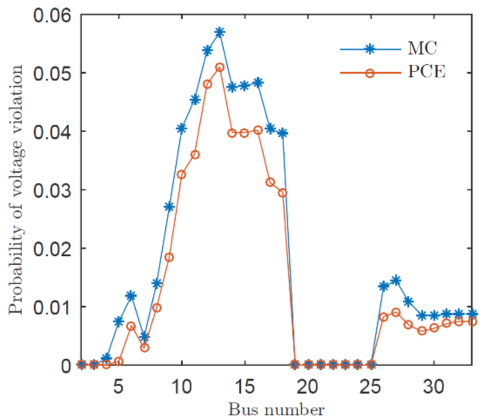



Table 2 Sensitivity analysis results for Bus 18 when a proportion of DGs can be controlled

| | Bus number | Change in violation probability by reducing PF to 0.7 (MC) | Change in violation probability by reducing PF to 0.7 (PCE) |
|---|---|---|---|
| 1st | 18 | -0.096 | -0.101 |
| 2nd | 16 | -0.075 | -0.074 |
| 3rd | 17 | -0.068 | -0.068 |
| 4th | 14 | -0.066 | -0.064 |
| 5th | 15 | -0.049 | -0.050 |

Figure 8 shows that the only remaining bus with voltage probability slightly larger than 0.05 is bus 13. Again, Sobol' indices can be used to decide power factor of DGs at which buses must be modified to achieve failure probability smaller than 0.05 at bus 13. Also, in the case that the probability of voltage violation at some buses remain larger than 0.05 even by reducing the power factor of controllable DGs at all buses to their minimum allowable power factors, Sobol' indices can be used to identify those buses whose active power generation must be curtailed.

Figure 8 Monte Carlo estimation of voltage violation probability after reducing the power factors at buses 18, 16 and 17 vs. approximated voltage violation probability using PCE surrogate




## 5. Conclusion

In this work, we considered a distribution network with multiple distributed generators, where power consumption and generation at all buses were random and highly correlated. We used compressive sampling to construct PCE surrogates for power flow model with a relatively small number of simulations. PCE surrogates estimate voltage profile of the system as functions of active power consumption and generation, and power factors of distributed generators. The approximated PCE surrogates are then used, with minimal computational cost, for prediction, deriving voltage probability distributions and efficient sensitivity analysis. The sensitivity analysis results can be used to identify influential distributed generators for which power factor must be adjusted to reduce the chance of voltage violation. We demonstrated the validity of our proposed method on IEEE 33-bus system. The results show that the proposed approach provides efficient and accurate probabilistic voltage analysis.

## 6. References

Alanne K and Saari A (2006) *Distributed energy generation and sustainable development.* Renewable and sustainable energy reviews, 10(6), 539-558.

Bruckstein AM *et al*. (2009) *From sparse solutions of systems of equations to sparse modeling of signals and images.* SIAM review, 51(1), 34-81.

Candès EJ and Wakin MB (2008) *An introduction to compressive sampling*. IEEE signal processing magazine, 25(2), 21-30.

Coster EJ *et al*. (2011) *Integration issues of distributed generation in distribution grids*. Proceedings of the IEEE, 99(1), 28-39.

Doostan A and Owhadi H (2011) *A non-adapted sparse approximation of PDEs with stochastic inputs*. Journal of Computational Physics, 230(8), 3015-3034.

Hung DQ. and Mithulananthan N (2013) *Multiple distributed generator placement in primary distribution networks for loss reduction*. IEEE Transactions on industrial electronics, 60(4), 1700-1708.

Jhala K *et al*. (2018) *Probabilistic voltage sensitivity analysis (PVSA)—A novel approach to quantify impact of active consumer*s. IEEE Transactions on Power Systems, 33(3), 2518-2527.

Lou Y (2017) https://sites.google.com/site/louyifei/Software (accessed14/04/2019).

Ni F *et al* (2017) *Basis-adaptive sparse polynomial chaos expansion for probabilistic power flow*. IEEE Transactions on Power Systems, 32(1), 694-704.

Ni F *et al* (2018) *Variance-based global sensitivity analysis for power systems*. IEEE Transactions on Power Systems, 33(2), 1670-1682.

Ren Z *et al* (2016) *Probabilistic power flow analysis based on the stochastic response surface method*. IEEE Transactions on Power Systems, 31(3), 2307-2315.

Shin Y and Xiu D (2016) *Nonadaptive quasi-optimal points selection for least squares linear regression*. SIAM Journal on Scientific Computing, 38(1), A385-A411.

Sobol IM (2001) *Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates*. Mathematics and computers in simulation, 55(1-3), 271-280.

Sudret B (2008) *Global sensitivity analysis using polynomial chaos expansions*. Reliability engineering & system safety, 93(7), 964-979.

Sudret B and Caniou Y (2013) *Analysis of covariance (ANCOVA) using polynomial chaos expansions* (pp. 3275-3281). ETH Zürich.

Tamp F and Ciufo P (2014) *A sensitivity analysis toolkit for the simplification of MV distribution network voltage management*. IEEE Transactions on Smart Grid, 5(2), 559-568.

Valverde G and Van Cutsem T (2013) *Model predictive control of voltages in active distribution networks*. IEEE Transactions on Smart Grid, 4(4), 2152-2161.

Weckx S *et al.* (2015) *Voltage sensitivity analysis of a laboratory distribution grid with incomplete data.* IEEE Transactions on Smart Grid, 6(3), 1271-1280.

Xiu D (2010) *Numerical methods for stochastic computations: a spectral method approach*. Princeton university press.

Zad BB *et al* (2015) O*ptimal reactive power control of DGs for voltage regulation of MV distribution systems using sensitivity analysis method and PSO algorithm*. International Journal of Electrical Power & Energy Systems, 68, 52-60.

Zhang Z *et al* (2018) *A novel voltage sensitivity approach for the decentralized control of DG plants*. IEEE Transactions on Power Systems, 33(2), 1566-1576.