THESIS

NOVEL METHODS TO QUANTIFY ALEATORY AND EPISTEMIC UNCERTAINTY IN HIGH SPEED NETWORKS

Submitted by

Ishan Deepak Kapse

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2017

Master's Committee:

Advisor: Sourajeet Roy

Sudeep Pasricha
Charles Anderson

ABSTRACT

NOVEL METHODS TO QUANTIFY ALEATORY AND EPISTEMIC UNCERTAINTY IN HIGH SPEED NETWORKS

With the sustained miniaturization of integrated circuits to sub-45 nm regime and the increasing packaging density, random process variations have been found to result in unpredictability in circuit performance. In existing literature, this unpredictability has been modeled by creating polynomial expansions of random variables. But the existing methods prove inefficient because as the number of random variables within a system increase, the time and computational cost increases in a near-polynomial fashion. In order to mitigate this poor scalability of conventional approaches, several techniques are presented, in this dissertation, to sparsify the polynomial expansion. The sparser polynomial expansion is created, by identifying the contribution of each random variable on the total response of the system. This sparsification is performed primarily using two different methods. It translates to immense savings, in the time required, and the memory cost of computing the expansion. One of the two methods presented is applied to aleatory variability problems while the second method is applied to problems involving epistemic uncertainty. The accuracy of the proposed approaches is validated through multiple numerical examples.

ACKNOWLEDGEMENTS

I thank my advisor, Dr. Sourajeet Roy for his timely guidance and also for his support, not only as an academic advisor, but also as a mentor. I thank my committee members Dr. Sudeep, Dr. Ali and Dr. Charles for their valuable guidance. I would also like to thank my colleagues Aditi Krishna Prasad and Majid Ahadi as my work would not have been complete without their contributions. I would like to thank the staff at Colorado State University for guiding me and helping me whenever I needed it. Lastly, I would also like to thank all the people of Colorado who helped me feel at home and made my experience at CSU a memorable one.

DEDICATION

I dedicate this work to my parents whose unconditional love and support will always be my greatest strength.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

As the technology we use advances, there is an ever-growing need to get high performance in electronic devices and to make it compact at the same time. One of the ways to achieve high performance is to put increasingly higher number of transistors on the chip. To put more number of transistors on a chip that has a smaller area constraint, as compared to the previous generation of the same technology, the transistors themselves need to be made smaller. As of today, transistors of the 10nm technology node are being processed. But with this level of miniaturization, new challenges are raised which motivates new areas of research. This work aims at addressing to resolve some of these challenges.

1.1 Problem statement

With the sustained miniaturization of integrated circuits to sub-45 nm regime and the increasing packaging density, random process variations have been found to result in unpredictability in the performance of high-speed circuits. As a result, contemporary computer aided design (CAD) tools need to be flexible enough to be able to predict the impact of parametric uncertainty on general circuit responses. This unpredictability has been traditionally modeled using brute-force Monte Carlo (MC) method, where a large set of inputs are generated based on the PDFs of the random variables and each input is simulated through some commercial circuit solver such as SPICE [1]-[6]. The results of these simulations are then collected and any statistical information required about the response of the circuit is acquired from these set of results. Although the method is simplistic, the slow convergence of this approach requires an increasingly large number of deterministic simulations of the network model to achieve accurate statistical results, as the number of dimensions in the system go on increasing. This method is particularly

inefficient when a single simulation of the network is computationally expensive and time consuming. This makes the MC approach computationally infeasible for analyzing large networks [7].

To address the slow convergence of the MC approach, new robust uncertainty quantification techniques have been introduced based on the generalized polynomial chaos (PC) theory, that model the uncertainty in the network response as an expansion of predefined orthogonal polynomial basis functions of the input random variables [7]-[48]. The coefficients of the PC expansion form the new unknowns of the system and are evaluated through either intrusive [7]-[22] or non-intrusive approaches [23]-[42].

Among intrusive approaches, the Stochastic Galerkin (SG) approach is extensively used, since this approach is highly accurate and it only requires the simulation of a single augmented coupled deterministic network model to determine the PC coefficients [8]-[22]. The problem with this method though is that the computation cost of the creation of a coupled deterministic network model grows near-exponentially with the number of random variables in the circuit. The problem is further exacerbated for non-linear networks where the non-linearity in the network has to be further augmented in the network model using additional dependent voltage / current sources. While recent works such as the decoupled PC algorithm [17] and the stochastic testing method [43] can mitigate the time and memory costs of the standard SG approach, both approaches require the development of intrusive codes that preclude the direct exploitation of SPICE-like legacy circuit simulators. These bottlenecks have limited the applicability of the SG approach to problems featuring only low-dimensional random spaces [49], [50].

Among non-intrusive approaches, the pseudo-spectral collocation method [24], the linear regression approach [23], [26], [27] and the non-intrusive stochastic collocation approach [25]

are some of the most widely used methods for VLSI and EM problems. These methods are not as accurate as the intrusive methods, but they can make use of commercial circuit-solvers such as SPICE. Since these approaches create a meta-model of the original system, any value of the random variable located within the original random space can be probed into the PC expansion generated and any statistical result of the network response can be generated, with minimal loss in accuracy. An added benefit is that since the network is treated as if it was a decoupled network, the individual simulations can be parallelized.

Irrespective of what approach is used, they scale poorly as the number of random dimensions in the network increase [49], [50]. Thus, any of these approaches become computationally expensive for problems involving higher numbers of random variables.

1.2 Goals of the thesis

Among non-intrusive approaches, the linear regression approach is one of the most widely used approaches [26], [31], [32], [33], [50]. This approach probes the PC expansion of the network responses at an oversampled set of multi-dimensional nodes located within the random space, thereby leading to the formulation of an overdetermined set of linear algebraic equations. These equations can be solved in a linear least-squares sense to directly evaluate the PC coefficients of the network responses [50]. The multi-dimensional set of nodes is typically a subset of the set of the tensor product of the unidimensional quadrature nodes [32], [33]. This small subset of nodes is referred to as the design of experiments (DoE). However, blindly choosing the DoE can lead to inaccurate results as is demonstrated in the work of [31]. Conventionally, the linear regression approach has not identified any particular formal method to identify or optimize the DoE [32], [33]. Although the stochastic testing approach does have a reliable method where the number of DoE is equal to the number of unknown PC coefficients [43], [44], this technique does not

choose DoE based on any optimality criterion and hence does not guarantee the maximum possible accuracy of results. This work chooses a random set of DoE to begin with, and then uses the D-optimality criterion to optimize the chosen set of DoE [51], [52]. This work also uses the sparse linear regression (SPLINER) approach of [27] to further reduce the time required and the cost of computing the PC coefficients. Even then, the linear regression approach scales in a near-quadratic manner as the number of random dimensions increases.

This work attempts at further reducing the scalability issues of the above mentioned approaches. To do so, it uses the high dimensional model representation (HDMR) formulation of [53] to quantify the impact of each random dimension acting alone on the response of the network. This information is obtained by generating 1-D PC expansions for each random dimension and then analyzing the impact of each dimension by computing sensitivity indices based on the variance of these 1-D expansions. Once the sensitivity indices for each random dimension are obtained, this information can be used to intelligently prune some of the higher order terms of the PC expansion [29]. This results in a smaller set of coefficients in the resultant PC expansion, and thus reduces the time and memory cost of computing the PC coefficients. The way the terms are pruned is by computing the sensitivity index, for each dimension, in increasing orders of expansion. Once the enrichment in the sensitivity index for a particular dimension, from a lower order of expansion to the higher order of expansion, falls below a certain threshold, the previous order of expansion is ascribed to that random dimension. This creates an anisotropic expansion as opposed to an isotropic expansion discussed so far. The accuracy of this method and the speedup it provides is demonstrated using multiple numerical examples. Finally, the proposed method is combined with the method of [28] to extract an even sparser PC expansion [42].

The above mentioned methods can be applied effectively to problems where the PDF of the random variables is known or there is sufficient information about the probability distribution of the random variables. But for a large number of real-life problems, there is insufficient information about the probability distribution of the random variables and that, more information is acquired as more simulations of the system are carried out. This is classified as epistemic uncertainty. This work proposes a novel method to reduce the computation cost of generating a PC expansion for problems involving epistemic uncertainty, which uses the HDMR formulation [53] to analyze the impact of each epistemic random dimension. Since there is insufficient information about the random dimensions themselves, the impact of each random dimension cannot be quantified using the variance based method mentioned earlier [23], [47]. Instead, this work creates a new method to quantify this epistemic uncertainty and thus creates a set of sensitivity indices using this method. By using the information of the sensitivity indices, this work then eliminates all those random dimensions whose value of sensitivity index lies below a certain pre-determined threshold. By eliminating the random dimensions with the lowest impact on the response of the network, a reduced dimensional PC expansion is created that requires only a fraction of the time and memory cost it requires to compute the full dimensional PC expansion. Further this generated PC expansion can be probed to obtain the statistical results of any range of values of random dimensions that lie within the original random space.

1.3 Organization of the thesis

This thesis attempts at being self-explanatory for the reader without a major need for any prior knowledge in uncertainty quantification. Most of the state of the art PC approaches are reviewed. Exploited techniques are explained in detail, and novel ideas are supported with extensive numerical examples and discussions. The rest of the text is organized as follows.

Chapter 2 provides a review of basics of the generalized PC (gPC) theory and describes the general intrusive and non-intrusive approaches used to create a gPC expansion. Since the thesis primarily uses the linear regression approach, this particular non-intrusive technique is explained in greater detail. It also discusses the methods used to select the best set of design of experiments (DoE), including the D-optimality criterion and the Fedorov search algorithm. Chapter 3 deals with the improvements made to the conventional linear regression approach using the HDMR principles. It primarily talks about a novel method developed to reduce the number of terms in the PC expansion. Numerical examples are provided to demonstrate the accuracy of the proposed approach against conventional linear regression approach. Chapter 4 uses an HDMR formulation to formulate sensitivity indices which enable the truncation of a high dimensional PC model to a reduced model. Since the problems involve random variables of an epistemic nature, the formulation of sensitivity indices is done in a separate manner as compared to the one in Chapter 3. To evaluate the accuracy of the proposed method, this work compares the results of the proposed method against the results obtained from Monte Carlo simulations. The reason for choosing MC simulations is that since we are using high dimensional problems, generating full-blown PC expansions can prove to be more computationally expensive than MC simulations.

CHAPTER 2: KNOWN UNCERTAINTY QUANTIFICATION METHODS

This chapter is aimed at familiarizing the reader with some of the existing methods used to address the problem of uncertainty quantification. It also provides the information necessary to further help the reader understand the focus of this thesis. First, the basis of this entire research, the generalized polynomial chaos (gPC) theory is introduced. Next, a brief overview of the uncertainty quantification techniques involving various intrusive and non-intrusive approaches is presented. Since the thesis uses the linear regression technique extensively, this technique is elaborated in greater detail. Next, a few methods used to identify and remove statistically insignificant terms in the gPC expansion are discussed. These methods serve as the basis for future chapters in this thesis.

2.1 Generalized Polynomial Chaos (gPC) theory

Polynomial chaos, originally known as the Wiener chaos expansion, was introduced by Dr. Norbert Wiener [55], [56], where Hermite polynomials were used to model stochastic processes with Gaussian random variables. Although it was originally applied to only Gaussian distributions of random variables, a generalized framework was soon developed that was subsequently used for all kinds of distributions of random parameters, both known (Normal, Uniform, Beta, Gamma, etc.) [54], and unknown. This generalized framework was later termed as the generalized polynomial chaos theory. Today, the gPC theory is widely used in several engineering fields such as VLSI design, fluid dynamics and space research to name a few.

2.1.1 Basics of the gPC theory

If the uncertainty in a system can be modeled using one random variable λ, then the output of that system can be modeled as an expansion of orthogonal polynomials weighted by the

polynomial chaos (PC) coefficients. Mathematically, the system response X(t, λ) can be represented as

$$X(t,\lambda) = \sum_{k=0}^{\infty} c_k(t)\Phi_k(\lambda) \tag{2.1}$$

where c_k(t) is the kth time dependent PC coefficient and $\Phi_k(\lambda)$ is a unidimensional polynomial basis term orthogonal with respect to the probability density function (PDF) of λ. Notice that (2.1) can potentially have an infinite number of terms in it, but for all practical purposes, the expansion is truncated to m+1 terms.

$$X(t,\lambda) = \sum_{k=0}^{m} c_k(t)\Phi_k(\lambda) \tag{2.2}$$

where m is the maximum degree of expansion of the polynomial basis terms. It is necessary for the polynomials to be orthogonal with respect to the PDF of the random variable, because the inner product of two orthogonal basis terms exhibits an interesting property which simplifies the calculation of the PC coefficients. The inner product of two polynomial basis terms can be mathematically expressed as

$$< \Phi_i(\lambda)\Phi_j(\lambda) > = \int_{\Omega} \Phi_i(\lambda)\Phi_j(\lambda)\,\rho(\lambda)d\lambda = \alpha_i^2\delta_{ij} \tag{2.3}$$

where <,> is the inner product notation, Ω denotes the space of the random variable, ρ represents the PDF of λ, α_i^2 is a constant and δ_{ij} denotes the Kronecker's delta function. The interesting property mentioned earlier is that δ_{ij} = 0 for i ≠ j and δ_{ij} = 1 for i = j. This translates to the inner product of two orthogonal basis terms to be zero, for i ≠ j, and α_i^2, for i = j. Typically the orthogonal basis terms are normalized with the value of α_i and thus, the rest of this thesis uses

the term orthonormal and orthogonal interchangeably. The choice of the polynomial function is made based on the Wiener-Askey scheme [54], which states that there is a one-to-one correspondence between the class of the orthogonal polynomials and the distribution of λ. Table 2.1 shows the choice for the class of orthogonal polynomials, made for the corresponding standard distributions of λ. Although it is possible to use a different polynomial basis instead of the one mentioned in Table 2.1, it has been established that the Wiener-Askey scheme yields the optimal convergence for the PC expansion. In case of arbitrary unknown distributions, the orthogonal polynomials can still be theoretically constructed.

Table 2.1: Wiener-Askey Scheme for gPC

Distribution of λ	Class of orthogonal polynomials	Support Ω
Gaussian	Hermite	(-∞,+∞)
Uniform	Legendre	[-1,1]
Beta	Jacobi	[-1,1]
Gamma	Laguerre	[0, ∞)

2.1.2 Generation of one-dimensional orthonormal polynomials

For a Gaussian random variable, the PDF $\rho(\lambda)$ is mathematically expressed as

$$\rho(\lambda) = \frac{1}{\sqrt{2\lambda}} e^{\frac{-\lambda^2}{2}} \tag{2.4}$$

Referring to the Wiener-Askey scheme, the Hermite polynomials are chosen to form the basis of the PC expansion since they are orthogonal to the Gaussian distribution. The Hermite polynomials can be computed, either using an analytic formula [54]

$$\Phi_k(\lambda) = (-1)^k e^{\frac{\lambda^2}{2}} \frac{d^k}{d\lambda^k} e^{\frac{-\lambda^2}{2}} \tag{2.5}$$

or alternatively, using a three-term recurrence relation

$$\Phi_{k+1}(\lambda) = \lambda\Phi_k(\lambda) - k\Phi_{k-1}(\lambda) \tag{2.6}$$

where $\phi_0 = 1$ and $\phi_1 = \lambda$. The normalizing factor α_i^2 for Hermite polynomials can be found as

$$\alpha_i^2 = < \Phi_i(\lambda), \Phi_i(\lambda) > = i! \tag{2.7}$$

Similarly, for a uniform random variable, the PDF $\rho(\lambda)$ is mathematically expressed as

$$\rho(\lambda) = \begin{cases} 0.5, & -1 \leq \lambda \leq 1 \\ 0, & Otherwise \end{cases} \tag{2.8}$$

Now the Legendre polynomials are chosen since they are orthogonal with respect to the PDF of λ. To compute the polynomials analytically, the following formula is used [54].

$$\Phi_k(\lambda) = \frac{1}{2^k k!}\frac{d^k}{d\lambda^k}(\lambda^2 - 1)^k \tag{2.9}$$

Alternatively, one can also use the three-term recurrence relation.

$$\Phi_{k+1}(\lambda) = \frac{2k+1}{k+1}\lambda\Phi_k(\lambda) - \frac{k}{k+1}\Phi_{k-1}(\lambda) \tag{2.10}$$

where $\phi_0 = 1$ and $\phi_1 = \lambda$. The normalizing factor α_i^2 for Legendre polynomials can be found as

$$\alpha_i^2 = < \Phi_i(\lambda), \Phi_i(\lambda) > = \frac{1}{2i+1} \tag{2.11}$$

For illustration purposes, the first six unidimensional Hermite and Legendre polynomials are listed in Table 2.2.

Table 2.2: First six unidimensional Hermite and Legendre polynomials

Bases	Orthonormal Hermite polynomial	Orthonormal Legendre polynomial
$\Phi_0(\lambda)$	1	1
$\Phi_1(\lambda)$	λ	$\sqrt{3}\,\lambda$
$\Phi_2(\lambda)$	$(\lambda^2-1)/\sqrt{2}$	$\sqrt{5}\left(\frac{3}{2}\lambda^2-\frac{1}{2}\right)$
$\Phi_3(\lambda)$	$(\lambda^3-3\,\lambda)/\sqrt{6}$	$\sqrt{7}\left(\frac{5}{2}\lambda^3-\frac{3}{2}\lambda\right)$
$\Phi_4(\lambda)$	$(\lambda^4-6\,\lambda^2+3)/(2\sqrt{6})$	$3\left(\frac{35}{8}\lambda^4-\frac{30}{8}\lambda^2+\frac{3}{8}\right)$
$\Phi_5(\lambda)$	$(\lambda^5-10\,\lambda^3+15\,\lambda)/(2\sqrt{30})$	$\sqrt{7}\left(\frac{63}{8}\lambda^5-\frac{70}{8}\lambda^3+\frac{15}{8}\lambda\right)$

2.1.3 Generation of multi-dimensional orthonormal polynomials

In most practical applications, there are multiple random variables in a stochastic network. In these networks, it is necessary to account the effects of the random parameters, acting together, on the total response of the system. To do so, a multi-dimensional PC expansion is constructed, where each random variable is represented by its own unique dimension. In case of multi-dimensional PC, λ represents a vector of n random variables $\lambda = [\lambda_1,\ \lambda_2, \dots,\ \lambda_n]^T$. The vector of random variables is designated with a transpose sign (T) to represent n mutually uncorrelated random variables. Similar to (2.2), a multidimensional PC expansion is written as

$$X(t,\lambda) = \sum_{k=0}^{P} c_k(t)\phi_k(\lambda) \tag{2.12}$$

where $\phi_k(\lambda)$ represents a multi-dimensional orthonormal basis term and the expansion is truncated to P terms such that

$$P + 1 = \binom{m+n}{m} = \frac{(m+n)!}{m!\,n!} \tag{2.13}$$

The multi-dimensional basis terms $\phi_k(\lambda)$ are the product of the unidimensional polynomial basis term across each dimension.

$$\phi_k(\lambda) = \prod_{i=1}^{n} \Phi_{k_i}(\lambda_i) \; where \sum_{i=1}^{n} k_i \leq m \tag{2.14}$$

In case of multi-dimensional polynomial bases, it is possible that different random variables within a network have different distributions. In this case, the unidimensional basis for each random variable is independently chosen according to the Wiener-Askey scheme and then multiplied to form the multidimensional bases. These multidimensional polynomials are now orthonormal to the joint PDF of the input random variables. Since the variables are mutually uncorrelated, their joint PDF can be expressed as the product of their individual PDFs.

$$< \phi_i(\lambda)\phi_j(\lambda) > = \int_{\Omega} \phi_i(\lambda)\phi_j(\lambda)\,\rho(\lambda)d\lambda = \delta_{ij} \tag{2.15}$$

Since the unidimensional polynomials were normalized, the inner product of two multi-dimensional orthonormal polynomials is equal to one, for i = j, and zero, for i ≠ j.

For illustration purposes, the P+1 multi-dimensional polynomial bases for a 2 dimensional, fourth order PC expansion are listed below.

$$\Phi_0(\lambda_1)\Phi_0(\lambda_2)$$
$$\Phi_1(\lambda_1)\Phi_0(\lambda_2) \quad \Phi_0(\lambda_1)\Phi_1(\lambda_2)$$
$$\Phi_2(\lambda_1)\Phi_0(\lambda_2) \quad \Phi_1(\lambda_1)\Phi_1(\lambda_2) \quad \Phi_0(\lambda_1)\Phi_2(\lambda_2)$$
$$\Phi_3(\lambda_1)\Phi_0(\lambda_2) \quad \Phi_2(\lambda_1)\Phi_1(\lambda_2) \quad \Phi_1(\lambda_1)\Phi_2(\lambda_2) \quad \Phi_0(\lambda_1)\Phi_3(\lambda_2)$$
$$\Phi_4(\lambda_1)\Phi_0(\lambda_2) \quad \Phi_3(\lambda_1)\Phi_1(\lambda_2) \quad \Phi_2(\lambda_1)\Phi_2(\lambda_2) \quad \Phi_1(\lambda_1)\Phi_3(\lambda_2) \quad \Phi_0(\lambda_1)\Phi_4(\lambda_2)$$

The total degree across all dimensions remains constant in each row and is incremented row by row [54].

2.1.4 Generation of statistical moments using PC coefficients

The main advantage of generating a PC expansion to model the uncertainty in a system is that the statistical moments for that system can be generated by using the PC metamodel. The mean and standard deviation of the output can be obtained as a function of the coefficients by integrating over the random space Ω_n. The PDF and other higher order statistical moments can be generated by probing the PC metamodel at a large number of Monte Carlo (MC) samples. Note that it is not necessary to simulate the network at the MC samples, but rather the output can be computed analytically using the PC metamodel.

2.1.4.1 Calculation of the arithmetic mean

The arithmetic mean is the first order statistical moment. The mean shows the value of the response when no parametric uncertainty acts on the network. Mathematically, the expected value or the mean of the output $X(\lambda)$ is expressed as

$$E\big(X(\lambda)\big) = \int_{\Omega_n} X(\lambda)\rho(\lambda)d\lambda \tag{2.16}$$

By replacing (2.12) in (2.16), we get

$$E\big(X(\lambda)\big) = \sum_{k=0}^{P} \int_{\Omega_n} c_k \phi_0(\lambda)\phi_k(\lambda)\rho(\lambda)d\lambda \tag{2.17}$$

$\phi_0(\lambda)$ is always one for all orthonormal polynomials. $\phi_0(\lambda)$ is introduced in (2.17) to simplify the equation further.

$$E\big(X(\lambda)\big) = \sum_{k=0}^{P} \int_{\Omega_n} c_k < \phi_0(\lambda), \phi_k(\lambda) > \; = \; c_0 \tag{2.18}$$

In (2.18), all the terms in the expansion are equal to zero, except when k = 0. Hence, the mean of the network response is represented by the zeroth order PC coefficient.

2.1.4.2 Calculation of the Variance

The variance is the second order statistical moment. It shows the maximum possible deviation of the response from the mean value, given that the parametric uncertainty is maximum in the network. The variance of the output $X(\lambda)$ is calculated as

$$Var\big(X(\lambda)\big) = E(X(\lambda) - E(X(\lambda)))^2$$

$$= \int_{\Omega_n} (\sum_{i=0}^{P} c_k \phi_k(\lambda) - \; c_0\phi_0(\lambda))^2 \, \rho(\lambda)d\lambda$$

$$= \sum_{k=0}^{P}\sum_{j=0}^{P} c_k c_j \int_{\Omega_n} \phi_i(\lambda)\phi_j(\lambda)\, \rho(\lambda)d\lambda - \; c_0^2$$

$$= \sum_{k=0}^{P} c_k^2 - c_0^2 = \sum_{k=1}^{P} c_k^2 \tag{2.19}$$

Hence the variance is the sum of squares of all the coefficients in the expansion, except the zeroth order PC coefficient. The standard deviation is another popular statistical measurement tool. Standard deviation or σ is the square root of the variance. Standard deviation is typically used for assessment since it is directly comparable to the mean of the response.

$$\sigma = \sqrt{\sum_{k=1}^{P} c_k^2} \tag{2.20}$$

2.1.4.3 Higher order moments and PDF

Higher order moments are not typically used, but they provide more information about the random behavior of the response of a network. The general formula for an M^{th} order moment is [57]

$$\mu_M\big(X(\lambda)\big) = E(X(\lambda) - E(X(\lambda)))^M$$

$$= \int_{\Omega_n} (X(\lambda) - E(X(\lambda)))^M \rho(\lambda) d\lambda \tag{2.21}$$

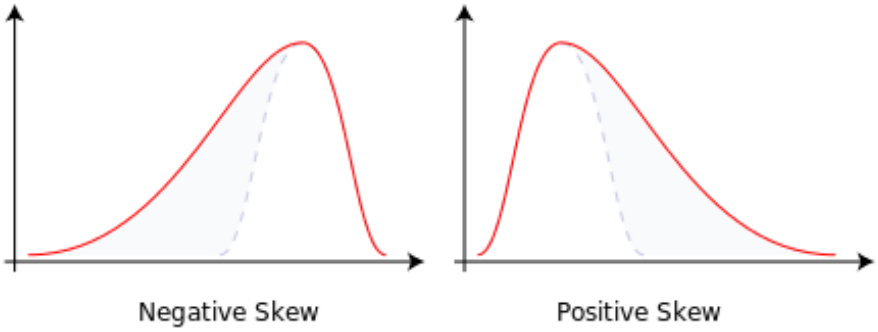


Fig. 2.1: Illustration of a skew. Image used from [58].

The third order moment is known as skewness. Skewness is the measure of the asymmetry of the probability distribution of the response about its mean value. Consider the two distributions shown in Fig. 2.1.

The right side of the distribution tapers differently from the left side. These tapering sides are called the tails. One can observe the distribution of the response to identify whether the response has a positive, negative or no skew at all [58].

The fourth order moment is called the kurtosis. This statistical moment gives information about the shape of the tails of the probability distribution of the response [59].
In case of statistical moments having an order greater than two, the inner product approach cannot be implemented since the integral now involves third degree polynomials. However, an easier method to generate higher order statistical moments is to generate a large number of MC samples according to the random variable distribution. The PC metamodel can then be used to generate the responses of the network at each of these samples, with a very small amount of loss of accuracy. The values of the responses can then be substituted in (2.21) to easily obtain any higher order statistical moment.

Lastly, the PDF is one of the most important statistical moments as it holds the information of all the statistical moments discussed so far. It is sometimes called the zeroth order statistical moment. To generate a PDF from a PC expansion, we simply start with a large random set of MC samples and probe the PC metamodel to generate the responses at each sample node. Then a histogram of all the responses is created to acquire the PDF of the response at any given time/frequency point. Conversely, one can perform network simulations at each of those sample nodes to get the responses. Although it has to be noted that running large number of simulations may require a lot of time, especially when one simulation itself takes a while to complete.

2.2 Intrusive and non-intrusive approaches

So far, we have established that finding the PC coefficients are the key to performing uncertainty quantification using gPC. The logical step then is to evaluate the PC coefficients. There are several methods used to achieve this goal. They are primarily divided into intrusive and non-intrusive methods.

2.2.1 Intrusive methods

Intrusive methods are methods which require intrusive coding and cannot be done in a black box manner [7]-[22]. These methods typically require the development of a new circuit solver. These methods exhibit a higher accuracy compared to the non-intrusive methods. The Stochastic Galerkin (SG) projection is one of the widely used intrusive methods. This method creates an augmented and coupled deterministic network, based on the equations governing the system. The unknown PC coefficients can then be obtained from a single simulation of this augmented coupled system. The higher accuracy of this system allows SG to choose lower orders of expansion. But even then, since the augmentation is of P+1 terms, the time and computation costs required for this approach scale in a near exponential manner, as the number of random variables increase. Further non-linear elements are modeled using lumped dependent sources and this further adds to the augmentation [15]. Thus, the SG approach is typically good for smaller circuits with a fewer number of random variables.

The shortcomings of the SG approach are addressed by the intrusive Stochastic Testing (ST) formulation [43], [44]. The ST approach solves the coupled system of equations of the augmented circuit, but in a decoupled manner at each time point. The coupled equations are solved at P+1 sampling nodes, but the choice of nodes affects the accuracy of the approach. A poor selection of nodes results in ill conditioned matrices, which in turn makes the solution

inaccurate or even impossible to obtain. Hence, a node selection algorithm is used to obtain the optimum set of sampling nodes. The node selection algorithm works by first creating a tensor product of $(m+1)^n$ nodes, based on the Wiener-Askey scheme of polynomials. The first node is selected, from this tensor product, which has the highest quadrature weight among all the nodes. The other P nodes are also selected based on their quadrature weights but with an additional condition that the node needs to have a large enough orthogonal component to the already generated set of nodes. Since the node selection algorithm uses the quadrature weights as a basis for evaluation, the algorithm does not guarantee the best selection of nodes, especially for higher dimensional problems. Also, since ST is an intrusive formulation, it cannot make use of commercial circuit solvers like SPICE. The advantage of ST approach over SG is that it can be solved in a decoupled manner and thus, the P+1 simulations required, can be executed in a parallelized manner.

2.2.2 Non-intrusive methods

The main advantages of non-intrusive methods over intrusive methods is that existing deterministic solvers, such as SPICE, can be used to determine the network response at the selected nodes. The simulations themselves can also be parallelized as they are independent from each other, leading to an even higher speedup as compared to intrusive methods like SG.

2.2.2.1 Pseudo spectral PC approach

Pseudo spectral PC approach is the most basic and yet the most accurate of all the non-intrusive approaches [24]. Being non-intrusive, this approach does not require the knowledge of the internal equations governing the circuit and it eliminates the need to design a deterministic circuit solver. In the pseudo spectral PC approach, the output response is expressed as an expansion of a series of orthogonal polynomials, as in (2.12), to generate a PC meta-model,

where the PC coefficients are the new unknowns of the system. The PC coefficients are determined using numerical integration techniques. Gaussian quadrature techniques are numerical integration tools, used to approximate the integral of a function $F(\lambda)$ as a weighted sum of function values computed at predetermined sample points.

$$\int_{\Omega_n} F(\lambda)\rho(\lambda)d\lambda = \sum_{i=1}^{Q} F(\lambda^i)w(\lambda^i) \#(2.22)$$

where Q is the number of simulation nodes, $\lambda^i = [\lambda_1^{(i)}, \lambda_2^{(i)}, \dots, \lambda_n^{(i)}]$ is a simulation node in the random space Ω_n, $w(\lambda^i)$ is the corresponding quadrature weight to the node λ^i and $F(\lambda^i)$ is the output of the function $F(\lambda)$ at λ^i. λ^i for a unidimensional problem are the roots of the polynomials generated according to the Wiener-Askey scheme. The number of unidimensional roots thus generated, are $(\mathrm{m}+1)$, m being the maximum order of expansion. For a multi-dimensional problem, Q nodes are generated by creating the tensor product of all the unidimensional nodes. Therefore $Q = (m+1)^n$. Similarly, $w(\lambda^i)$ is the product of all the weights corresponding to those unidimensional nodes which are a part of λ^i.

Unidimensional nodes can be generated from the roots of polynomials. An alternative way of generating the nodes and their corresponding weights is by solving an eigenvalue problem, which is known as the Golub-Welsch algorithm [60]. The method to generate nodes and weights, using the Golub-Welsch algorithm, for Hermite and Legendre polynomials is described below.

In case of Hermite polynomials, first a matrix A is constructed in the following manner.

$$A(i,j)_{(q+1)*(q+1)} = \begin{cases} \sqrt{i} & j = i-1 \\ \sqrt{j} & i = j-1 \\ 0 & otherwise \end{cases} \quad (2.23)$$

Now, if an eigenvalue decomposition is performed on A, such that $A = W\Lambda W^T$, where W is a unitary matrix, then the nodes will coincide with the eigenvalues of A, i.e. $\lambda^i = \Lambda(i,i)$ and the corresponding weights are the squares of the first element of each eigenvector. i.e. $w(\lambda^i) = W_{1i}^2$.

In case of Legendre polynomials, the matrix A is constructed as

$$A(i,j)_{(q+1)*(q+1)} = \begin{cases} \dfrac{0.5}{\sqrt{1-\dfrac{1}{\left(2(j-1)\right)^2}}} & j = i-1 \\ \dfrac{0.5}{\sqrt{1-\dfrac{1}{\left(2(i-1)\right)^2}}} & i = j-1 \\ 0 & otherwise \end{cases} \tag{2.24}$$

Similar to the case of Hermite polynomials, if an eigenvalue decomposition is performed such that $A = W\Lambda W^T$, then the nodes will coincide with the eigenvalues of A, i.e. $\lambda^i = \Lambda(i,i)$ and the corresponding weights are the squares of the first element of each eigenvector. i.e. $w(\lambda^i) = W_{1i}^2$.

The unknown PC coefficients c_k can be calculated, using the orthogonal projection technique, which involves the calculation of the inner product of a function and a polynomial

$$c_k = < X, \phi_k > = \int_{\Omega_n} X(t,\lambda)\,\phi_k(\lambda)\rho\lambda d\lambda = \sum_{k=1}^{Q} X\left(t,\lambda^i\right)\phi_k\left(\lambda^i\right)w\left(\lambda^i\right) \tag{2.25}$$

As the number of simulations required in the pseudo-spectral approach is $(m+1)^n$, this approach is good for solving problems involving lower number of dimensions. For a high dimensional problem though, the number of terms in the PC expansion and subsequently, the number of simulations required to calculate the coefficients scales in a near exponential manner

and thus, this method does not provide any significant advantages over using MC method for high dimensional problems.

2.2.2.2 Conventional linear regression approach

Conventional linear regression approach is a non-intrusive approach that starts with a randomly selected, smaller $2(P+1)$ subset of nodes from a larger $(m+1)^n$ tensor product of nodes, and uses the linear least squares technique to optimize the subset of nodes [61]. Since the thesis relies heavily on the linear regression approach, this particular approach is explained in greater detail in this section.

i. Linear least squares technique

As stated in (2.12), a response of a network can be modeled using the gPC theory as

$$X(t,\lambda) = \sum_{k=0}^{M} c_k(t)\phi_k(\lambda) \tag{2.26}$$

where M, in the case of linear regression, is an over-sampled set of nodes located within the random space Ω_n and $M = 2(P+1)$. The above equation can be expressed, for M nodes, in the matrix form as

$$Ac + \varepsilon = X \tag{2.27}$$

where

$$A = \begin{bmatrix} \phi_0(\lambda^{(1)}) & \cdots & \phi_N(\lambda^{(1)}) \\ \vdots & \ddots & \vdots \\ \phi_0(\lambda^{(M)}) & \cdots & \phi_N(\lambda^{(M)}) \end{bmatrix}; c = \begin{bmatrix} c_0 \\ \vdots \\ c_N \end{bmatrix}; X = \begin{bmatrix} X^{(1)} \\ \vdots \\ X^{(M)} \end{bmatrix}; \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_M \end{bmatrix} \quad (2.28)$$

ε is the random truncation error. Ideally, ε should be zero but practically, there exists some small non-zero error. Next, the linear least squares technique models the data in such a way that the maximum number of nodes are approximated to fit the model. Once this model is generated, the model can then be probed to generate a response for any input belonging to the random space Ω_n. For illustration purposes, refer to Fig. 2.2 where a linear model is created from a discrete set

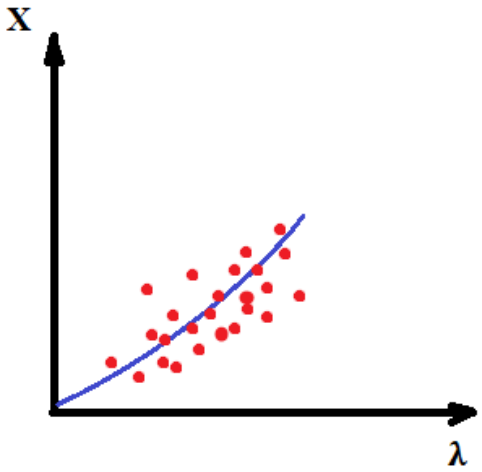


Fig 2.2: Illustration of curve fitting done by linear least squares approach

of points, in such a way that the error of approximation is minimized.

The linear least squares algorithm does not provide a unique answer to the given problem and the final selection of nodes is partly based on the initial random selection of nodes, but it is designed

to ensure that the most optimum model is created, by minimizing the sum of squares of errors term.

$$\tilde{c} = arg_c min S(c) \tag{2.29}$$

where

$$S(c) = \left||X - Ac|\right|^2 = \sum_{k=1}^{M} r_k^2 = \sum_{k=1}^{M} \left(\left| y^{(k)} - \sum_{j=1}^{N} c_j \phi_j(\lambda^{(k)}) \right| \right)^2 \tag{2.30}$$

Here, we assume an ideal case where ε is zero. (2.30) is minimum when the gradient of that equation is zero, since it is a convex function. Hence

$$\frac{\delta S}{\delta c_j} = 2 \sum_{k=1}^{M} r_k \frac{\delta r_i}{\delta c_j} = 2 \sum_{k=1}^{M} (y^{(k)} - \sum_{j=1}^{N} c_j \phi_j(\lambda^{(k)}))(-\phi_j(\lambda^{(i)})) = 0 \tag{2.31}$$

The simplified form of (2.31) would be

$$\sum_{k=1}^{M} \sum_{j=1}^{N} \tilde{c}_j \phi_j(\lambda^{(k)}) \phi_k(\lambda^{(k)}) = \sum_{k=1}^{M} \phi_k(\lambda^{(k)}) X_k \tag{2.32}$$

which can be further written in the matrix form as

$$(A^T A)\tilde{c} = A^T X \tag{2.33}$$

(2.33) can only have a solution when A^T is full column rank which makes $A^T A$ positive definite. Hence

$$\tilde{c} = (A^T A)^{-1} A^T X \tag{2.34}$$

This is the vector of coefficients from (2.28). Solving the above equation yields the PC coefficients. Any statistical moments' information can be calculated from the PC coefficients. It

is important to note that in the above equation, the term $(A^T A)$ is known as the Fisher information matrix or simply information matrix.

ii. Benefits and shortcomings

The major advantages of linear regression approach are that any commercial circuit solver can be used to generate the values of the responses required in (2.34). Moreover, since the simulation of each node is independent of the other, the required simulations can be parallelized. The number of simulations too scale, with the number of random numbers $M = 2(P + 1)$, in a polynomial fashion. It should also be noted that the matrix A can be stored and used for the same or any other problem, with the same number of random variables and maximum order of expansion. All these factors amount to a substantial speedup over any intrusive approach or the MC approach. It is also important to point out that by oversampling the initial set of nodes, it is ensured that this approach also provides results with acceptable accuracy.

On the other hand, the linear regression approach so far has been imperfect, in the sense that it does not provide a way of optimizing the subset of nodes selected. Randomly choosing M nodes, and not optimizing them, may or may not yield good results depending on the choice of nodes. Hence, there are several methods used in addition to the linear regression approach that work on optimizing some attribute of the (2.34).

iii. D-optimality criterion

There are several methods used to optimize the subset of regression nodes. One of the most popular techniques used to optimize the selection of the M regression nodes is the D-optimality criterion [51], [52]. The basis of this optimization technique is to minimize the $(A^T A)^{-1}$

component of the (2.34), and thus, essentially maximizing the determinant of the Information matrix.

As discussed in (2.27), the linear regression approach tries to solve the following system of algebraic equation.

$$Ac + \varepsilon = X \tag{2.35}$$

The importance of the D-optimal criterion to the accuracy of the evaluated PC coefficients of (2.34) is revealed using the following Lemma.

Lemma: Assuming that the truncation error $\varepsilon_j, 1 \leq j \leq M$ at all M design of experiments (DoE) of (2.28) are independent of each other and exhibit a normal distribution of zero mean and same variance σ^2, then in order to achieve the maximum accuracy of the PC coefficients, the DoE must be chosen such that the determinant of the Information matrix $A^T A$ is maximized.

Proof: Based on the PC expansion of the network responses of (2.35), it is understood that the presence of the random truncation error ε makes the PC coefficients themselves random variables. The variance of the evaluated PC of (2.34) can be computed as

$$Var(\tilde{c}) = Var((A^T A)^{-1} A^T X) = (A^T A)^{-1} A^T Var(X)((A^T A)^{-1} A^T)^T \tag{2.36}$$

Knowing that the truncation error for each DoE is independent (ε_j) and has a constant variance $\sigma^2, Var(X) = \sigma^2 I$, I being the identity matrix. Replacing this is (2.36), the variance of the PC coefficients can be compactly expressed as

$$Var(\tilde{c}) = (A^T A)^{-1} \sigma^2 \tag{2.37}$$

It is understood that to ensure the maximum accuracy of the PC coefficients, it is necessary to reduce the uncertainty in the solution $\tilde{c}$ (i.e. the variance of $\tilde{c}$). Since the variance of $\tilde{c}$ is

inversely proportional to the determinant of the information matrix $A^T A$, a simple way to minimize the variance of $\tilde{c}$ is to maximize the determinant. Therefore the M DoE for the linear regression of (2.34) must be chosen such that the determinant of the information matrix is maximized. This criterion is referred to as the D-optimal criterion. Other types of optimality criteria do exist, but the D-optimal criterion has been deemed the most effective and popular to date. The next challenge is to develop a search algorithm that can efficiently identify the D-optimal nodes from the tensor product space of nodes.

iv. Fedorov search algorithm

The selection of the D-optimal nodes is commonly performed using the Classical Fedorov search algorithm, which is a greedy search algorithm [61], [62]. This greedy search algorithm is commonly used in the field of estimation theory and data analysis. This algorithm begins by considering a set of $M = 2(P + 1)$ starting DoE selected from the tensor product grid of $(m + 1)^n$ multi-dimensional quadrature nodes and creating the corresponding information matrix $A^T A$. Thereafter, each DoE in the starting set is replaced by the best possible substitute DoE taken from the remaining $(m + 1)^n - M$ quadrature nodes such that the determinant of the information matrix increases by the maximum amount in the process. This step-by-step refinement of the starting DoE continues till all the initial set of nodes has been replaced [26], [31].

As per the above description, at the rth step, it is assumed that the first $r - 1$ nodes have been replaced by their best possible substitutes. Now if the rth DoE of the starting set is removed from A, then the new determinant of the information matrix can be expressed as

$$\det(A^T A)_{new} = \det\left((A^T A) - R\left(\lambda^{(r)}\right)R'\left(\lambda^{(r)}\right)\right)$$

$$= \det(A^T A)\left(1 - R\left(\lambda^{(r)}\right)(A^T A)^{-1} R'\left(\lambda^{(r)}\right)\right) \tag{2.38}$$

where $R\left(\lambda^{(r)}\right)$ is the row vector contributed by the r^{th} DoE in A. Similarly, if any arbitrary k^{th} DoE from the remaining $(m+1)^n - M$ quadrature nodes is included into A, the new determinant of the information matrix can be expressed as

$$\det(A^T A)_{new} = \det\left((A^T A) + R\left(\lambda^{(r)}\right) R^T\left(\lambda^{(r)}\right)\right)$$

$$= \det(A^T A)\left(1 + R\left(\lambda^{(r)}\right)(A^T A)^{-1} R'\left(\lambda^{(r)}\right)\right) \tag{2.39}$$

Combining the results of (2.38) and (2.39), after exchanging the r^{th} DoE of the starting set with any arbitrary k^{th} DoE from the remaining $(m+1)^n - M$ quadrature nodes, the new determinant of the new information matrix can be mathematically expressed as the recursive function

$$\det(A^T A)_{new} = det(A^T A)(1 + d_{kk} - d_{rr} + d_{kr}^2 - d_{kk} d_r)$$

$$d_{kr} = R\left(\lambda^{(k)}\right) \Psi^{(r-1)} R^T\left(\lambda^{(r)}\right) \tag{2.40}$$

where $\Psi^{(r-1)}$ represents the inverse of the information matrix obtained after the previous (i.e., $(r-1)^{th}$) exchange. From (2.40), it is understood that in to order to achieve D-optimality, the k^{th} node $\lambda^{(k)}$ needs to be chosen to satisfy the optimization criterion

$$\max(d_{kk} - d_{rr} + d_{kr}^2 - d_{kk} d_r) \tag{2.41}$$

Once the best possible node $\lambda^{(k)}$ has been found to satisfy (2.41) and the relevant exchange has been made, the new determinant can be directly updated using (2.40) and the substitution process moves on to the $(r+1)^{th}$ node. Once all M starting DoE have been replaced, the new set of

DoE will represent the D-optimal selection. In the next subsection, the computational cost associated with the node selection has been derived.

v. Computational cost

It is noted that the total computational cost of the search algorithm is due to two main factors. Firstly, identifying the D-optimal DoE requires searching through $(m+1)^n - M$ quadrature nodes for each DoE in the starting set – in other words, a total of $M((m+1)^n - M)$ searches. The associated CPU cost can be expressed as

$$C_a = 2(P+1)((m+1)^n - 2(P+1))C_1 \approx 2(P+1)(m+1)^n C_1 \#(2.42)$$

where C_1 is the CPU cost of computing the terms in the parenthesis of (2.42), assuming that the inverse $\Psi^{(r-1)}$ is known. It is noted that based on (2.40) and (2.42), C_1 can be expressed as

$$C_1 = 3k\left((P+1)^2 + (P+1)\right) \tag{2.43}$$

where the first term is the cost of the matrix-vector multiplication $\Psi^{(r-1)}R^T\left(\lambda^{(r)}\right)$, the second term is the cost of the vector-vector multiplication of $R\left(\lambda^{(r)}\right)$ with $\Psi^{(r-1)}R^T\left(\lambda^{(r)}\right)$, and the factor 3 is due to the fact that the above operations need to be performed for 3 scalars d_{rr}, d_{kk} and d_{kr} of (2.41). k is assumed to be the cost of each floating point operation. Combining (2.40) and (2.41), it can be concluded that the overall search cost (C_a) scales in an exponential manner with the number of random dimensions (n), quantified as

$$O((P+1)^3(m+1)^n) \approx O(n^{3m}(m+1)^n) \tag{2.42}$$

The other source of computational effort arises from the fact that for each substitution, the information matrix changes and the inverse $\Psi^{(r-1)}$ has to be reevaluated. This CPU cost is expressed as

$$C_b = 2(P+1)C_2 \tag{2.43}$$

where C_2 is the CPU cost of each matrix inversion. It is noted that for direct inversion methods, C_2 scales as $O((P+1)^3)$ thereby ensuring that the cumulative cost of the matrix inversions (C_b) scales as $O((P+1)^4) \approx O(n^{4m})$ with respect to the number of random dimensions (n). Given that for typical PC problems, $2 \leq m \leq 5$, and this suggests a near exponential scaling of the associated CPU costs.

CHAPTER 3: ANISOTROPIC POLYNOMIAL CHAOS

The previous chapter discussed several intrusive and non-intrusive methods to generate the PC expansion, and these methods were compared against each other, in terms of their individual benefits and shortcomings. It was evident from the last chapter that the linear regression approach is one of the most robust and fairly accurate method for solving uncertainty quantification problems. To recap a few advantages of this approach, it can be constructed in a black box manner, can be parallelized and can make use of commercial circuit solvers such as SPICE. But, even with all the speedup advantages that the approach demonstrates, the approach still suffers from the fact that the time and memory costs of computing the PC coefficients scales in a near exponential manner with the number of random variables in the system and the maximum degree of the gPC expansion. To address this issue, the original PC expansion was scrutinized in greater detail. It was observed that there was a lot of redundancy in the terms of the PC expansion, and that some of the terms in the expansion provided a significantly lower amount of statistical information as compared to some of the other terms in the expansion. To scrutinize the PC expansion in a greater detail, it is formulated with an alternate representation known as the high dimensional model representation (HDMR) [53]. HDMR quantifies the relative effect of each random dimension on the network response surface. A certain parameter, known as the sensitivity index, is then computed based on the contribution of each random dimension to the total response of the system [29]. A sensitivity index is always a value between zero and one. The closer the value of an index is to zero, the lower is the statistical information provided by the corresponding random dimension. The values of these sensitivity indices are used to identify and prune the statistically insignificant terms from the original PC expansion. Removal of these statistically insignificant terms further sparsifies the expansion and thus leads

to reduction in the time and memory costs required for computation of the corresponding PC coefficients. This reduction, in turn, leads to an even higher speedup as compared to any of the traditional intrusive and non-intrusive approaches. In the subsequent sections, the proposed method is described in greater detail. The validity of the proposed method is demonstrated using multiple numerical examples.

3.1 Proposed anisotropic polynomial chaos (APC) formulation

In this section, we first examine the HDMR formulation and we elaborate on the anisotropic truncation scheme. Next, we explain the method to generate the sensitivity indices and the application of the indices to create the sparse Anisotropic PC formulation. Finally, we prove the effectiveness of the proposed method using a few examples.

3.1.1 High dimensional model representation (HDMR) formulation

To recap, the traditional PC expansion is expressed mathematically as [53]

$$X(t,\lambda) = \sum_{k=0}^{P} c_k(t)\phi_k(\lambda) \quad where \quad P+1 = \binom{m+n}{m} = \frac{(m+n)!}{m!\,n!} \tag{3.1}$$

The HDMR of any output $X(t,\lambda)$ of the network is expressed as a hierarchical sum of component functions as follows.

$$X(t,\lambda) = X_0(t) + \sum_{i=1}^{N} X_i(t,\lambda_i) + \sum_{1\le i,j\le N} X_{ij}(t,\lambda_i,\lambda_j) + \cdots + X_{12\ldots N}(t,\lambda_1,\ldots,\lambda_N) \tag{3.2}$$

In the above equation, $X_0(t)$ represents the mean value of $X(t,\lambda)$. In other words, it is the value of $X(t,\lambda)$ when all the random variables are set to zero. $X_i(t,\lambda_i)$ represents the contribution of each λ_i acting alone on the total response of the network. $X_{ij}(t,\lambda_i,\lambda_j)$ represents the contribution of two random variables λ_i and λ_j acting together on the response of the system. This gives us

information about the effects of interactions of two random variables acting simultaneously on a system. Each subsequent term in the HDMR expansion of (3.2) represents the interactions, of increasing numbers of dimensions acting simultaneously, on the total response of the system. Thus, we can see that there are $N + 1$ terms in the expansion, with each term, except the mean and the final term, being a combination of several terms themselves. The number of terms in the expansion can be significantly reduced by using the Sparsity-of-effects principle. This principle states that the response of a statistical system is usually dominated by the mean and lower order interaction terms. In other words, the majority of statistical information is contained within the zeroth and the first order component function of (3.2) and the amount of information supplied by each subsequent higher order component function significantly decreases from the previous order function [23]. The zeroth and the first order functions from (3.2) are represented as

$$X_0(t) = X\left(t, \lambda^{(0)}\right) \quad and \quad X_i(t, \lambda_i) = X(t, \lambda)|_{\lambda^{(0)} \setminus \lambda_i} - X_0(t) \quad for \quad 1 \leq i \leq N \tag{3.3}$$

where the notation $\lambda^{(0)}$ denotes the case where the value of each random variable is set to zero, and the notation $\lambda^{(0)} \setminus \lambda_i$ denotes the case where each random variable except λ_i is set to zero. Each first order interaction term can be approximated, as a unidimensional PC expansion as

$$X_i(t, \lambda_i) \approx \sum_{k=1}^{m} c_i^{(k)}(t)\, \Phi_k(\lambda_i) \quad for \quad 1 \leq i \leq N \tag{3.4}$$

The 1-D coefficients can be computed by using the inner product operation as

$$c_i^{(k)}(t) = < X_i(t, \lambda_i)\, , \Phi_k(\lambda_i) > = \int_{\Omega_N} \left(X(t, \lambda)|_{\lambda^{(0)} \setminus \lambda_i} - X_0(t) \right) \Phi_k(\lambda_i) \rho(\lambda_i) d\lambda_i \tag{3.5}$$

where $\rho(\lambda_i)$ represents the marginal PDF of the random dimension λ_i. The PC coefficients from the integral in (3.5) are thus evaluated using Gaussian quadrature rules as

$$c_i^{(k)}(t) \approx \sum_{j=0}^{m} w_j X_i(t, \lambda_i) \Phi_k \left(\lambda_i^{(j)}\right) \tag{3.6}$$

w_j represents the j-th Gaussian quadrature weight corresponding to the j-th quadrature node $\lambda_i^{(j)}$. Although we need $m+1$ simulations to compute the 1-D coefficients for each i^{th} random variable, the zeroth order simulation has already been performed to calculate $X_0(t)$ and this simulation can be reused while computing $c_i^{(k)}(t)$. Thus, in total, $mN+1$ simulations are carried out to compute the 1-D coefficients in the HDMR expansion of (3.2).

3.1.2 Anisotropic truncation scheme

In case of the traditional isotropic PC expansion, the maximum degree of expansion along all random dimensions is set to the same value m, where m is the optimal degree required for the N-dimensional problem. Hence, this work calls it the isotropic PC expansion. This means that all multi-dimensional PC bases satisfy the following constraint, from (2.14)

$$\frac{1}{m}\left|\left|K\right|\right|_1 = \frac{1}{m}\left(\sum_{i=1}^{N} k_i\right) \leq 1 \tag{3.7}$$

where $K = [k_1,\ k_2, \ldots,\ k_N]$ is the vector of the 1-D PC degrees of expansion. However, a close inspection of the unidimensional PC expansion reveals that the amount of statistical information provided by each random variable varies significantly from one random dimension to the other. Essentially, the uncertainty provided by each random parameter is computed using the same degree of accuracy, irrespective of the amount of statistical information provided by that random parameter.

To correct this, the anisotropic truncation scheme is introduced. This scheme requires the maximum degree of expansion, along the i^{th} random dimension to be set to m_i, where m_i is independent of the maximum degree of expansion along the other random dimensions. The maximum degree of the polynomials used in a PC expansion directly translates to the accuracy of the PC expansion. The higher the degree of the PC expansion, the greater is the accuracy of results. By assigning a different degree of expansion to each random variable, we ensure that the random dimensions that provide statistically significant information are computed with a higher degree of expansion, and the contributions of the statistically insignificant terms are computed with a lower degree of expansion, without significantly affecting the accuracy of the total response of the system. Thus, the basis terms in an anisotropic PC expansion are a subset of the $P + 1$ terms of (2.14) with the additional constraint of

$$k_1 \leq m_1, k_2 \leq m_2, \dots, k_N \leq m_N \tag{3.8}$$

This ensures that only those multi-dimensional terms are included in the PC expansion whose degree of each 1D basis is below the corresponding maximum $m_i; 1 \leq i \leq N$. In other words, the number of bases in an anisotropic PC expansion will always be less than or equal to $P + 1$. However, since the contribution of each i^{th} random dimension has been adequately captured by the expansion of the degree m_i, the loss in anisotropic expansion is minimal.

3.1.3 Evaluation of sensitivity indices

In order to determine an optimal value of m_i, an adaptive approach is employed [29]. In this approach, the value of m_i is initially set to one and the coefficients of eq. (3.4) are evaluated using the pseudo-spectral collocation method. Once the initial coefficients are obtained, then the value of m_i is iteratively increased in increments of one. In each iteration, the new coefficients of

eq. (3.4) are evaluated using pseudo-spectral collocation method [24]. After the computation of the coefficients, the enrichment in the variance, predicted using the 1-D expansion, arising from the increment in the degree of expansion is evaluated as

$$S_i(r_i, t) = \left| \frac{\sum_{j=1}^{r+1} \left(c_i^{(j,r_i)}\right)^2 - \sum_{j=1}^{r+1} \left(c_i^{(j,r_i-1)}\right)^2}{\sum_{j=1}^{r+1} \left(c_i^{(j,r_i)}\right)^2} \right| \tag{3.9}$$

where $c_i^{(j,r_i)}$ refers to j^{th} coefficient computed during the ${r_i}^{th}$ (i.e. current) iteration. The term $\sum_{j=1}^{r+1} \left(c_i^{(j,r_i)}\right)^2$ is the variance computed from those PC coefficients. It is also noted that since the PC coefficients of eq. (3.4) are dynamic quantities, so is the enrichment of eq. (3.8). Note that the sensitivity indices thus generated are time-dependent quantities. Essentially the sensitivity index is a vector for any random variable. For a more objective assessment of the sensitivity index, the normalized integral of eq. (3.8) will be computed as

$$L_i = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} S_i(t)\, dt \tag{3.10}$$

where $t_2 - t_1$ represents the time window of simulation and the above integral can be performed using any numerical technique. This work evaluates the integral using the trapezoidal rule. For PC expansions, the numerator approximation of eq. (3.4) converges in an L_2 sense with increasing degree of expansion [49], [54]. This convergence will be manifested in the progressive decay of the value of L_i. Once the integral of eq. (3.10) falls below a prescribed tolerance ε, the iterations are halted and the maximum degree of expansion is set to $m_i = r_i - 1$. As expected, the tolerance ε is problem dependent. It is noted that the total number of SPICE simulations required to determine the optimal r_i for the i^{th} dimension using the adaptive

approach is $({r_i}^2 + r_i + 2)/2$. The total number of SPICE simulations for repeating this process for n dimensions is

$$N_0 = 1 + \sum_{i=1}^{n} \left(\frac{(r_i^2 + r_i + 2)}{2} - 1 \right) \tag{3.11}$$

If multiple network responses need to be probed, then the maximum degree of expansion m_i for the i^{th} dimension needs to be computed for all network responses separately using this approach. Then, for each dimension, the degree of expansion chosen for any random dimension λ_i is the one that has the maximum value among all the m_i generated for each network response.

3.1.4 Recovery of multi-dimensional PC coefficients

Once the degree of expansion along all random dimensions is known using the above methodology, an anisotropic PC expansion of the network response $X(t, \lambda)$ can be computed as

$$X(t, \lambda) \approx \sum_{k=0}^{Q} c_k(t) \phi_k(\lambda) \tag{3.12}$$

where the multi-dimensional basis $\phi_k(\lambda)$ is a product of the unidimensional basis as

$$\phi_k(\lambda) = \prod_{i=1}^{N} \Phi_{k_i}(\lambda_i) \ \ where \sum_{i=1}^{N} k_i \leq m_i \tag{3.13}$$

It is noted that the total number of coefficients in the anisotropic expansion of (3.12) Q is significantly smaller than that of the conventional isotropic PC expansion from (2.12) i.e. P. The order of the conventional isotropic expansion for the same problem would be

$$m = \max(m_1, m_2, \dots, m_N) \tag{3.14}$$

Next, in order to intelligently evaluate the coefficients of (3.12), instead of directly adopting the linear regression approach, first the expansion of (3.12) is separated into the mean term, the 1-D terms and the multi-dimensional interaction terms.

$$X(t,\lambda) = X_0(t) + \sum_{i=1}^{N} X_i(t,\lambda_i) + \sum_{k=1}^{Q-R} c_k(t)\phi_k(\lambda) \tag{3.15}$$

where

$$R = \sum_{i=1}^{N} m_i \leq mN \tag{3.16}$$

The mean is already known from (3.3) while the 1-D terms are also known from the 1-D PC expansions of (3.4), thereby leaving only the multi-dimensional terms to be represented as

$$\sum_{k=1}^{Q-R} c_k(t)\phi_k(\lambda) = X(t,\lambda) - X_0(t) - \sum_{i=1}^{N} X_i(t,\lambda_i) \tag{3.17}$$

The coefficients of the multi-dimensional terms of (3.17) can now be computed using the linear regression approach. In this way, the proposed approach would require evaluation of only $Q - R$ coefficients, instead of the $Q + 1$ coefficients of (3.12), thereby achieving a further reduction of $R + 1$ simulations above and beyond that already achieved due to the anisotropic feature. This reduction stems from the reuse of the 1-D expansions of (3.4) in (3.12). Thus, the total number of SPICE simulations required for the anisotropic approach is $(Q - R) + N_0$.

It is remarked that the proposed approach has two main sources of computational expense. One is the search algorithm used to generate the M multidimensional regression nodes from the tensor product grid of 1-D quadrature nodes [27], [32], [33]. The other is the costs incurred during the

SPICE simulations done to evaluate the 1-D coefficients [27], [41]. The major efficiency in the proposed approach over the conventional isotropic PC approach is largely due to the reduction in the time required for SPICE simulations. This numerical efficiency can be expressed as

$$\eta = \frac{(P+1)}{(Q-R)+N_0} = 1 + \frac{\Delta}{(Q-R)+N_0} \tag{3.18}$$

where Δ refers to additional number of bases included in the isotropic PC expansion over the APC expansion. It is always true that $\Delta \geq 0$ where the actual value of Δ is always problem dependent. This makes the proposed anisotropic PC expansion more efficient than all the works on conventional isotropic PC expansion. Furthermore, since this approach supplements other sparse PC approaches like [23], [46]-[48], it can still be incorporated into these formulations to further enhance their numerical efficiency.

3.1.5 Advantages and disadvantages of the proposed approach

Since the proposed approach itself relies on a non-intrusive approach, all the advantages of non-intrusive methods apply to the proposed method. In addition, the proposed method does not scale in a near-polynomial fashion like the conventional PC expansion and thus mitigates the shortcomings of the isotropic approach to some extent.

The proposed approach does not help for problems involving lower numbers of random dimensions. The reason for this is that for lower numbers of dimensions, the speedup achieved by an anisotropic expansion is not substantial over an isotropic expansion, but at the same time, a lot of simulations are needed to compute the 1-D coefficients, which in turn increase the time and memory costs beyond that required to generate an isotropic expansion. This will be demonstrated by example 3 in the numerical examples section that follows. The proposed approach also does not provide a substantial speedup over an isotropic approach, if for a

particular example, it is observed that a higher number of random dimensions within a network affect the network response significantly. The proposed approach works best a significant gradient is observed between the values of m_i for different random dimensions within the network, which is the case with a huge number of examples we find in the world but not all.

3.2 Numerical examples

In this section, three examples are presented to compare the accuracy and scalability of the proposed APC approach against the conventional isotropic PC approach. All relevant PC computations are performed using Matlab 2013b while the deterministic transient simulations are performed using HSPICE [63]. In particular, for example 2 and 3, the transmission line networks are modeled using W-element transmission line model provided by HSPICE which can automatically consider frequency-dependent per-unit-length parameters. The above simulations are run on a workstation with 8 GB RAM, 500 GB memory and an Intel i5 processor with 3.4 GHz clock speed.

3.2.1 Example 1: BJT low noise amplifier

In this example, the low noise amplifier (LNA) network of Fig. 3.1 is considered. This LNA utilizes an NXP BFG425W wideband BJT, which is represented as level-1 (Gummel-Poon) SPICE model. The RF input to the network is a sinusoidal wave with amplitude of 1V and a frequency of 1 GHz. The uncertainty in the network is introduced via $N = 10$ normal random variables whose characteristics are listed in Table 3.1.

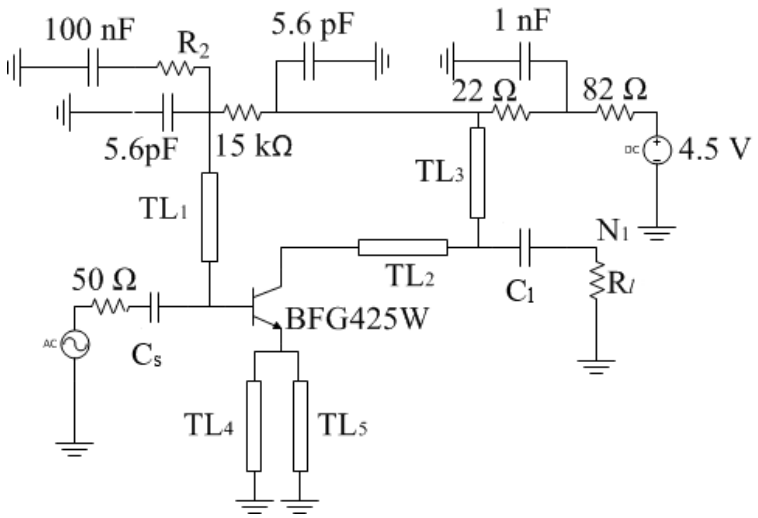


Fig. 3.1: BJT LNA network used in example 1

Table 3.1: Characteristics of Random Variables used in Example 1

No.	Random Variable	Mean	% Relative SD
1	W_1 (Width of TL_1)	0.2 mm	10
2	W_2 (Width of TL_2)	0.25 mm	20
3	W_3 (Width of TL_3)	0.3 mm	10
4	W_4 (Width of TL_4)	0.7 mm	20
5	W_5 (Width of TL_5)	0.9 mm	20
6	B_f (BJT Current Gain)	145	30
7	C_{js} (Substrate capacitance)	667.5 fF	10
8	R_l (Load resistance)	50 Ω	10
9	C_l (Load capacitance)	2.7 pF	20
10	C_s (Source capacitance)	4.7 pF	20

In order to evaluate the accuracy of the proposed approach, the mean and standard deviation of the steady-state response at the output node N_1 of Fig. 3.1 is computed using two Hermite PC expansions – the proposed anisotropic approach and the conventional isotropic expansion. For both expansions, the sparse linear regression algorithm(SPLINER) is used to determine the regression nodes [27]. For this example, the decay in the integral of eq. (3.9) with increasing degrees of expansion for each random dimension is noted in Table 3.2. For an enrichment tolerance of $\varepsilon = 0.002$, the maximum degrees of expansion along each dimension is provided in

Table 3.2. The comparison of the Mean $\pm 3\sigma$ results are shown in Fig 3.2 where the proposed APC expansion is found to exhibit good agreement with the conventional isotropic expansion.

Table 3.2: Decay of integral Example 1

Random Variable	Value of integral of (3.10) for different values of 'r'				Degree of expansion
	r=2	r=3	r=4	r=5	
1	0.0038	9.32e-6	-	-	2
2	0.0016	-	-	-	1
3	0.0122	0.0019	-	-	2
4	0.0010	-	-	-	1
5	1.10e-4	-	-	-	1
6	6.25e-4	-	-	-	1
7	0.0123	0.0024	1.28e-4	-	3
8	0.0011	-	-	-	1
9	0.0255	0.0040	0.0025	1.61e-4	4
10	0.0018	-	-	-	1

Next, in order to test the accuracy for the higher order statistical moments, the probability distribution function of the steady state response at node N_1 evaluated at the time point of

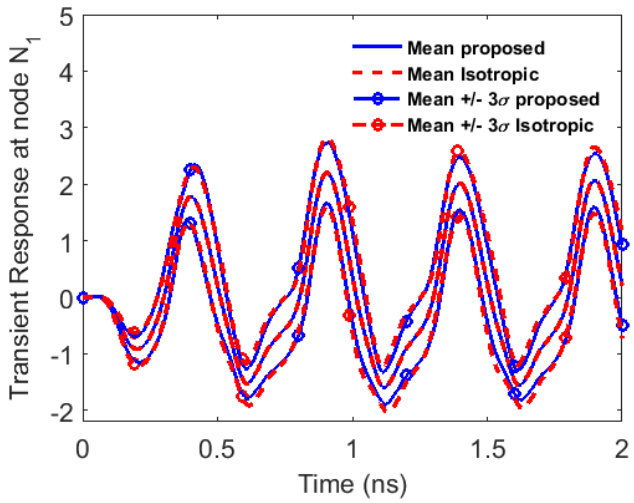


Fig 3.2: Comparison of the mean and statistical corners at node N_1

maximum standard deviation (t = 0.99 ns) is computed using the above two approaches and the results are displayed in Fig 3.3. As expected, the probability distribution results for the APC expansion exhibit good agreement with that of the conventional isotropic expansion.

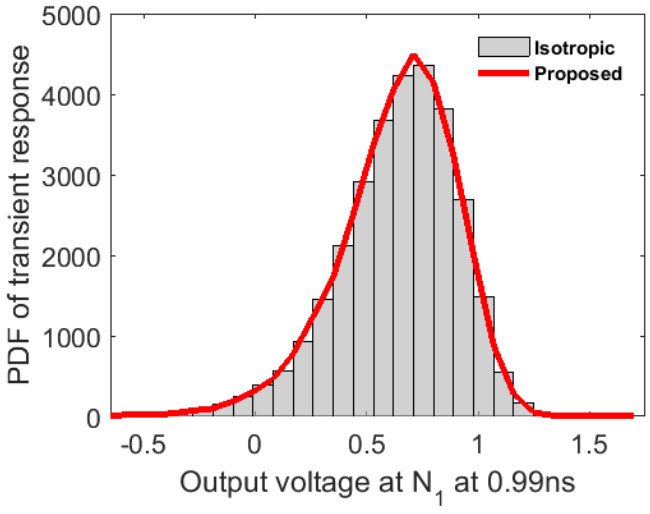


Fig 3.3: PDF of transient response at node N_1 at the time point of maximum standard deviation

Finally, it is noted that the anisotropic expansion requires 1933.42 seconds for the 661 SPICE simulations. This includes $N_0 = 83$ simulations and $Q - R = 579$ SPICE simulations. On the other hand, the isotropic expansion requires 2927.93 seconds to perform the necessary $P + 1 = 1001$ SPICE simulations. This amounts to a reduction of $\Delta = 340$ SPICE simulations for this example.

The current example only consisted of 10 random variables. The beauty of the proposed method is that, as the complexity of the problem increases in terms of number of random variables and the order of expansion required, we see that the savings in terms of time and computation costs also increases. This is demonstrated by the next example.

3.2.2 Example 2: Multi-conductor transmission line

In this example, the proposed approach and the conventional approach are compared for a large distributed network. For this purpose the 16 multi-conductor transmission line (MTL) network loaded with SPICE level-49 CMOS inverters, as shown in Fig 3.4 is considered. The layout of the stripline network is also illustrated in Fig 3.4. The input to the network is a trapezoidal

waveform with rise/fall time T_r= 0.1 ns, pulse width T_w = 0.5 ns and amplitude equal to 1V. The uncertainty in the network is introduced via N = 13 normal random variables whose characteristics are listed in Table 3.3.

For this example, the decay in the integral of eq. (3.9) with increasing degrees of expansion for nodes N_1 and N_2 is noted in Table 3.4 and Table 3.5 respectively. In order to establish the accuracy of the proposed approach, the mean and standard deviation of the steady-state response at the output node of Fig. 3.4 is computed using two Hermite PC expansions – the proposed anisotropic approach and the conventional isotropic expansion. As before, the SPLINER algorithm is used to identify the regression nodes for both expansions [27]. The comparison of the statistical results using the two approaches is shown in Fig 3.5 where the proposed APC approach is found to exhibit good agreement with the conventional isotropic APC approach.

Furthermore, to test the accuracy for higher order statistical moments, the probability distribution function of the transient response at node N_1 and N_2 evaluated at the time point of maximum standard deviation (t = 5.58 ns and t = 1.44 ns respectively) is computed using the above two approaches and 30,000 MC samples. The corresponding results are displayed in Fig. 3.6. As expected, the probability distribution results for the APC expansion exhibit good agreement with that of the conventional isotropic expansion.

It is appreciated that the proposed APC approach requires 14781.62 seconds for performing the 1627 SPICE simulations. This includes N_0 = 112 simulations and $Q - R$ = 1515 SPICE simulations. On the other hand, the isotropic expansion requires 21622.78 seconds to perform the necessary $P + 1$ = 2380 SPICE simulations. This amounts to a reduction of Δ = 753 SPICE simulations for this example.

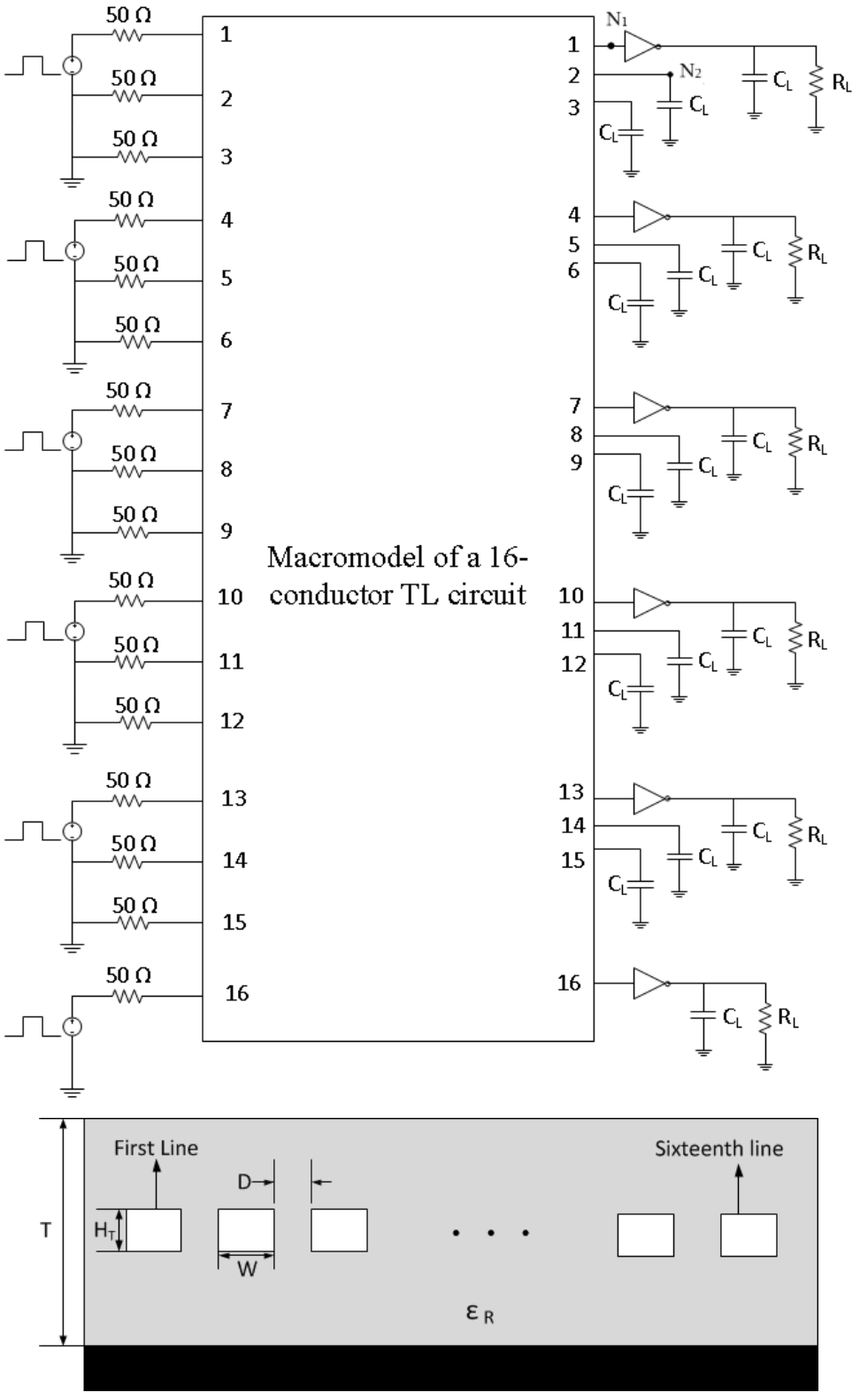


Fig 3.4: Multi-conductor Transmission Line Network used in Example 1 and 2

Table 3.3: Characteristics of random variables of transmission line network

No.	Random Variable	Mean	% Relative SD
1	W(Line Width)	0.15 mm	10
2	H_T(Line Height)	0.03 mm	10
3	T (Height of dielectric)	0.45 mm	10
4	D (Separation between lines)	0.15 mm	10
5	ε_r (Relative Permittivity)	4.1	15
6	L (Length of line)	6 cm	15
7	G (Conductance of line)	5.8e+7 S/m	10
8	PL (Channel length PMOS)	0.1 μm	10
9	PW (Channel Width PMOS)	10 μm	10
10	NL (Channel length NMOS)	0.1 μm	10
11	NW (Channel width NMOS)	10 μm	10
12	R_l (Load resistance)	1.5 kΩ	5
13	C_l (Load capacitance)	1pF	5

Table 3.4: Decay of integral of (3.10) with increasing degrees for node N_1 of example 2

Random Variable	Value of integral of (3.10) for different values of 'r'				Degree of expansion
	r=2	r=3	r=4	r=5	
1	0.0380	0.0072	0.0113	0.0069	4
2	8.56e-4	-	-	-	1
3	1.66e-4	-	-	-	1
4	0.0228	0.0039	-	-	2
5	0.0264	0.0363	0.0048	-	3
6	0.0883	0.0315	0.0118	0.0043	4
7	8.84e-4	-	-	-	1
8	0.0044	-	-	-	1
9	0.0039	-	-	-	1
10	0.0041	-	-	-	1
11	0.0035	-	-	-	1
12	6.7e-7	-	-	-	1
13	0.0105	0.0021	-	-	2

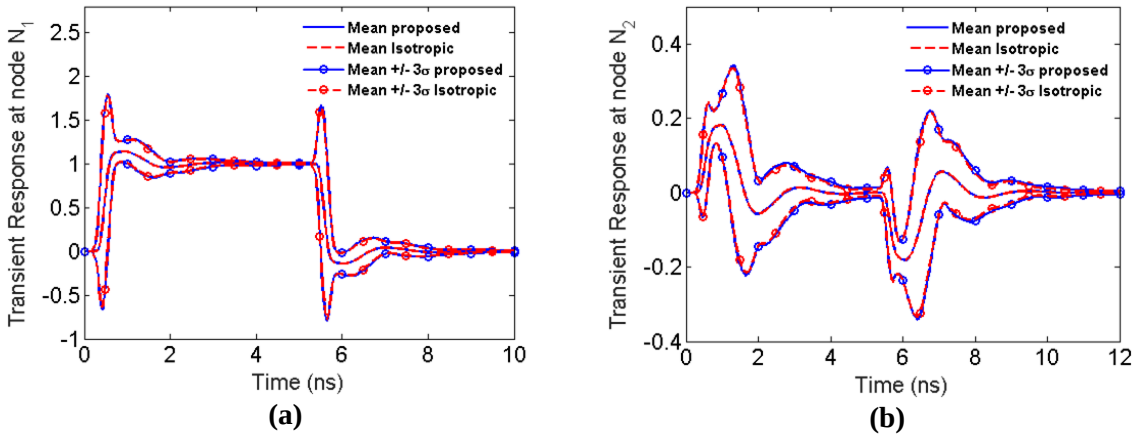
3.2.3 Example 3: Test of scalability

In this example, the same MTL network of example 2 is considered. The objective of this example is to demonstrate the superior scalability of the proposed APC approach over the conventional isotropic PC approach as the number of random dimensions increases.

To demonstrate the numerical efficiency offered by the proposed APC approach, the number of random dimensions is progressively increased from 3 to 13 as shown in Table 3.6. For each test case, the proposed APC expansion and the isotropic PC expansion are implemented and the

Table 3.5: Decay of integral of (3.10) with increasing degrees for node N_2 of example 2

Random Variable	Value of integral of (3.10) for different values of 'r'				Degree of expansion
	r=2	r=3	r=4	r=5	
1	0.0376	0.0080	0.0108	0.0068	4
2	2.35e-4	-	-	-	1
3	3.79e-4	-	-	-	1
4	0.0349	0.0026	-	-	2
5	0.0128	0.0377	0.0073	-	3
6	0.0877	0.0378	0.0125	0.0049	4
7	2.66e-4	-	-	-	1
8	0.0041	-	-	-	1
9	0.0028	-	-	-	1
10	0.0020	-	-	-	1
11	0.0025	-	-	-	1
12	5.08e-8	-	-	-	1
13	0.0140	0.0025	-	-	2

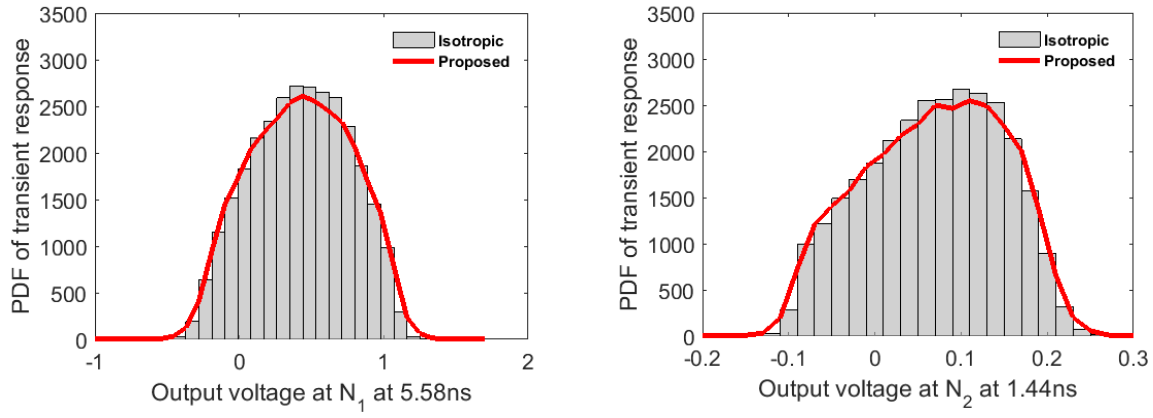


(a) **(b)**

Fig. 3.5(a): Comparison of the mean and statistical corners at node N_1.

(b): Comparison of the mean and statistical corners at node N_2.

incurred time cost is noted in Table 3.6. These incurred costs include both the cost to extract the regression nodes using non-intrusive stochastic testing as well as the cost to perform the required SPICE simulations. The scaling of the CPU costs from Table 3.6 is illustrated in Fig. 3.7 and demonstrates the savings in CPU time achieved by the proposed approach. This improvement in scalability with increasing number of random dimensions is what is theoretically expected.



(a) **(b)**

Fig. 3.6(a): PDF of the transient response at node N_1.

(b): PDF of the transient response at node N_2.

Table 3.6: Scaling of CPU time costs using proposed APC and conventional isotropic PC approach

Random Variable	Isotropic		Proposed	
	Number of simulations	Cost of simulations (secs)	Number of simulations	Cost of simulations (secs)
3	35	317.98	77	699.56
5	126	1144.74	170	1544.48
7	330	2998.12	306	2780.07
9	715	6495.92	553	5024.12
11	1365	12401.3	968	8794.47
13	2380	21622.78	1627	14781.62

3.3 Anisotropic hyperbolic PC formulation

In this section, the anisotropic truncation scheme is combined with a hyperbolic truncation scheme to create a hybrid which results in an even sparser PC expansion and thus provides additional speedup over each of the individual methods. The basis of the hyperbolic polynomial chaos truncation scheme is explained and the approach is then combined with the anisotropic PC approach. Finally, we evaluate the accuracy of the hybrid approach with a numerical example.

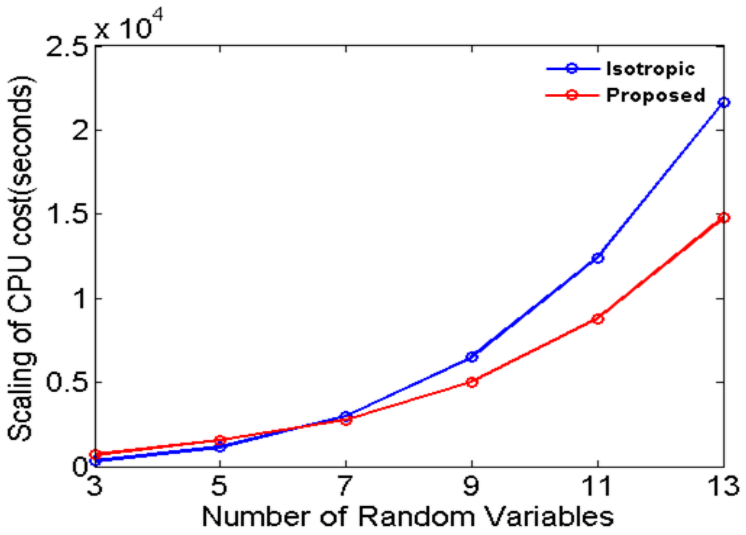


Fig 3.7: Scaling of overall CPU time costs for the proposed approach and conventional isotropic approach in Example 3 with increasing number of random dimensions

3.3.1 Basis of hyperbolic polynomial chaos (HPC) formulation

Recalling the sparsity-of-effects principle, it states that in a stochastic network the main effect and the lower order interactions primarily impact the response of a network and that most of the higher order interactions provide relatively insignificant impact to the total response of a network. Here, the primary change caused in the response by the individual change in a random variable is called a $main\ effect$, and the variation caused in impact of a random variable by changes in other variables is called an $interaction$. This phenomenon is validated through several examples in the literature [23], [45].

In a conventional isotropic PC expansion, there are $(m+1)^n$ possible combination of terms but the PC expansion is truncated to $P+1$ terms. Using the HDMR principle, the $P+1$ term PC expansion is further truncated by selecting a lower degree of expansion for the statistically insignificant random dimensions. Now, using the sparsity-of-effects principle, it is possible to prune more polynomial bases without a major loss in accuracy. In order to reduce number of polynomial bases with loss of accuracy, the works of [23] and [45] suggest considering the rank

of the polynomials, where rank is defined as the number of dimensions in a polynomial basis, for example $\phi_{[1,1,1]}(\lambda) = \Phi_1(\lambda_1)\Phi_1(\lambda_2)\Phi_1(\lambda_3)$ is a rank 3 polynomial base while $\phi_{[2,0,0]}(\lambda) = \Phi_2(\lambda_1)$ is a rank 1 polynomial base.

The understanding of ranks and sparsity-of-effects is helpful to note that the lower rank bases always contribute the most significant amount of information. Further, within the higher rank terms themselves, the terms which use a lower degree of unidimensional polynomial bases used still contributes more information than the terms which use a higher degree of unidimensional polynomial bases. Using this knowledge, it is possible to intelligently select and prune those terms from a PC expansion which have a higher rank and a lower statistical contribution to the response of a network. This is the basis of HPC truncation scheme.

If degrees of the expansion of the unidimensional polynomial bases that make up a PC expansion are denoted by $d_1, d_2, \dots, d_n$, then we know that a conventional isotropic expansion is guided by the constraint that

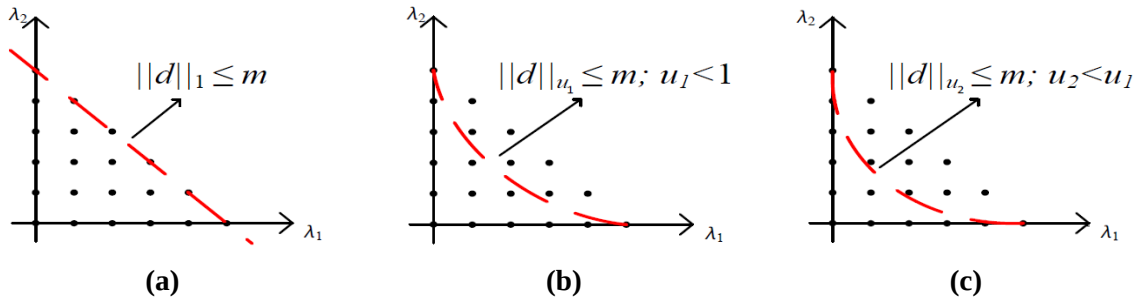
$$d_1 + d_2 + \dots + d_n \leq m \tag{3.19}$$

where m is the maximum degree of the PC expansion. In other words, the constraint is on an L_1 norm. In HPC, the constraint is put on the L_u^{th} norm of the indices vector $d = [d_1, d_2, \dots, d_n]$, where $u \leq 1$ and the L_u^{th} norm is defined as

$$\left|\left|d\right|\right|_u = (d_1^u + d_2^u + \dots + d_n^u)^{\frac{1}{u}} \leq m \tag{3.20}$$

u is called the hyperbolic factor. When u = 1, then the HPC expansion is identical to the conventional isotropic PC expansion. Different values of the factor u result in different number of polynomial bases being selected. For illustration purposes, an example with N = 5 and m = 4

is considered and Fig. 3.8a depicts the case of an isotropic expansion where Fig. 3.8b and 3.8c depict the case of an HPC expansion for two different values of u. A larger value of u results in an expansion having greater accuracy but also requires a higher computational time and memory cost. A smaller value of u trades off accuracy for a smaller expansion requiring comparatively lower computation time and memory cost. Typically, we find that a value of u which is close to 0.6 or 0.7 finds a good balance between the loss in accuracy and the amount of speedup achieved.



(a) **(b)** **(c)**

Fig 3.8: Graphical representation of the hyperbolic truncation scheme using a 2D example (N=2, m=5) showing **(a)** the traditional isotropic truncation scheme and **(b,c)** the effect of decreasing the hyperbolic factor on the proposed approach, from left to right. Image used from [64].

3.3.2 Adding anisotropicity

In our case though, since we add anisotropicity to the expansion, there would be an additional constraint on the indices that

$$d_1 \leq m_1, d_2 \leq m_2, \ldots, d_N \leq m_N \tag{3.21}$$

It is emphasized that since the different degrees of eq. (3.20) have been deemed adequate to capture the impact of the corresponding random dimension acting alone, the loss in accuracy due to the extra anisotropic feature is marginal [42]. The overall sparse PC expansion arising from the anisotropic HPCE can be formulated as

$$X(t,\lambda) = \sum_{k=0}^{Q} c_k(t)\phi_k(\lambda) \tag{3.22}$$

where $Q \ll P$. The sparse coefficients of eq. (3.22) can be computed using any intrusive or non-intrusive approach, although this work relies on either a non- intrusive linear regression approach or a non-intrusive stochastic testing approach.

3.3.3 Advantages and disadvantages of anisotropic HPCE

Anisotropic HPCE shares all the advantages of the anisotropic method mentioned earlier. The advantage of the hybrid method is that it provides an even greater reduction in terms of number of coefficients to be computed, with only a minimal loss of accuracy.
The disadvantage of the hybrid method or the HPCE method in particular is that, for high dimensional problems, the HPCE method sometimes ends up pruning out a lot of information that is still significant to the response of the network. This is because in the case of high-dimensional problems, the number of terms between two separate values of u goes on increasing rapidly. That leads to a high amount of pruning that can compromise the accuracy of the generated PC expansion beyond a certain level of tolerance.

3.3.4 Example 4: BJT LNA

In order to validate the proposed approach, the LNA network of Fig. 3.1 is considered. The input signal is a sinusoid with an amplitude of 1V and frequency of 2 GHz. The uncertainty in the network is introduced via N = 10 normally distributed random variables described in Table 3.2. The lengths of the transmission lines $TL_1, TL_2, TL_3, TL_4, and\ TL_5$ in Fig. 3.1 are 8.9 mm, 3.9 mm, 6.6 mm, 3.0 mm, and 3.0 mm respectively. The transmission lines are copper microstrip traces with thickness 2 μm located on top of a dielectric plate of thickness 0.5 mm and relative permittivity of 4.6 [42].

For this example, both the full-blown PC expansion and the anisotropic HPCE approaches are implemented. The maximum degree of expansion required is m = 4. For the full-blown PC approach, a total of $P + 1$ = 1001 basis terms, or in other words a total of 1001 full model SPICE simulations of the network of Fig 3.1 are required. On the other hand, the proposed anisotropic HPCE approach, used u = 0.79 with the anisotropic degrees of expansion listed in Table 3.2. Overall, the hybrid approach required only 284 SPICE simulations. Thus, for this example, the proposed approach achieves a CPU speedup of roughly 3.5 times over the full-blown approach. As expected, the accuracy of the proposed approach with that of the full-blown approach is exhibited as shown in Fig. 3.9.

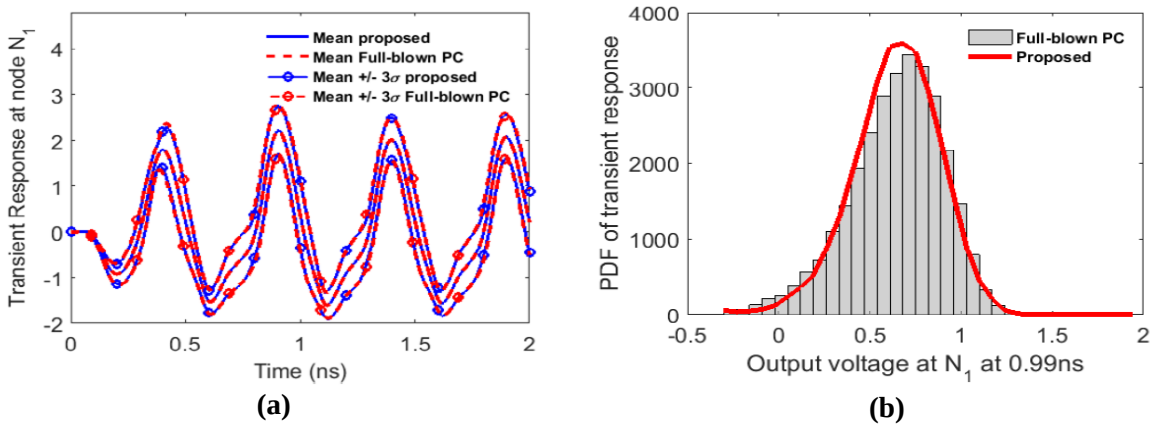


Fig 3.9: Statistics of the LNA network of Fig. 3.1 evaluated using conventional isotropic method and the proposed anisotropic HPCE method **(a)** Statistical analysis of the transient response at node N_1. **(b)** PDF of transient response at node N_1 at the time point of maximum standard deviation. [42]

Comparing with the anisotropic results generated for the same example earlier in section 3.2.1, we can see that the anisotropic approach required 661 SPICE simulations while the hybrid approach required only 284 SPICE simulations which translates to a speedup of a little over 2.

CHAPTER 4: ANALYSIS OF EPISTEMIC UNCERTAINTY

So far, this work has talked about intrusive and non-intrusive approaches to model parametric uncertainty. The linear regression approach was primarily chosen to do so and this work also elaborated on methods to reduce the time and memory costs of the computation necessary to create the PC metamodels. But so far, the random variables are themselves assumed to be known and that their distribution is assumed to be one of the standard distributions. When the information about the parameters is entirely known and the uncertainty in the problem reduces to only the randomness in the values of the parameters themselves, it is known as an aleatory variability problem. But this is not the only kind of uncertainty that exists. This chapter elaborates on a separate kind of uncertainty known as the epistemic uncertainty, and how this kind differs from the aleatory kind of uncertainty [65]-[67], which has been discussed extensively so far. Traditionally the Monte Carlo approach has been used to model epistemic uncertainty. This chapter briefly presents some of the existing techniques used to analyze epistemic uncertainty and then talks about the proposed method used to do the same. The chapter then develops a hybrid method based on the application of a dimension reduction technique to the proposed method to get a huge amount of reduction to the time and memory costs of computation of the PC expansion. Finally the accuracy, of the proposed method, including the hybrid method, is evaluated using a few examples.

4.1 Types of uncertainty

Predominantly, there are two kinds of uncertainties. To understand these uncertainties better, let us consider an example of rolling a dice [65]. Assuming that no information about the dice is given, one can assume the dice to be either a fair dice or a loaded dice. Next, let's say the dice is

rolled four times and the numbers that come up are 2, 3, 3 and 4. Now, if asked to model the dice, one can approach the problem in two ways. As one approach to modelling the dice, the previous observations can be disregarded and the dice can be modeled using a uniform distribution between 1 and 6, where the number only holds an integer value. The second approach to developing a model can be purely empirical. Given the four observations are only between two and four, one can also start with an assumption that this is a loaded die where the number 3 occurs more often, 2 and 4 a little less so, and the rest of the numbers have the lowest probability of occurrence. As more and more rolls of the dice are placed, we gain more information about the outcomes of the dice and thus, can even more accurately predict the probabilities of the outcomes of the six faces of the dice.

The first approach treats the problem as an aleatory uncertainty problem. The term aleatory comes from the Latin word $Alea$ meaning a dice [66]. Aleatory random variables are those variables whose probability distribution functions are known. Generating more samples of aleatory random variables do not add to the information provided by these samples. Essentially, modeling aleatory random variables is a based on a frequentist approach to solving uncertainty problems. A frequentist approach states that given an infinite number of samples, the value of the random variable results in a limited set of outcomes which is guided by a probability distribution function [67].

The second approach treats the problem as an epistemic uncertainty problem. The term epistemic comes from the Greek word $episteme$ [66], which means knowledge. Epistemic uncertainty is the scientific uncertainty which arises due to limited data and knowledge. The PDF of an epistemic variable is not completely known. Instead of being modeled by probabilities in the frequentist sense, epistemic variables are modeled by what is called confidence intervals.

Confidence intervals are probabilities as well, but as opposed to probabilities of aleatory random variables which are frequentist in nature, the probabilities of epistemic variables are Bayesian probabilities [67]. In Bayesian statistics, the probability of a proposition simply represents a degree of belief in the truth to that proposition. This degree of belief is nothing but the confidence interval mentioned earlier, which can be updated as more and more information is added to the model created.

Of the two, aleatory uncertainty can be modeled with ease while epistemic uncertainty cannot. So, a good way to model epistemic uncertainty is to convert the epistemic uncertainty into an aleatory uncertainty and then analyze it. While traditionally, the brute force MC method has been used to model epistemic uncertainty, it is severely time and computationally expensive. The following sections elaborate more on the methods used to tackle epistemic uncertainty.

4.2 Methods to deal with epistemic uncertainty

In case of epistemic uncertainty, statistical assessment of the network response is not directly possible since there is limited information regarding the PDFs of the epistemic dimensions. To address this problem, the worst-case bounds of the network response is extracted to understand the spread of uncertainty in the network response. The methodologies to propagate parametric uncertainty for worst case analysis can be broken into two forms – the probabilistic approaches and the interval analysis approaches. The probabilistic approaches ascribe a uniform PDF to the input parameter so that selecting any value of the parameter can be done without any penalty. Thereafter, methods such as Monte Carlo or surrogate models such as PC metamodels [7]-[42], Gaussian process models [68], and radial basis function models [69] are used to propagate the parametric uncertainty to the network responses. By generating a large number of samples of the input parameter space, these approaches can be used to generate an ensemble of the response.

From this ensemble of responses, the maximum and minimum bounds of the response can be determined.

On the other hand, interval methods, as the name suggests, use the knowledge of the maximum and minimum bounds of the input parameter to evaluate the induced bounds of the response surface without the need of assuming a uniform PDF [70]. Among the interval methods, interval arithmetic and affine arithmetic are the common approaches for propagating parametric uncertainty for worst case analysis [71]-[73]. More recently, the Taylor models have also been developed for worst case analysis [74], [75]. This thesis however, uses the Fuzzy logic approach of [76], [77] to quantify epistemic uncertainty. The fuzzy logic approach is explained in greater detail in the following sections.

4.3 Fuzzy logic approach

Each epistemic dimension is described as a pure interval of possible values. However, for many problems, additional information such as the likelihood of a dimension to assume values in a particular sub-interval within the larger interval of support may be elicited from expert or prior knowledge regarding the model. One way to embed this likelihood information into the problem is via a fuzzy logic framework. Specifically, a fuzzy logic framework attempts to denote the sub-intervals of support corresponding to a particular confidence level as an alpha cut (α-cut) made to the membership function of the dimension. Thereafter, the subinterval of support at each α-cut needs to be propagated to the response surface – in other words, the worst case analysis of the network needs to be performed at each α-cut. Once the induced subinterval of support of a network response are known at each α-cut, this information will be collated to describe the membership function of the same response. This response membership function reveals the

confidence of the response surface to assume values within a specific subinterval – information that is not available in traditional worst case analysis.

Unfortunately, both conventional probabilistic and intervals methods are unsuitable for performing worst case analysis using fuzzy frameworks. This is because for probabilistic methods, the number of SPICE simulations required to evaluate the PC or other metamodels scale in a near-exponential manner with the number of epistemic dimensions and the number of α-cuts. Moreover, while sparse and reduced dimensional representations of PC has been developed for purely aleatory problems, they are not viable for epistemic UQ [23], [48]. This is because the reported methods of [23], [48] utilize the statistical information of the network response to decide which bases/dimensions to remove. However, due to lack of knowledge regarding the PDFs of epistemic dimensions, no statistics of the response surface can be mathematically defined. As a result, probabilistic methods are not applicable for high-dimensional problems. For interval analysis methods, because the subinterval of support changes for each α-cut, these analyses have to be performed for each α-cut anew. This again leads to very slow worst case analysis for large number of α-cuts.

In this paper, a reduced dimensional PC based probabilistic approach for the efficient worst case analysis for very high-dimensional networks have been developed. The key feature of this approach is the ability to utilize the HDMR formulation to directly quantify the impact of each epistemic dimension on the maximum and minimum bounds of the response surface. This impact factor or the sensitivity of the network response on each epistemic dimension is mathematically described as the shrinkage in the area enclosed by the maximum and minimum bounds of the response surface as each dimension is rendered inactive. Such sensitivity indices are applicable for only epistemic problems since they do not involve the computation of statistical moments

such as variance/covariance/PDFs of the response as suggested in the previous chapter. Similar to the previous chapter, an appropriate threshold is selected and those dimensions which have a sensitivity index below the threshold can be safely discarded, without significant loss in accuracy. Performing the PC expansion on the resultant low-dimensional random subspace will lead to the recovery of a very sparse set of coefficients with negligible loss of accuracy. Evaluating these sparse set of coefficients will scale much more favorably than conventional full-blown PC expansions. The above methodology can be easily adapted when considering multiple responses of interest.

To solve a problem involving epistemic uncertainty using fuzzy logic approach, it is first necessary to develop the PC techniques for the worst case analysis of the problem's simulations and then to develop a fuzzy framework based on the worst case analysis.

4.3.1 Worst case analysis

Consider a general interconnect network where the epistemic parameters are described as

$$e_{min,i} \leq e_i \leq e_{max,i} \tag{4.1}$$

where $[e_{min,i}, e_{max,i}]$ represent the closed interval of possible values the epistemic parameter e_i can assume. For a probabilistic representation, each epistemic parameter of (4.1) is modeled as a random parameter with a uniform PDF as

$$e_i = \frac{\left(e_{max,i} + e_{min,i}\right)}{2} + \frac{\left(e_{max,i} - e_{min,i}\right)}{2}\lambda_i \tag{4.2}$$

where $\lambda_i \in [-1,1]$. The set of random variables $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_N]$ located within the hypercube $[-1,1]^N$ represents the entire epistemic uncertainty of the problem. The behavior of the overall network can be characterized by the stochastic modified nodal analysis (MNA) equations as

$$G(\lambda)X(t,\lambda) + C(\lambda)\frac{dX(t,\lambda)}{dt} + F\big(X(t,\lambda)\big) + \sum_{i=1}^{N_t}(T_i Y_i(t,\lambda)T_i^T) * X(t,\lambda) = B(t) \qquad (4.3)$$

where G, C matrices contain the stamp of all the memoryless and memory lumped circuit elements respectively, X is the vector of stochastic voltage/current responses, F contains the stamp of nonlinear circuit elements, T_i is the selector matrix mapping the vector of port currents $i_i(t)$ for the i^{th} distributed network into the nodal space of the circuit, Y_i is the corresponding time-domain Y-parameter macro-model of the i^{th} distributed network, B represents the input vector of independent voltage and current sources, and '$*$' denotes the temporal convolution which is performed in a recursive manner in SPICE.

The goal of worst case analysis is to evaluate the maximum and minimum bounds of the response described as

$$x_{min,j}(t) \leq x_j(t,\lambda) \leq x_{max,j}(t) \qquad (4.4)$$

where $[x_{min,j}(t), x_{max,j}(t)]$ represent the dynamic bounds enclosing the possible variations of the j^{th} response $x_j(t,\lambda)$. In order to do so, the response $x_j(t,\lambda)$ is first expressed as a PC expansion as

$$x_j(t,\lambda) = \sum_{k=0}^{P} c_{j,k}(t)\phi_k(\lambda) \qquad (4.5)$$

where $\phi_k(\lambda)$ is the k^{th} degree Legendre polynomial, $c_{j,k}(t)$ is the corresponding coefficient, and the number of terms in the expansion of (4.5) is truncated to $P+1 = (n+m)!/(n!\,m!)$, m being the maximum degree of the expansion. The coefficients of (4.5) can be evaluated using

intrusive or non-intrusive approaches. Once the coefficients are known, the metamodel of (4.5) becomes a closed form surrogate of the network response. By choosing an exhaustive set of Monte Carlo samples $\lambda^k = \left[{\lambda_1}^{(k)}, {\lambda_2}^{(k)}, \dots, {\lambda_N}^{(k)}\right], 1 \le k \le M$ and replacing them in the metamodel of (4.5) will yield a large ensemble of the network response. Provided the M set of points provide a dense coverage over the hypercube space $[-1, 1]^N$, the maximum and minimum bounds $[x_{min,j}(t), x_{max,j}(t)]$ can be extracted from this ensemble of the response.

4.3.2 Fuzzy framework for worst case analysis

In the above subsection, no additional information regarding the likelihood of the epistemic parameter e_i assuming values in different subintervals within the interval of support $[e_{min,i}, e_{max,i}]$ has been considered. However, if known, then it is necessary to include this likelihood or confidence information in the UQ as it will reduce the subjectivity of the worst case analysis. One methodology to include this information is by using a fuzzy logic framework.

In a fuzzy framework, the confidence that the epistemic parameter e_i assumes values in any subinterval within the support $[e_{min,i}, e_{max,i}]$ is described using a membership function as shown in Fig. 4.1. The shape and characteristics of the membership function is often elicited using either expert knowledge of the network or from additional experiments/simulations. In Fig. 1, the parameter is shown to exhibit a triangular membership function although other shapes such as trapezoidal, Gaussian, or nonlinear triangular shapes are also possible.

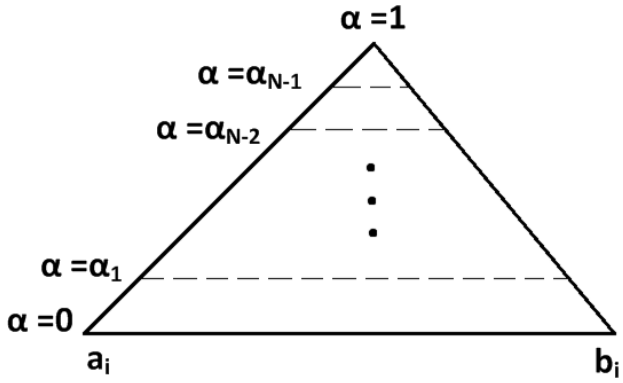


Fig 4.1: Typical Input Membership Function

Once the membership function is known, in order to assign different confidence levels to e_i, the membership function is next cut at different horizontal levels (or α-cut levels) as shown in Fig. 4.1. For any general j^{th} α-cut level α_j, the corresponding subinterval of support is $C_j^{(i)}$ as indicated in Fig. 1. This range $C_j^{(i)}$, also known as the confidence interval, encloses the set of possible fuzzy values e_i can assume with a confidence of $1-\alpha_j$ [76]. For a general multidimensional problem, the j^{th} confidence interval will be the hyperrectangle $C_j^{(1)}$ x $C_j^{(2)}$ … x $C_j^{(N)}$. Now, the worst-case analysis has to be performed for each α-cut separately and the corresponding induced confidence level of the network response evaluated. By stitching together, the knowledge of the response confidence level for each α-cut, the membership function of the network response can be determined. This membership function can be used to determine the worst-case response for any given value of the α-cut or confidence level. It is noted that the more number of α-cuts there are, greater will be the resolution of the response membership function.

It is observed that whether using probabilistic or interval methods for propagating the confidence levels, they have to be performed anew for each α-cut. A more efficient approach for probabilistic methods have been reported [78]. In these works, rather than laboriously constructing a new PC metamodel for each α-cut, it is posited that a solitary global PC metamodel be constructed for the overall interval of support (i.e., for $\alpha = 0$) [78]. This is because from Fig. 1 it is clear that the confidence interval for any epistemic parameter and any α-cut will always fall within the overall interval of support. Thus, the metamodel for $\alpha = 0$ can be directly reused for all the higher α-cuts. Nevertheless, the main drawback of such an approach is the fact that as the number of epistemic dimensions increases (i.e., *N* increases), the number of SPICE simulations required to non-intrusively evaluate the PC coefficients for $\alpha = 0$ still increases in a near exponential manner. Furthermore, as explained in the previous section, the well-known sparse or reduced dimensional PC representations developed for purely aleatory problems cannot be used for purely epistemic problems. This is because all of these works depend on statistical information, usually the variance information of the network response, to objectively determine which bases or dimensions are least important and can be removed. However, by definition, purely epistemic problems do not involve the knowledge of the PDF and hence statistical information to guide the recovery of a sparse PC metamodel is unavailable. To address this poor scalability with respect to the number of epistemic dimensions, a non-statistical approach based on global sensitivity analysis is developed in this paper.

4.4 Proposed reduced dimensional PC approach

Since the sensitivity indices cannot be developed in the same manner as those developed for aleatory random variables, due to the lack of information about the PDF of the random variables, a new methodology is needed to analyze the sensitivity indices of the epistemic random

variables. However, there are two major challenges to achieving this objective. First, a forward model needs to be identified which can be used to determine the relative impact of each epistemic dimension on the response. This forward model must be efficient enough to limit the overhead costs in performing the global sensitivity analysis yet accurate enough to accurately capture the relative impact of each dimension. Second, the impact factors (or sensitivity indices) must be global in the sense that they should capture the impact of the dimension over the entire [-1, 1] interval of support rather than being based on derivative information which captures only the local impact of the dimension at any point. These challenges will be addressed in the next few subsections.

4.4.1 Forward model for global sensitivity analysis

In the previous chapter, it has been shown that considering only the isolated effects of the aleatory dimensions is sufficient for accurately computing the relative sensitivity indices in high-speed interconnect networks. This claim is supported by two reasons. First, the sparsity of effects principle claims that the lower order interactions between the random dimensions are richer in their information content than the higher order interactions. This means that considering the isolated effects of the uncertain dimensions (i.e., the first order terms) can still yield an accurate relative measure of the impact of each dimension if not the accurate total impact. Second, since at this stage only the relative impact of the uncertain dimensions is required as opposed to their accurate absolute vales, neglecting the higher order interactions between the dimensions does not bias the identification of the unimportant dimensions.

Based on the above reasoning, in this method too, the forward model to be used for global sensitivity analysis consists of only the isolated effects of the epistemic dimensions, mathematically expressed as

$$x(t,\lambda) = x_0(t) + \sum_{i=1}^{N} x_i(t,\lambda_i) \tag{4.6}$$

where

$$x_0(t) = x\left(t,\lambda^{(0)}\right) \quad and \quad x_i(t,\lambda_i) = x(t,\lambda)|_{\lambda^{(0)}\setminus\lambda_i} - x_0(t) \quad for \quad 1 \le i \le N \tag{4.7}$$

where the notation $\lambda^{(0)}$ denotes the case where the value of each random variable is set to zero, and the notation $\lambda^{(0)}\setminus\lambda_i$ denotes the case where each random variable except λ_i is set to zero. The first function of (4.6) represents the response in absence of all epistemic dimensions while the second function of (4.6) represents the contribution of each epistemic dimension on the network response acting alone. These contributions are then described using one dimensional (1D) PC expansions as

$$x_i(t,\lambda_i) \approx \sum_{k=1}^{m} c_i^{(k)}(t)\,\Phi_k(\lambda_i) \quad for \quad 1 \le i \le N \tag{4.8}$$

The coefficients of (4.8) satisfies the inner product operation

$$c_i^{(k)}(t) = < x_i(t,\lambda_i)\,, \Phi_k(\lambda_i) > = \int_{\Omega_N} \left(x(t,\lambda)|_{\lambda^{(0)}\setminus\lambda_i} - x_0(t)\right) \Phi_k(\lambda_i)\rho(\lambda_i)d\lambda_i \tag{4.9}$$

where $\rho(\lambda_i)$ represents the marginal uniform probability density function of λ_i. The integral of (4.9) is thereafter approximated using Gaussian quadrature rules as

$$c_i^{(k)}(t) \approx \sum_{j=0}^{m} w_j\, x_i(t,\lambda_i)\Phi_k\left(\lambda_i^{(j)}\right) \tag{4.10}$$

where w_j represents the j^{th} 1D Gaussian quadrature weight corresponding to the j^{th} quadrature node $\lambda_i^{(j)}$. From (4.9) and (4.10) it is noted that only $(m+1)N$ deterministic SPICE simulations are required to find the PC coefficients of (4.8), thereby ensuring that the forward model can be constructed very efficiently even for high-dimensional problems.

4.4.2 Global sensitivity analysis in epistemic problems

In purely epistemic problems, the imprecision in the network responses can be well represented by finding the maximum and minimum bounds $[x_{min,j}(t), x_{max,j}(t)]$ of (4.4). Thus, in this paper, the impact of each dimension is measured on the response bounds as opposed to the response variance computed in the earlier chapter.

For this purpose, let us assume α = 0 and the PC representation of the forward model of (4.6) is known. Now, let an exhaustive set of Monte Carlo samples, $\lambda^k = \left[{\lambda_1}^{(k)}, {\lambda_2}^{(k)}, \dots, {\lambda_N}^{(k)}\right], 1 \leq k \leq M$ be chosen and replaced in (4.6). Thereafter, at any time point, the dynamic bounds of the network response can be evaluated as

$$U_{min,0}(t) = \min_{1 \leq k \leq M}\left(x(t,\lambda^k)\right); \quad U_{max,0}(t) = \max_{1 \leq k \leq M}\left(x(t,\lambda^k)\right) \tag{4.11}$$

It is noted that in order to be accurate, the number of samples *M* much be large enough to cover the multidimensional epistemic space $[-1,1]^N$ in fine grained manner. However, since for each sample point only a single analytic solution of (4.6) is required, choosing a pessimistic value of *M* does not incur significant computational costs. The dynamic bounds $[U_{min,j}(t), U_{max,j}(t)]$ are known as the unconditional response bounds. Next, the same exercise is repeated for the same set of MC samples with the additional condition that ${\lambda_1}^{(k)} = 0$. This will yield the unconditional response bounds $[U_{min,1}(t), U_{max,1}(t)]$ where

$$U_{min,1}(t) = \min_{1\le k\le M, \lambda_1^{(k)}=0} \left(x(t,\lambda^k)\right); \quad U_{max,1}(t) = \max_{1\le k\le M, \lambda_1^{(k)}=0} \left(x(t,\lambda^k)\right) \tag{4.12}$$

From the knowledge of the unconditional and conditional dynamic response bounds, the raw sensitivity of the response of the dimension λ_1 is denoted as the average difference between the maximum and minimum bounds using the unconditional and conditional results

$$S_i(t) = \frac{\left(U_{max}(t) - U_{max,1}(t)\right) + \left(U_{min,1}(t) - U_{min}(t)\right)}{2} \tag{4.13}$$

It is important to note that the above raw sensitivity index being time dependent, the true sensitivity index is the integral described as

$$s_i(t) = \frac{1}{2}\int \left(U_{max}(t) - U_{max,1}(t)\right) dt + \frac{1}{2}\int \left(U_{min,1}(t) - U_{min}(t)\right) dt \tag{4.14}$$

The integral of (4.14) is nothing other than the average area enclosed between the maximum bounds and minimum bounds of the unconditional and conditional results. This integral can always be evaluated using numerical integration rules such as the trapezoidal rule [80]. Once the true sensitivity indices for all N dimensions have been evaluated, the relative sensitivity indices can be measured as

$$r_i(t) = \frac{s_i}{\sum_{i=1}^{N} s_i} \tag{4.15}$$

The relative sensitivity indices serve as problem dependent impact factors used to rank the random dimensions in decreasing order of their effect on the network bounds. Those dimensions with the relative sensitivity index of (4.15) falling below a prescribed tolerance ε are considered to be non-impactful and will be pruned from the original N-dimensional epistemic space to

contract it into a n-dimensional subspace Ω_n where $n < N$. Thus, the original set of epistemic dimensions will be replaced by a new set of the most impactful n dimensions.

The above methodology can also be easily adapted to multiple, say K network responses. In such scenarios, the relative sensitivity indices of each dimension has to be measured with respect to dynamic bounds of each of the K network responses. Let the set of important epistemic dimensions for any i^{th} network response be give as λ_i. In that case, the overall set of important dimensions are those that occur in the union set $\lambda = \lambda_1 \cup \lambda_2 \cup ... \cup \lambda_N$. Finally, the construction of the n-dimensional reduced PC metamodel and recovery of the PC coefficients of that metamodel can proceed using the well-known non-intrusive linear regression approach of [7]-[22].

4.5 Numerical examples

In this section, three examples are presented to demonstrate the accuracy and efficiency of the proposed reduced dimensional PC approach. For comparison of the maximum and minimum values of the response at each α-cut, the network of the example presented has been simulated at each α-cut for a certain number of random MC samples. The reason for using MC technique is that the number of simulations required for each α-cut reduces as the value of α increases from zero to one. All relevant PC computations are performed using MATLAB 2013b while the deterministic transient simulations are performed using HSPICE. In particular, for example 1 and 2, the transmission line networks are modeled using W-element transmission line model provided by HSPICE [63] which can automatically consider frequency-dependent per-unit-length parameters. The above simulations are run on a workstation with 8 GB RAM, 500 GB memory and an Intel i5 processor with 3.4 GHz clock speed.

4.5.1 Example 1: Multi-conductor transmission line

In this example, the proposed approach and the conventional MC approach is compared for a large distributed network. For this purpose the 16 multi-conductor transmission line (MTL) network loaded with SPICE level-49 CMOS inverters, as shown in Fig 4.2 is considered. The layout of the strip-line network is also illustrated in Fig 4.2. The input to the network is a trapezoidal waveform with rise/fall time T_r= 0.1 ns, pulse width T_w = 0.5 ns and amplitude equal to 5V. The uncertainty in the network is introduced via N = 22 normal random variables whose characteristics are listed in Table 4.1.

In order to establish the accuracy of the proposed approach, the maximum and minimum of the steady-state response of the output node N_1 and N_2 of Fig. 4.2 is computed using the proposed reduced dimensional approach and the conventional brute force MC simulations. This is done for each α-cut. The membership function of the inputs is designed as follows. Random variables 1 through 8 have a negative 10% skew, meaning that for a random variable with a value x at $\alpha = 0$, will have a value $1.1x$ at $\alpha = 1$. The value of the corresponding random variable at all other α – cuts is computed by using a triangular function as shown in Fig. 4.1. Random variables 1 through 12 have a negative 5% skew, while random variables 13 through 22 have a positive 10% skew. As before, the SPLINER algorithm is used to identify the regression nodes for both expansions. Considering the voltage at node N_1 and node N_2 to be the two network responses of interest, the average of the sensitivity indices of the random dimensions with respect to this response is computed using the methodology of section 4.4.2 and listed in Table 4.2. Considering a tolerance of ε = [0.005] for the output of node N_1 and node N_2, the original $(N = 32)$ dimensional space is truncated to an $(n = 5)$ dimensional subspace where the dimensions retained are W, T, d, ε_r and L (RVs 1, 3, 4, 5 and 6). The maximum and minimum values of the

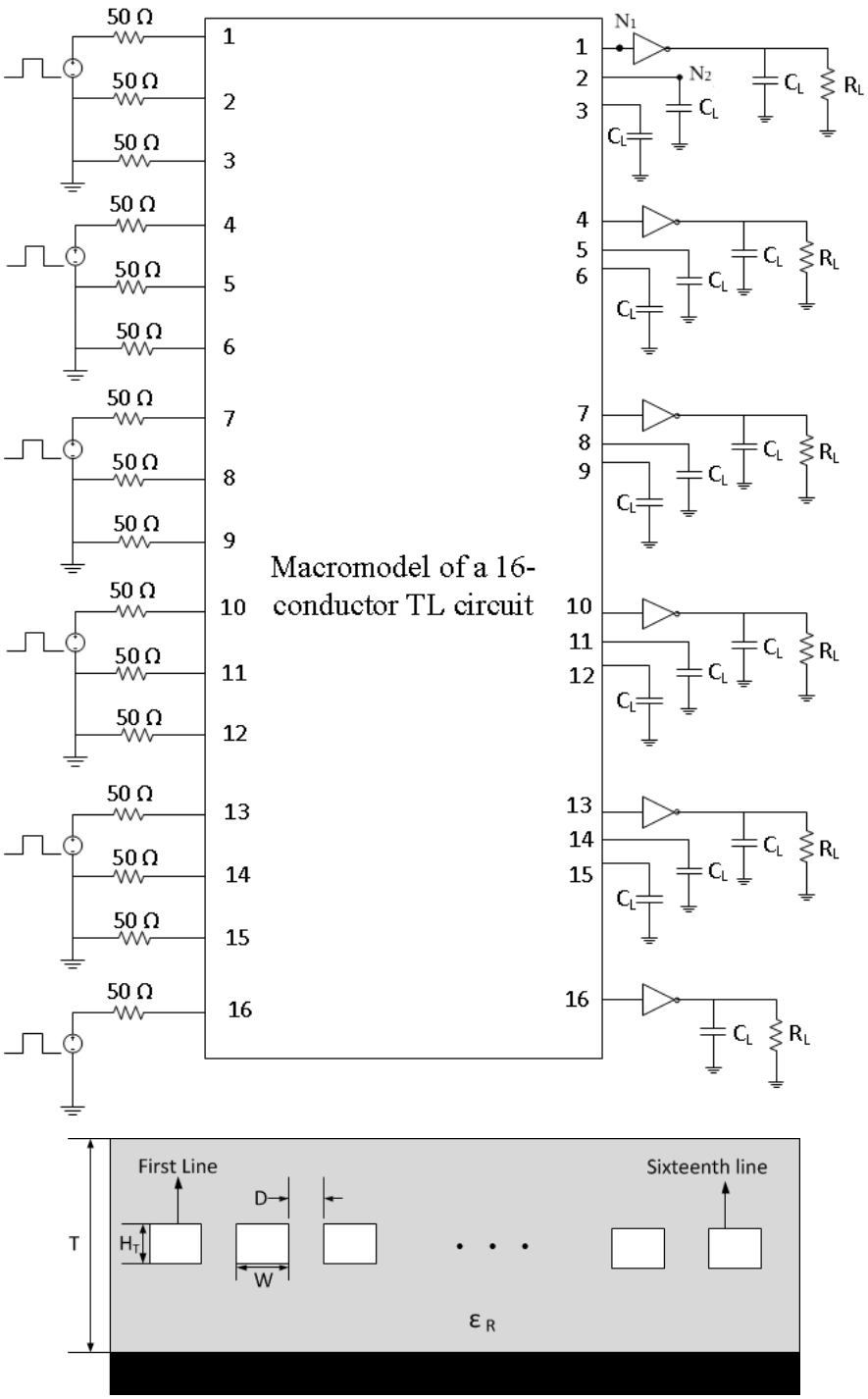


Fig 4.2: Transmission Line Network used in Example 1 and 2

response at node N_1 and node N_2 are then computed using the above two methods and the results for a few α-cuts are compared in Fig. 4.3. The accuracy of the proposed approach is demonstrated by generating the membership function of the response at node N_1 and node N_2

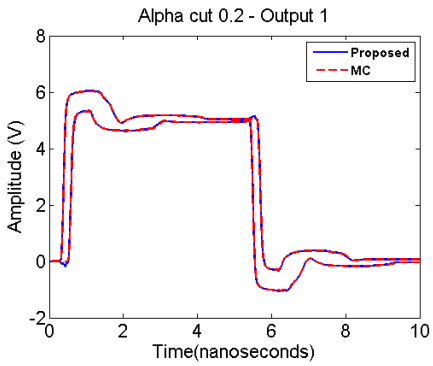
Table 4.1: Characteristics of random variables of transmission line (TL) network of Example 1 and 2.

No.	Random Variable	Mean	% Relative SD
1	W (TL width)	0.15 mm	20
2	ht (TL height)	0.03 mm	20
3	T (Dielectric height)	0.45 mm	20
4	d (TL separation)	0.15 mm	20
5	ε_r (Dielectric permittivity)	4.1	20
6	L (TL length)	6 cm	20
7	g (TL conductance)	5.8e+7 S/m	20
8	PL (Channel length PMOS)	0.1 μm	10
9	PW (Width PMOS)	10 μm	10
10	NL (Channel length NMOS)	0.1 μm	10
11	NW (Width NMOS)	10 μm	10
12-17	$R_{L1} – R_{L6}$ (Load Resistances)	1.5 kΩ	20
18-22	$C_{L1} – C_{L5}$ (Load Capacitances)	1pF	20

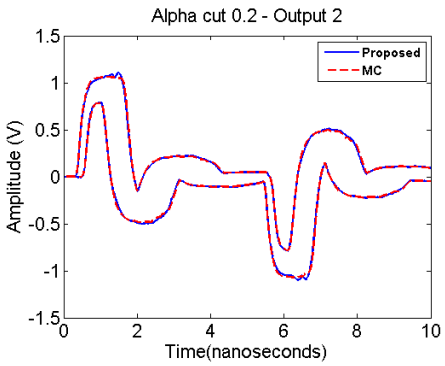
Table 4.2: Sensitivity indices computed as per (4.15) for Example 1 and 2.

Random Variable	Sensitivity index at node N_1	Sensitivity index at node N_2
1	0.0103	0.0107
2	8.11e-4	0.0015
3	0.0052	0.0038
4	0.0027	0.0353
5	0.2343	0.1942
6	0.7461	0.7942
7	3.3e-4	3.33e-4
8	2.6e-7	3.43e-7
9	2.13e-6	3.18e-6
10	2.53e-7	3.17e-7
11	3.32e-6	4.33e-6
12	1.12e-8	6.09e-7
13	6.7e-11	1.34e-9
14	3.6e-11	9.53e-10
15	3.2e-11	8.98e-10
16	3.2e-11	8.34e-10
17	3.3e-12	5.42e-11
18	8.3e-7	4.69e-7
19	2.1e-4	0.0040
20	5.4e-5	0.0010
21	5.4e-9	8.75e-8
22	5.5e-6	7.68e-5

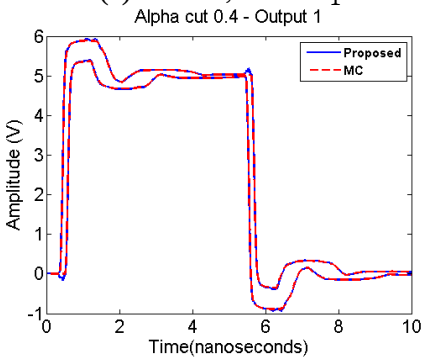
and this is demonstrated in Fig. 4.4. To generate the membership function of the output, the time point selected is the time point where the response of node N_1 and node N_2 has the maximum value at $\alpha = 1$ ($t =$ 1.125 ns for node N_1 and $t = 1.08$ ns for node N_2). The figures shown



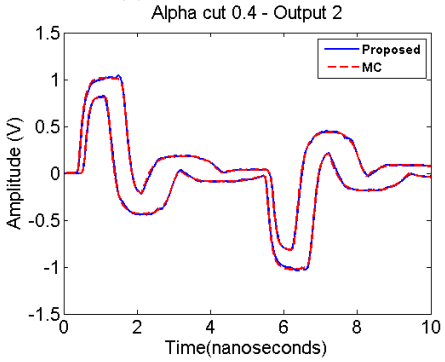
(a) α = 0.2, Node N_1



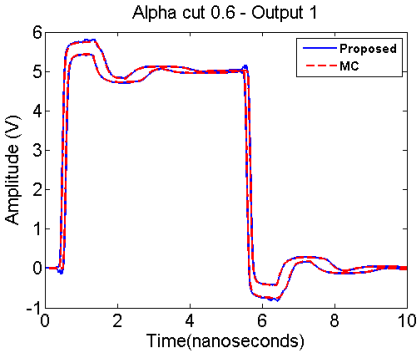
(a) α = 0.2, Node N_2



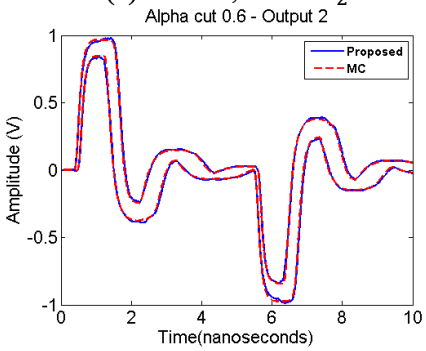
(a) α = 0.4, Node N_1



(a) α = 0.4, Node N_2



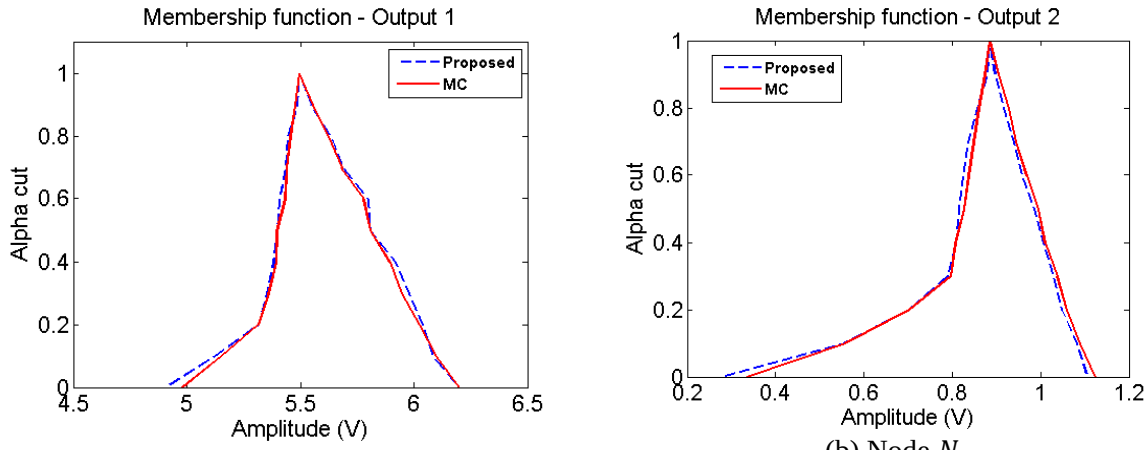
(a) α = 0.6, Node N_1



(a) α = 0.6, Node N_2

Fig 4.3: Maximum and Minimum values of the response at α-cuts 0.2, 0.4 and 0.6 for nodes N_1, N_2 for Example 1.

demonstrate that the reduced dimensional PC approach shows good agreement with the Monte Carlo approach. This not only demonstrates the accuracy of the reduced dimensional PC approach but also validates the use of only the first order terms of HDMR to identify the non-impactful dimensions in Table 4.2. For generating the PC coefficients, the SPLINER method is used and it takes 155 seconds to locate the regression nodes, which is a fraction of the time required for 615 SPICE simulations, which is 5096.2 seconds. The time needed for simulations



(a) Node N_1 (b) Node N_2

Fig 4.4: Membership function of the output at nodes N_1 and N_2.

includes the 111 simulations needed to perform dimension reduction. On the other hand, the number of MC simulations carried out at each α-cut are described in Table 4.3. In total, the time required to carry out all the MC simulations of all the α-cuts combined is 821,708.3 seconds. This translates to a speedup of about 161x. This is as expected from the discussion made in the previous section.

4.5.2 Example 2: Multi-conductor transmission line with irregular input membership functions

In the previous example, the membership functions considered for the inputs were triangular,

Table 4.3: Number of MC simulations carried out at each α-cut for example 1,2.

Alpha cut	Number of MC Simulations
0	20k
0.1	15k
0.2	12k
0.3	12k
0.4	10k
0.5	10k
0.6	8k
0.7	5k
0.8	4k
0.9	3k
1	1
Total	99,001

which has a certain regularity to it. To prove the efficacy of the proposed approach for any kind of α-cut boundaries, we perform the experiment by considering parabolic input membership functions. For the second example, we use a parabolic input membership function the shape of which is described below. The percentage skew for each of the random variables is kept the same as the first example. For random variables 1 through 8, the left tail and the right tail of the input membership function are defined by the equations

$$Left\ tail: \ [(Mean\ at\ \alpha_1) * (1.0909 - 0.04545 * \alpha_i - 0.04545 * {\alpha_i}^2)]$$
$$Right\ tail: \ [(Mean\ at\ \alpha_1) * (0.7273 + 0.2727 * {\alpha_i}^2)] \quad (4.16)$$

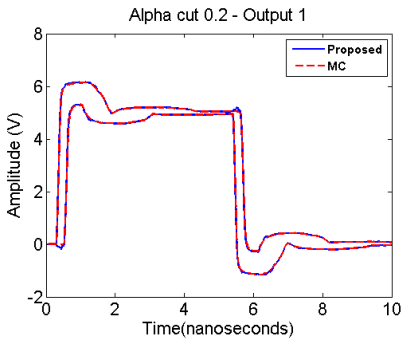
where α_i takes a value from zero to one and corresponds to the α-cut under inspection. Similarly, for random variables 9 through 12

$$Left\ tail: \ [(Mean\ at\ \alpha_1) * (1.0476 - 0.0238 * \alpha_i - 0.0238 * {\alpha_i}^2)]$$
$$Right\ tail: \ [(Mean\ at\ \alpha_1) * (0.8572 + 0.1428 * {\alpha_i}^2)] \quad (4.17)$$

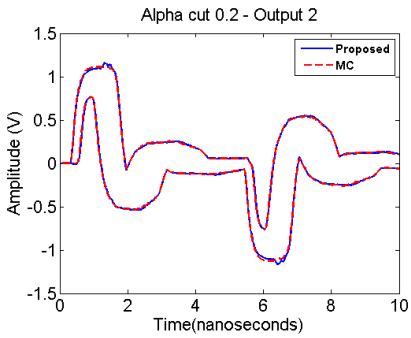
Lastly, for random variables 13 through 22

$$Left\ tail: \ [(Mean\ at\ \alpha_1) * (1.3333 - 0.6333 * \alpha_i + 0.3 * {\alpha_i}^2)]$$
$$Right\ tail: \ [(Mean\ at\ \alpha_1) * (0.8889 + 0.2211 * \alpha_i - 0.11 * {\alpha_i}^2)] \quad (4.18)$$

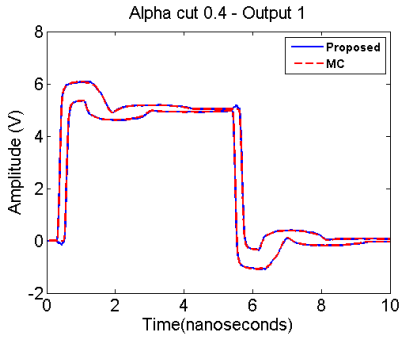
Since the example's α-cut zero values do not change from example 1, the same zeroth α-cut reduced dimensional PC expansion involving five random dimensions can be used for this example. The maximum and minimum values of the response at node N_1 and node N_2 are then computed using the above two methods and the results for a few α-cuts are compared in Fig. 4.5. The accuracy of the proposed approach is demonstrated by generating the membership function of the response at node N_1 and node N_2 and this is demonstrated in Fig. 4.6. The time point



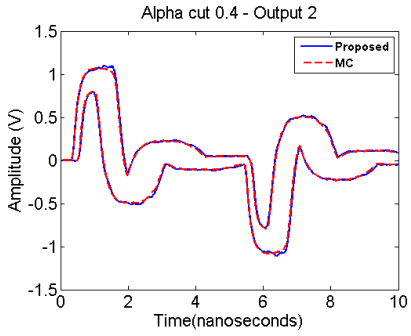
(a) α = 0.2, Node N_1



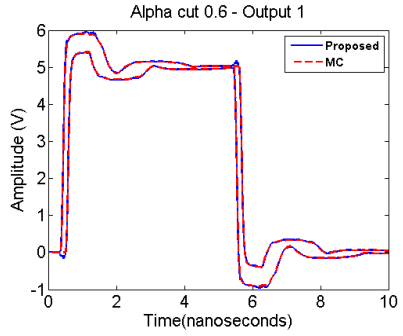
(a) α = 0.2, Node N_2



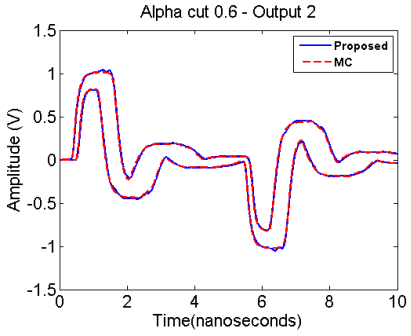
(a) α = 0.4, Node N_1



(a) α = 0.4, Node N_2



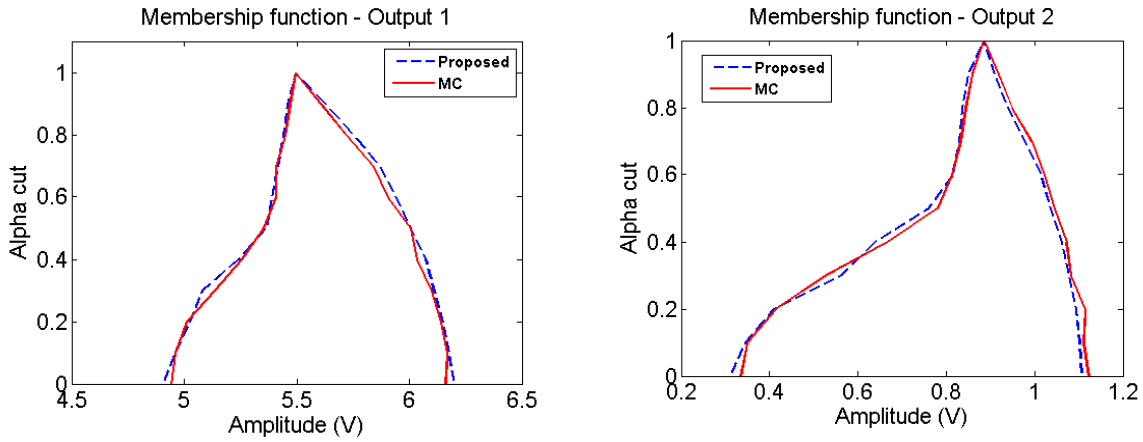
(a) α = 0.6, Node N_1



(a) α = 0.6, Node N_2

Fig 4.5: Maximum and Minimum values of the response at α-cuts 0.2, 0.4 and 0.6 for nodes N_1, N_2 for Example 2.

selected to generate the output membership function is the same as the previous example. The number of simulations required to generate the MC samples at each α-cut is the same as example 1 and is depicted in Table 4.3. And thus, the speedup mentioned is also the same as example 1. The objective of doing this particular example, is to show, that if a reduced dimensional PC expansion is generated at the zeroth α-cut, then the results for any other α-cut can be generated irrespective of what the confidence intervals are for that α-cut; provided the boundary of values



(a) Node N_1 (b) Node N_2

Fig 4.6: Membership function of the output at nodes N_1 and N_2 for Example 2.

for every random variable lies within the boundary of values of the respective random variables at the zeroth α-cut.

4.5.3 Single conductor transmission line network

So far the examples we have accounted for, have all considered random variables whose α = 1 values have been a single value rather than being a range of values. In this example, we consider a set of random variables, to have a range of values, even at α = 1. The example considers the transmission line network of Fig. 4.7. The characteristics of the random variables are listed in Table 4.4. The voltage source is a trapezoidal pulse with an amplitude of 5V, rise/fall times of 200 ps and a width of 2.6 ns. The models used for the inverters in the network are SPICE level-49 CMOS inverters.

There are three kinds of membership functions of the inputs. Random variables 1 through 7 and random variables 15 through 21 have a negative 10% skew, while all the other random variables have a negative 5% skew. Variables 1 through 14 have a triangular distribution for their membership functions, much like example 1. Variables 15 through 21 have a parabolic distribution the shape of which is guided by equations of (4.16). Variables 22 through 28 also

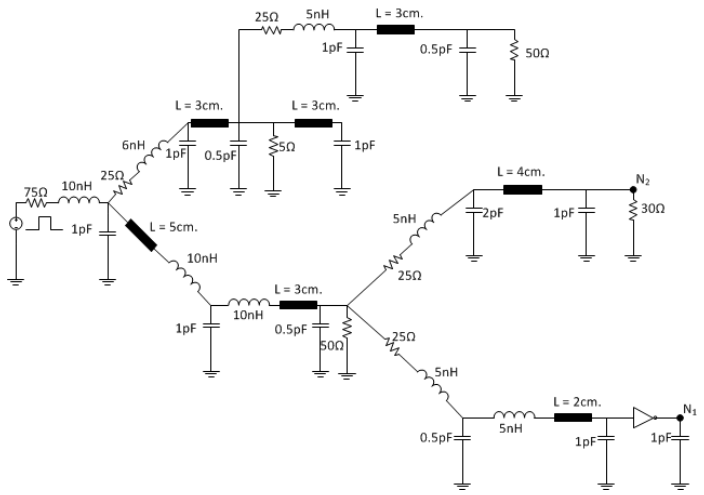


Fig 4.7: Schematic of the network used for Example 3

Table 4.4: Characteristics of random variables of Example 3 network

No.	Random Variable	Mean	% Relative SD
1-7	$\varepsilon_{r1} - \varepsilon_{r7}$ (Dielectric relative permittivity)	4.1	20
8-14	$w_1 - w_7$ (TL width)	150µ	10
15-21	$ht_1 - ht_7$ (TL height)	20µ	20
22-28	$H_1 - H_7$(Dielectric height)	100µ	10
29	LP (Channel length PMOS)	1µ	20
30	LN (Channel length NMOS)	2.5µ	20
31	WP (Width PMOS)	20µ	20
32	WN (Width NMOS)	20µ	20

have a parabolic distribution, but guided by the equations (4.17). Variables 29 through 32 have a trapezoidal distribution, which means that unlike the other variables discussed so far, these set of variables has a range of values at α = 1. At α = 1, these variables hold values that would have been the values of these variables if these variables had a triangular distribution and α was equal to 0.9. The values of the sensitivity indices for N_1 and N_2 is presented in Table 4.5. Considering a tolerance of ε = [0.015] for the output of node N_1 and node N_2, the original $(N = 32)$ dimensional space is truncated to an $(n = 12)$ dimensional subspace where the dimensions retained are $\varepsilon_{r1} - \varepsilon_{r5}$, w_1, H_1, H_3, LP, LN, WP and WN. For generating the PC coefficients, the SPLINER method is used and it takes 114,108 seconds to locate the regression nodes, and it takes 8800.2 seconds for 4889 SPICE simulations. The time needed for simulations includes the

129 simulations needed to perform dimension reduction. The comparison of the maximum and minimum results of the output response at N_1 is presented in Fig. 4.8. The number of MC simulations performed for example 3 is listed in Table 4.6. To prove the accuracy of the proposed approach against MC simulations, the membership function of the response at N_1 and N_2 is also presented in Fig. 4.9. The time point selected to generate the output membership function is the time point where the α = 1 results have maxima. In total, the time required to carry out all the MC simulations of all the α-cuts combined is 180,000 seconds. The speedup achieved for this problem is about 20.5x.

4.6 Advantages and disadvantages of the hybrid approach

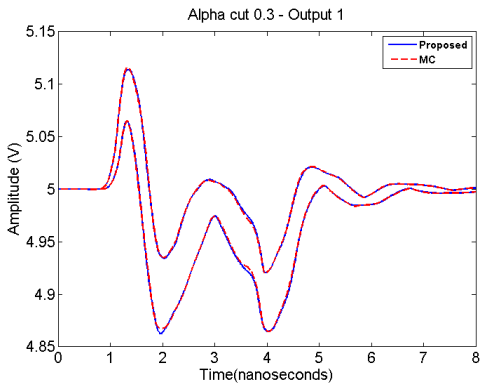
With the use of dimension reduction techniques using HDMR principles, the size of the PC expansion can be greatly diminished than the actual full-blown PC expansion with very little loss in accuracy as is evident from the examples. The strength of the proposed method also lies in the fact that no matter what confidence interval gets chosen as more and more information about the random variables becomes available to the user, the user can simply probe the new values into the PC expansion and go about creating the PDF of the epistemic variable, so long as the new confidence interval lies within the interval of α = 0. This saves a huge amount of computation and simulation cost for the user.

On the other hand, this method fails to generate any kinds of savings if the new information that the user receives about the random variables happens to be outside the confidence interval of α = 0. In that case, the method actually fails to deliver as much benefit as expected. Since a full PC using reduced number of dimensions is created, the number of terms in the PC expansion can go on increasing significantly if more and more random variables contribute significantly. Although rarely, but it is possible that a higher number of simulations, can be needed to generate a reduced

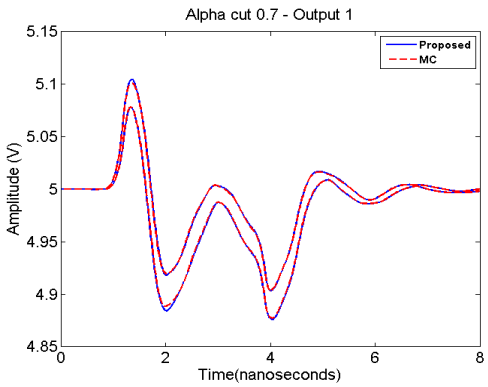
Table 4.5: Sensitivity indices computed as per (4.15) for Example 3

Random Variable	Sensitivity index at node N_1	Sensitivity index at node N_2
1	0.3697	0.5116
2	0.0376	0.1449
3	0.1110	0.0426
4	0.0033	0.1246
5	0.0052	0.0439
6	3.56e-5	1.82e-4
7	9.52e-7	2.35e-6
8	0.0264	0.0388
9	0.0106	0.0058
10	0.0035	0.0026
11	0.0024	0.0036
12	0.0055	0.0087
13	4.58e-6	1.67e-5
14	1.57e-6	1.23e-5
15	5.17e-4	0.0020
16	3.42e-4	7.04e-4
17	1.53e-4	3.69e-5
18	1.15e-4	0.0012
19	6.04e-4	0.0010
20	3.08e-6	8.11e-6
21	3.90e-7	1.88e-6
22	0.0320	0.0444
23	0.0037	0.0071
24	0.0150	0.0039
25	0.0029	0.0026
26	0.0059	0.0087
27	3.36e-6	2.70e-5
28	1.29e-5	3.14e-6
29	0.1465	1.07e-4
30	0.0983	2.13e-4
31	0.0455	1.51e-4
32	0.0733	2.90e-4

order PC expansion, than simply running MC simulations at each alpha cut. The other problem with this method is that, the choice of the threshold for sensitivity index becomes subjective in some cases and an intelligent judgement has to be made.



(a) α = 0.3, Node N_1

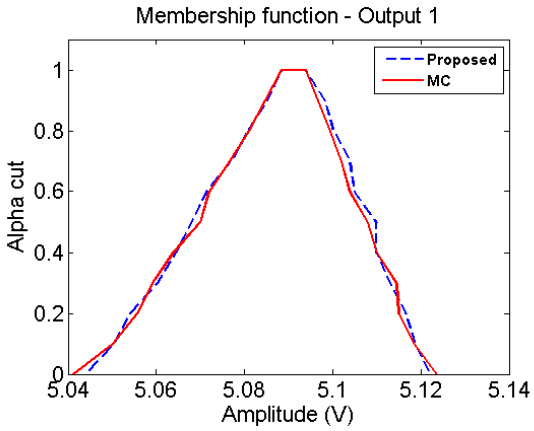


(b) α = 0.7, Node N_1

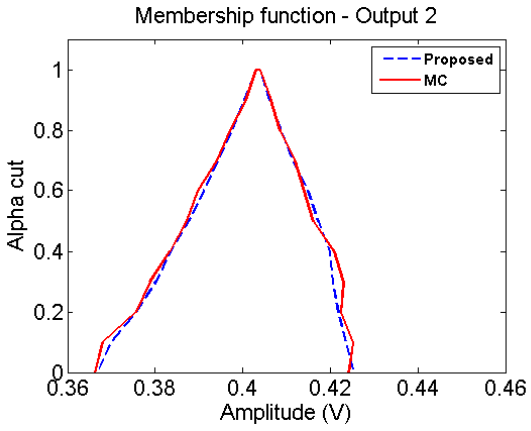
Fig 4.8: Maximum and Minimum values of the response at α-cuts 0.3 and 0.7 for node N_1 for Example 3.

Table 4.6: Number of MC simulations carried out at each α-cut for example 3

Alpha cut	Number of MC Simulations
0	30k
0.1	30k
0.2	25k
0.3	20k
0.4	20k
0.5	15k
0.6	12k
0.7	10k
0.8	8k
0.9	5k
1	1k
Total	176k



(a) Node N_1



(b) Node N_2

Fig 4.8: Membership function of the output at nodes N_1 and N_2 for Example 3.

CONCLUSION

In this thesis, the need to introduce novel methods to quantify aleatory and epistemic parametric uncertainty was discussed. The work presented a brief overview of the traditional and contemporary methods to analyze parametric uncertainty. It presented the pros and cons of using each method and stated why the linear regression approach was chosen to solve for the gPC coefficients. The work discussed some of the shortcomings of the traditional linear regression approach. Once some of the sources for the shortcomings were identified, new methods were developed to remove the redundancies in the PC expansion. By removing these redundancies, the number of the terms in the expansion were reduced, thereby reducing the time and computation cost required for both, generating the linear regression nodes and computing the coefficients by running MC simulations. It was ensured while removing the redundancies that the accuracy of the system was not compromised beyond a certain threshold.

For aleatory variables, an anisotropic scheme of PC expansion was developed, which relied on the HDMR formulation to accurately identify those random dimensions which did not provide as much statistical information to the total response of the system as some of the other random dimensions. The contribution of each random dimension could be identified by a generated scalar value, known as the sensitivity index. By setting a different degree of expansion for each random dimension based on the corresponding sensitivity index, the higher degree multi-dimensional terms, for those random dimensions which provided insignificant statistical information, could be eliminated, with minimal loss in accuracy.

Unlike aleatory random variables, there is insufficient information about the probability distribution of epistemic random variables. While there are many approaches to solve epistemic

uncertainty, this work is solely focused on the use of fuzzy logic to quantify epistemic uncertainty. In the fuzzy logic approach, rather than discussing the probability distribution of the random variables, these epistemic random variables were modeled using possibility windows, known as confidence intervals. These confidence intervals can be updated as new information about the system is obtained. The novel approach of this method is to generate a global PC expansion at the largest confidence interval and then to probe this global PC expansion to compute the statistical moments at all the other confidence intervals, given the condition that the values the random variables take at all the other confidence intervals is a subset of the values they hold at the confidence interval where the global PC was generated. Furthermore, since these intervals were guided by the maximum and minimum values that a particular random dimension holds at that interval, a good way to quantify the impact of these variables was to the study the maximum and minimum values of the response at the corresponding α – cut. Since complete information about the PDF of the random dimensions was not available, a new HDMR based formulation scheme was developed to quantify the impact of each epistemic variable on the response of the network. Further, a novel way to generate sensitivity indices for this approach was also presented. Unlike the previous method which set a lower degree of expansion for statistically insignificant terms, this method focuses on eliminating those random dimensions altogether. The result was the creation of a reduced dimensional PC expansion that only required a fraction of the time and memory cost to compute.

The accuracy of all the methods presented above was verified using multiple numerical examples, by comparing them against conventional quantification methods. As expected, the proposed methods displayed huge savings in computation time for minimal loss in accuracy.

REFERENCES

[1] G.S. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*. New York, NY: Springer-Verlag, 1996

[2] Q. Zhang, J. J. Liou, J. McMacken, J. Thomson and P. Layman, "Development of robust interconnect model based on design of experiments and mulitobjective optimization," *IEEE Trans. Electron Devices*, vol. 48, no. 9, pp. 1885–1981, Sept. 2001

[3] Y. Massoud and A. Nieuwoudt, "Modeling and design challenges and solutions for carbon-nanotube based interconnect in future high performance integrated circuits," *ACM J. Emerg. Technol. Comput. Syst.*, vol. 2, no. 3, pp. 155–196, July 2006

[4] A. Nieuwoudt and Y. Massoud, "On the impact of process variations for carbon nanotube bundles for VLSI interconnect," *IEEE Trans. Electron. Devices*, vol. 54, no. 3, pp. 446–455, March 2007

[5] A. Nieuwoudt and Y. Massoud, "On the optimal design, performance, and reliability of future carbon nanotube-based interconnect solutions," *IEEE Trans. Electron. Devices*, vol. 55, no. 8, pp. 2097–2110, Aug. 2008

[6] J. F. Swidzinski and K. Chang, "Nonlinear statistical modeling and yield estimation technique for use in Monte Carlo simulations [microwave devices and ICs]," *IEEE Trans. Microw. Theory Tech.*, vol. 48, no. 12, pp. 2316-2324, Dec. 2000

[7] P. Manfedi, "High-speed interconnect models with stochastic parameter variability," Ph.D. dissertation, Informat. Comm. Tech., Politecnico di Torino, Turin, Italy, 2013

[8] S. Vrudhula, J. M. Wang and P. Ghanta, “Hermite polynomial based interconnect analysis in the presence of process variations,” *IEEE Trans. Computer Aided Design*, vol. 25, no. 10, pp. 2001–2011, Oct. 2006

[9] I. S. Stievano, P. Manfredi and F. G. Canavero, “Stochastic analysis of multiconductor cables and interconnects,” *IEEE Trans. Electromagn. Compatibility*, vol. 53, no. 2, pp. 501–507, May 2011

[10] I. S. Stievano, P. Manfredi and F. G. Canavero, “Parameters variability effects on multiconductor interconnects via Hermite polynomial chaos,” *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 1, no. 8, pp. 1234–1239, Aug. 2011

[11] D. Vande Ginste, D. De Zutter, D. Deschrijver, T. Dhaene, P. Manfredi and F. Canavero, “Stochastic modeling-based variability analysis of on-chip interconnects,” *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 3, no. 7, pp. 1182–1192, Jul. 2012

[12] I. S. Stievano, P. Manfredi and F. G. Canavero, “Carbon nanotube interconnects: process variation via polynomial chaos,” *IEEE Trans. Electromagn. Compatibility*, vol. 54, no. 1, pp. 140–148, Feb. 2012

[13] A. Biondi, D. Vande Ginste, D. De Zutter, P. Manfredi and F. Canavero, “Variability analysis of interconnects terminated by general nonlinear loads,” *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 2, no. 7, pp. 1244–1251, Jul. 2013

[14] P. Manfredi, D. Vande Ginste, D. De Zutter and F. Canavero, “Uncertainty assessment in lossy and dispersive lines in SPICE-type environments,” *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 2, no. 7, pp. 1252–1258, Jul. 2013

[15] P. Manfredi, D. Vande Ginste, D. De Zutter and F. Canavero, "Stochastic modeling of nonlinear circuits via SPICE-compatible spectral equivalents," *IEEE Trans. Circuits Syst.*, vol. 61, no. 7, pp. 2057–2065, Jul. 2014

[16] M. R. Rufuie, E. Gad, M. Nakhla R. Achar, "Generalized Hermite polynomial chaos for variability analysis of macromodels embedded in nonlinear circuits," *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 4, no. 4, pp. 673–684, April 2014

[17] T.-A. Pham, E. Gad, M. S. Nakhla and R. Achar, "Decoupled polynomial chaos and its applications to statistical analysis of high-speed interconnects," *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 4, no. 10, pp. 1634–1647, Oct. 2014

[18] D. Spina, F. Ferranti, G. Antonini, T. Dhaene and L. Knockaert, "Efficient variability analysis of electromagentic systems via polynomial-chaos and model order reduction," *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 4, no. 6, pp. 1038–1051, June 2014

[19] D. Spina, F. Ferranti, T. Dhaene, L. Knockaert, G. Antonini, and D. Vande Ginste, "Variability analysis of multiport systems via polynomial-chaos expansion," *IEEE Trans. Microwave Theory Tech.*, vol. 60, no. 8, pp. 2329–2338, Aug. 2012

[20] A. Rong and A. C. Cangellaris, "Transient analysis of distributed electromagnetic systems exhibiting stochastic variability in material parameters," in *Proc. XXXth General Assembly and Scientific Symp.*, Aug. 2011, pp. 1-4

[21] Md. A. H. Talukder, M. Kabir, S. Roy and R. Khazaka, "Efficient generation of macromodels via the Loewner matrix approach for the stochastic analysis of high-speed passive distributed networks," in *Proc. IEEE 18th Workshop on Signal and Power Integrity*, May 2014, pp. 1-4

[22] Md. A. H. Talukder, M. Kabir, S. Roy and R. Khazaka, "Efficient stochastic transient analysis of high-speed passive distributed networks using Loewner matrix macromodels," in *Proc. IEEE Intl. Conf. Signal and Power Integrity*, August 2014, pp. 1-4

[23] A. K. Prasad, and S.Roy, "Global sensitivity based dimension reduction for fast variability analysis of nonlinear circuits," in *Proc, 24thIEEE Conf. Electric. Perform. Electron. Packag.*, October 2015, pp. 97-100.

[24] M. Ahadi, M. Kabir, E. Chobanyan, A. Smull, Md. A. H. Talukder, S. Roy, R. Khazaka and B. M. Notaros, "Non-intrusive pseudo spectral approach for stochastic macromodeling of EM systems using deterministic full-wave solvers," in *Proc. 23rd IEEE Conf. Electric. Perform. Electron. Packag.*, October 2014, pp. 235-238

[25] A. K. Prasad, and S. Roy, "Multidimensional variability analysis of complex power distribution networks via scalable stochastic collocation approach," *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 5, no. 11, pp. 1656-1668, Nov. 2015

[26] A. K. Prasad, M. Ahadi, and S. Roy, "Multidimensional uncertainty quantification of microwave/RF networks using linear regression and optimal design of experiments," *IEEE Trans. Microwave Theory Tech*, July 2016

[27] M. Ahadi and S. Roy, "Sparse linear regression (SPLINER) approach for efficient multidimensional uncertainty quantification of high-speed circuits," *IEEE Trans. Computer Aided Design*, Jan. 2015

[28] M. Ahadi, A. K. Prasad, and S. Roy, "Hyperbolic polynomial chaos expansions (HPCE) and its application to statistical analysis for nonlinear circuits," in *Proc. IEEE 20th Workshop on Signal and Power Integrity*, May 2016, pp. 1-4

[29] I. Kapse, A. K. Prasad, and S. Roy, "Generalized anisotropic polynomial chaos approach for expedited statistical analysis of nonlinear radio-frequency (RF) circuits," accepted for oral presentation in *IEEE 20th Workshop on Signal and Power Integrity,* May 2016, pp. 1-4

[30] M. Ahadi, M. Kabir, S. Roy and R. Khazaka, "Fast multidimensional statistical analysis of microwave networks via Stroud cubature approach," *Proc. IEEE International Conf. on Numerical Electromagn., Multiphysics Modeling and Optimiz.,* August 2015

[31] A. K. Prasad, M. Ahadi, B. S. Thakur and S. Roy, "Accurate polynomial chaos expansion for variability analysis using optimal design of experiments," *Proc. IEEE International Conf. on Numerical Electromagn., Multiphysics Modeling and Optimiz.,* August 2015

[32] D. Spina, F. Ferranti, G. Antonini, T. Dhaene and L. Knockaert, "Non-intrusive polynomial chaos-based stochastic macromodeling of multiport systems," in *Proc. IEEE 18th Workshop on Signal and Power Integrity*, May 2014, pp. 1-4

[33] D. Spina, D. De Jonghe, D. Deschrijver, G. Gielen, L. Knockaert, and T. Dhaene, "Stochastic macromodeling of nonlinear systems via polynomial-chaos expansion and transfer function trajectories," *IEEE Trans. Microwave Theory Tech*, vol. 62, no. 7, pp. 1454–1460, July 2014

[34] A. Rong and A. C. Cangellaris, "Interconnect transient simulation in the presence of layout and routing uncertainty," in *Proc. 20th IEEE Conf. Electric. Perform. Electron. Packag.*, October 2011, pp. 157-160

[35] J. S. Ochoa and A. C. Cangellaris, "Analysis of the impact of interconnect routing variability on signal degradation," in *Proc. 21st IEEE Conf. Electric. Perform. Electron. Packag.*, October 2012, pp. 315-318

[36] P. Sumant, H. Wu, A. Cangellaris, and N. Aluru, "Reduced-order models of finite element approximations of electromagnetic devices exhibiting statistical variability," *IEEE Trans. Antennas Propag.*, vol. 60, no. 1, pp. 301–309, Jan. 2012

[37] X. Chen, J. S. Ochoa, J. E. Schutt-Aine and A. C. Cangellaris, "Optimal relaxation of I/O electrical requirements under packaging uncertainty by stochastic methods," in *Proc. 64th Electronic Comp. Tech. Conf.*, May 2014, pp. 717–722

[38] A. C. Yucel, H. Bagci, J. S. Hesthaven and E. Michielssen, "A fast Stroud-based collocation for statistically characterizing EMI/EMC phenomena on complex platforms", *IEEE Trans. Electromagn. Compatibility*, vol. 51, no. 2, pp. 301-311, May 2009

[39] A. C. Yucel, H. Bagci and E. Michielssen, "An adaptive multi-element probabilistic collocation method for statistical EMC/EMI characterization," *IEEE Trans. Electromagn. Compatibility*, vol. 55, no. 6, pp. 1154–1168, Dec. 2013

[40] M. Ahadi, M. Vempa, and S. Roy, "Efficient multidimensional statistical modeling of high speed interconnects in SPICE via stochastic collocation using Stroud cubature," in *Proc. IEEE Symp. Electromagn. Compatibility and Signal Integrity,* March 2015, pp. 300-305

[41] P. Manfredi, D. Vande Ginste, D. De Zutter and F. Canavero, "Generalized decoupled polynomial chaos for nonlinear circuits with many random parameters," *IEEE Microwave and Wireless Components Letters*, vol. 25, no. 8, pp. 505–507, Aug. 2015

[42] I. Kapse, S. Roy "Anisotropic formulation of hyperbolic polynomial chaos expansion for high-dimensional variability analysis of nonlinear circuits" *IEEE 25th Conference on Electrical Performance of Electronic Packaging and Systems,* October 2016, pp. 123 – 126

[43] Z. Zhang, T. A. El-Moselhy, I. M. Elfadel and L. Daniel, "Stochastic testing method of transistor level uncertainty quantification based on generalized polynomial chaos," *IEEE Trans. Computer Aided Design*, vol. 32, no. 10, pp. 1533-1545, Oct. 2013

[44] P. Manfredi and F. Canavero, "Efficient statistical simulation of microwave devices via stochastic testing-based circuit equivalents of nonlinear components," *IEEE Trans. Microw. Theory Tech.*, vol. 63, no. 5, pp. 1502-1511, May 2015

[45] Z. Zhang, T. A. El-Moselhy, I. M. Elfadel and L. Daniel, "Calculation of generalized polynomial-chaos basis functions and Gauss quadrature rules in hierarchical uncertainty quantification," *IEEE Trans. Computer Aided Design*, vol. 33, no. 5, pp. 728-740, May 2014

[46] Z. Zhang, X. Yang, I. V. Oseledets, G. E. Karniadakis, and L. Daniel, "Enabling high-dimensional hierarchical uncertainty quantification by ANOVA and tensor-train decomposition," *IEEE Trans. Computer Aided Design*, vol. 34, no. 1, pp. 63-76, Jan. 2015

[47] A. C. Yucel, H. Bagci, and E. Michielssen, "An ME-PC enhanced HDMR method for efficient statistical analysis of multiconductor transmission line networks," *IEEE Trans. Comp., Packag. and Manuf. Technol.*, vol. 5, no. 5, pp. 685-696, May 2015

[48] J. S. Ochoa and A. C. Cangellaris, "Random-space dimensionality reduction for expedient yield estimation of passive microwave structures," *IEEE Trans. Microwave Theory Tech*, vol. 61, no. 12, pp. 4313–4321, Dec. 2013

[49] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*. New Jersey: Princeton University Press, 2010

[50] M.S. Eldred, "Recent advances in non-intrusive polynomial-chaos and stochastic collocation methods for uncertainty analysis and design", in *Proc. 50th*

AIAA/ASME/AHS/ASC Structures, Structural Dynamics, and Materials Conference, May 2010, pp. 1-37

[51] N.-K. Nguyen and A. J. Miller, "A review of some exchange algorithms for constructing discrete D-optimal designs," *Comput. Statistics and Data Analysis*, vol. 14, pp. 489-498, 1992

[52] A. J. Miller and N.-K. Nguyen, "A Fedorov exchange algorithm for D-optimal design," J. Royal Stat. Society, vol. 43, no. 4, pp. 669-677, 1994

[53] H. Rabitz and O. F. Alis, "General formulation of high-dimensional model representations" *Journal of Math. Chem.*, vol. 50, no. 2-3, pp. 197-233, 1999

[54] D. Xiu and G. E. Karniadakis, "The Weiner-Askey Polynomial Chaos for stochastic differential equations" *SIAM Journal Sci. Comput.*, vol. 24, no. 2, pp. 619-644, July 2002

[55] N. Wiener, "The homogeneous chaos", *Amer. J. Math.*, vol. 60, pp. 897–936, 1938.

[56] N. Wiener, "Non-linear problems in random theory", *MIT Press*, Cambridge, 1958.

[57] P. M. Congedo, G. Gianluca, I. Gianluca, "On the use of high-order statistics in robust design optimization", *6th European Conference on Computational Fluid Dynamics (ECFD VI)*, 2014.

[58] Skewness [https://en.wikipedia.org/wiki/Skewness]

[59] Kurtosis [https://en.wikipedia.org/wiki/Kurtosis]

[60] G. H. Golub and J. H. Welsch, "Calculation of Gauss Quadrature Rules", *Math. Comp.*, vol. 23, pp. 221–230, 1969.

[61] V. V. Fedorov, "Theory of Optimal Experiments", *NY: Academic Press*, New York, 1972.

[62] R. D. Cook and C. J. Nachtsheim, “A comparison of algorithms for constructing exact D-optimal designs”, *Amer. J. Math.*, vol. 60, pp. 897–936, 1938.

[63] *HSPICE Reference Manual Version E 2010.12*, Synopsys Inc., Santa Clara, CA, 2010.

[64] M. Ahadi, “Efficient Multi-dimensional uncertainty quantification of high speed circuits using advanced polynomial chaos approaches”, Master’s Thesis, Department of Electrical and Computer engineering, Colorado State University, July 2016.

[65] Web link [http://www.ce.memphis.edu/7137/PDFs/Abrahamson/C05.pdf]

[66] Armen Der Kiureghian, “Aleatory or Epistemic? Does it matter?”, *Special Workshop on Risk Acceptance and Risk Communication*, Stanford University, March 2017

[67] Tony O’ Hagan, “Dicing with the unknown”, *Department of Statistics*, Columbia University

[68] Rasmussen, C.E. and C.K.I. Williams. Gaussian Processes for Machine Learning. MIT Press, 2006

[69] G. J. A. Loeven, J. A. S. Witteveen, and H. Bijil, “A Probabilistic Radial Basis Function Approach for Uncertainty Quantification”, Dept. of Aerospace engineering, Delft University of Technology, 2007

[70] X. Chen, E. J. Park and D. Xiu, “A flexible numerical approach for quantification of epistemic uncertainty”, *J. Comput. Physics*, vol. 240, no. 12, pp. 211-224, May 2013

[71] A. Monti, F. Ponci and F. Valtorta, “Extending Polynomial Chaos to include interval analysis”, *IEEE Trans. Instrumentation and Measurement*, vol. 59, no. 1, pp. 48-55, Jan 2010

[72] T. Ding, R. Trinchero, P. Manfredi, I. S. Stievano, F. G. Canavero, "How affine arithmetic helps beat uncertainties in electrical systems", *IEEE Circuits Syst. Mag.*, vol. 15, no. 4, pp. 70-79, Nov. 2015

[73] J. D. Ma, R. A. Rutenbar, "Fast interval-valued statistical modeling of interconnect and effective capacitance", *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 25, no. 4, pp. 710-724, Apr. 2006

[74] P. Manfredi, R. Trinchero, F. G. Canavero, I. S. Stievano, "Application of Taylor models to the worst-case analysis of stripline interconnects", *IEEE 20th Workshop on Signal and Power Integrity*, vol. 240, no. 12, pp. 1-4, June 2016

[75] K. Makino, M. Berz, "Taylor models and other validated functional inclusion methods", *Int. J. of Pure and Appl. Math.*, vol. 4, no. 4, pp. 379-456, 2003.

[76] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338-353, 1965

[77] X. Chen, Y. He, and D. Xiu, "Efficient method for uncertainty propagation using fuzzy sets," *SIAM J. Scientific Comput.*, vol. 37, no. 6, A2488-A2507, Nov. 2015

[78] I. Kapse, A. K. Prasad and S. Roy, "Analyzing Impact of Epistemic Uncertainty in High-Speed Circuit Simulation Using Fuzzy Variables and Global Polynomial Chaos Surrogates", *IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modelling and Optimization for RF, Microwave, and Terahertz Applications*, pp. 320 – 322, July 2017

[79] S. O. Rice, "Efficient evaluation of integrals of analytic functions by the trapezoidal rule" *The Bell System Technical Journal*, vol. 52, no. 5, pp. 707-722, 1973