

SMART SIMULATION TECHNIQUES FOR THE EVALUATION OF PARAMETRIC  
UNCERTAINTIES IN BLACK BOX SYSTEMS

By

DINESH RAMAMURTHY

A thesis submitted in partial fulfillment of  
the requirements for the degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

**WASHINGTON STATE UNIVERSITY**  
School of Electrical Engineering and Computer Science

MAY 2005

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation/thesis of DINESH RAMAMURTHY find it satisfactory and recommend that it be accepted.

---

Chair

---

---

## ACKNOWLEDGEMENT

First and foremost, I would like to express my most sincere gratitude to Sandip Roy for being my advisor here at Washington State University and for guiding me through this thesis. As an advisor, Sandip epitomizes the spirit of the phrase “Friend, philosopher and guide”. Working under him has been a succulent experience that helped me hone my technical writing, mathematical and other research oriented skills. He is a great teacher and his classes are fun (except for his tests, he is a tough tester). I would like to thank him for giving me the opportunity to work on this project. I have learned a lot and enjoyed every moment of it. Great captains have the uncanny ability to bring out the best in their teams, in the same vein, Sandip managed to bring out the researcher in me. Sandip thanks a lot for being a great advisor and an even better role model.

A special thanks to Dr. Bernie Lesieutre for the funding and also for the various illuminating discussions. His insights, especially regarding electric power systems, have helped a lot. This work is supported by the Power Systems Engineering Research Center, through the project "Techniques for Evaluation of Parametric Variation in Time-Step Simulations". I am extremely grateful to them for their support.

I would like to thank Dr. Murali Medidi and Dr. Ben Belzer for serving on my committee. In particular, I would like to thank Dr. Medidi for his valuable suggestions regarding the computer science aspect of my thesis and Dr. Belzer for his interesting thoughts regarding the Information theory section. Their suggestions helped me improve this thesis.

I can't do justice to my family in an acknowledgement section. I am forever indebted to them for their love and sacrifice, hope I get the opportunity to repay them someday.

There is no greater wealth than good health and great friends. I am lucky to be blessed with a lovely set of friends, back in India, who have stuck with me through thick and thin and continue to do so. We call ourselves "The Red Army" (A self bestowed appellation, no relation to Moscow, Russia). I would like to thank them for their support, for their friendship and most of all for bearing with me all these years. I would also like to thank my friends in Pullman for their help and kindness.

Last but not the least I would like to thank my Alma Mater Washington State University, Pullman. Though I have always cribbed about Pullman, I think I am going to miss the place. Thanks for hosting me. Go Cougs!!

SMART SIMULATION TECHNIQUES FOR THE EVALUATION OF PARAMETRIC  
UNCERTAINTIES IN BLACK BOX SYSTEMS

Abstract

by Dinesh Ramamurthy, M.S.  
Washington State University  
May 2005

Chair: Sandip Roy

When parameters of complex processes are uncertain, it is often necessary to perform exhaustive simulations to characterize the outputs of these processes. If simulations are computationally intensive, characterization of outputs through exhaustive simulations may be infeasible. In such cases, intelligent approaches for choosing simulations based on probabilistic descriptions of uncertainties may be valuable. The Probabilistic Collocation Method is a probabilistic technique that can model the deterministic relationship between the uncertain parameters and an output of interest with a small set of simulations.

In this thesis, we review PCM, provide a new generalization of PCM for systems with multiple correlated uncertain parameters and also present an order selection algorithm for the technique. Although we tout PCM as a very economic technique, the number of simulation points nevertheless grows exponentially with the number of uncertain parameters. To overcome this difficulty, we develop some Information theory-based

techniques that can allow us to apply PCM using only a subset of the uncertain input parameters. We conclude our analytical development by discussing the possibility of applying PCM to solve optimization problems. In the penultimate chapter we illustrate a possible PCM application in Computer Science (specifically, in queueing theory) and also develop a larger electric power system example. We conclude the thesis by summarizing our results and discuss future directions. In our work and previous work, PCM has been used in such diverse areas as global climate evolution studies, chemical engineering applications and power systems analysis, which indicates the versatile nature of this algorithm.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	iii
ABSTRACT.....	v
LIST OF FIGURES.....	ix
LIST OF TABLES .....	xi
 CHAPTER	
1. INTRODUCTION .....	1
2. THEORY BEHIND THE PROBABILISTIC COLLOCATION METHOD .....	9
2.1 Orthogonal polynomials and Gaussian quadrature .....	10
2.2 One dimensional PCM .....	15
2.3 Improving PCM .....	26
3. MULTIPLE CORRELATED INPUTS: CONDITIONAL PCM.....	29
3.1 The two-dimensional PCM algorithm.....	30
3.2 Examples .....	33
3.3 Meshing PCM with Boundary Load flow .....	38
3.4 Order Selection Algorithm .....	43
4. OPTIMIZATION.....	49
4.1 Advantage of using PCM .....	54
4.2 Comparing PCM with a Traditional Minimization Technique .....	56

5. INFORMATION THEORETIC APPROACH FOR USING FEWER PARAMETERS IN PCM .....	60
5.1 Dependency measurement .....	61
5.2 Reducing the number of input variables .....	63
5.3 Justification.....	72
6. EXAMPLES .....	76
6.1 A larger example.....	76
6.2 Computer Science application .....	76
7. SUMMARY AND FUTURE WORK .....	87
BIBLIOGRAPHY .....	89



## LIST OF FIGURES

1.1	A general reduced order Black Box Model .....	1
2.1	Mappings with one and two stochastic inputs.....	10
2.2	Plot of the distribution for the pressure .....	18
2.3	PCM 1 <sup>st</sup> Order approximation vs. the actual function .....	22
2.4	PCM 2 <sup>nd</sup> Order approximation vs. the actual function .....	23
2.5	PCM 3 <sup>rd</sup> Order approximation vs. the actual function .....	24
2.6	Comparison of Output distributions .....	25
3.1	5-bus load flow example .....	33
3.2	Input distribution .....	34
3.3	Mapping between the two input parameters and the voltage output .....	37
3.4	The output distribution.....	37
3.5	PCM is meshed with a boundary load flow algorithm .....	40
3.6	Dynamic simulation example.....	41
3.7	PCM-generated mapping .....	42
3.8	Order selection, Case I.....	47
3.9	Order selection, Case II .....	47
4.1	Depiction of the optimization problem.....	49
4.2	Accuracy of PCM ls .....	52
4.3	Comparison of PCM polynomials .....	52
4.4	Minimum of a quartic function .....	55
5.1	Input distribution, Case I.....	67
5.2	Comparison of the output distributions, Case I.....	68
5.3	Comparison of the output distributions, Case II .....	70
6.1	IEEE 14-Bus test system.....	75
6.2	Jackson Network .....	81

6.3	Expected queue length plot.....	83
6.4	Busyness plot.....	84
6.5	Comparison of distributions for busyness.....	85
6.6	Comparison of distributions for queue length.....	83

## LIST OF TABLES

2.1	Orthogonal Polynomials and their roots .....	19
2.2	Comparison of mean and variance values .....	21
3.1	KL distance comparison .....	45
4.1	Comparison of PCM with Newton-Raphson .....	54
5.1	Mutual Information values.....	69
6.1	Distribution details .....	83

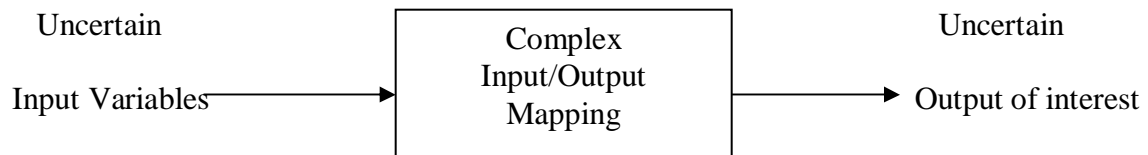
## **Dedication**

To my parents, Ramamurthy and Vaidehi,  
my Grandmother, Saraswathi  
and to my brother, Raja

## CHAPTER 1

### INTRODUCTION

Study of complex processes such as power networks dynamics, chemical reactions and global climate evolution often requires time-intensive simulation. When the parameters of such a process are uncertain, it is generally necessary to simulate the process over the range of parameter values to characterize the output(s) of the process. This thesis is concerned with intelligently choosing simulation points (parameter sets for simulations), so as to characterize the outputs of complex processes with minimal effort.



**Figure 1.1:** A general reduced order "black box" model representation.

We are specifically interested in characterizing the mapping between a set of uncertain parameters and an output of interest. We take the perspective that a low-order "black box" model can capture the mapping between the inputs and output. Such a "black box" model does not attempt to capture the operational intricacies of the system; rather, it tries to represent the relationship between the input variables and the output of interest based on observations of the system output at a finite set of input simulation points. Once the "black box" relationship between the input variables and the output of interest has been identified, such reduced order models could prove to be useful for analysis. However, to

come up with input/output mapping or to characterize the output, the system has to be simulated for a set of input values. “Economy” is the key word: techniques for coming up with such reduced order models with as few system simulations as possible could prove to be very useful. Traditionally, techniques such as brute force Monte Carlo simulation [16] were used for generating the mapping. The problem with such techniques is that they involve exhaustive simulations to characterize the outputs. If simulations are computationally intensive, characterization of outputs through exhaustive simulations may be infeasible.

Artificial Neural Networks (ANN) are also popular in the modeling arena; they are used for mimicking dynamic behavior of the system. Artificial Neural Networks map a set of input variables/patterns with corresponding output variables/patterns. A general ANN model consists of three layers viz. an input layer that carries the input information to the system, a hidden computational layer and an output layer. The input layer has connections, which has connection weights corresponding to it. The input values are multiplied by the weights and the weighted sum is formed. Each neuron has a threshold value called bias associated with it which is subtracted from the weighted sum. The computational layer applies an activation function to this weighted sum to produce the output. To determine the weights and biases and optimization procedure called training is used. We request the readers to refer to [20] and [21] for more information on ANN based modeling.

An alternative approach for intelligently choosing simulation points is to exploit probabilistic descriptions of the uncertain parameters. In other words, we would like to choose simulation points in such a way that the mapping between the parameter and output is accurately identified over the range of likely parameter values. The Probabilistic Collocation Method [1], [2], [3], [5] is a technique that can be used to model the deterministic relationship between uncertain parameters and an output of interest using polynomial functions. This is the approach that we shall take in this thesis.

The Probabilistic Collocation Method (PCM), also known as Deterministic Equivalent Modeling Method (DEMM), is a modeling technique that employs Gaussian quadrature [8] to characterize the relationship between uncertain input parameters and an output of interest. The output of interest is modeled as a polynomial of the uncertain input parameter(s). PCM was first used for global climate change studies [5]. In [1], [2] and [3] the authors apply PCM for modeling uncertainties in electric power systems. When probabilistic descriptions for the uncertain parameters are well known, it has been claimed that the Probabilistic Collocation Method (PCM) is more efficient compared to simulation techniques like Monte Carlo in terms of number of simulations required to capture the input-output relationship. For instance a simple power systems load flow analysis problem where we are want to find the effect of a particular uncertain parameter on the bus voltage or line flow may require 100's or sometimes even 1000's of simulations in the Monte Carlo approach whereas the same problem can be modeled with just a handful of simulations using PCM. PCM reduces the complexity by assuming a structured polynomial mapping between the uncertain input parameter(s) and the output

of interest and identifying a good set of simulations for correctly and robustly determining the mapping. The point selection is done based on Gaussian quadrature, which forms the crux of the theory behind PCM. Another interesting feature of PCM is that the same set of simulation points can be used for analyzing multiple output parameters.

It is interesting to consider some other modeling techniques for efficiently choosing simulation points under parameter uncertainty. The Stochastic Response Surface Method (SRSM) [14], [15] is an uncertainty modeling technique used mainly in the field of chemical and bio-medical engineering. In SRSM the inputs of the system are represented as functions of certain standard random variables (srvs) and each output under examination is expressed as a series expansion in terms of the srvs as multidimensional Hermite polynomials. The reasoning behind this representation is that it offers consistency, as the srvs are well behaved and mathematically tractable. The mapping between the input and the output can be established by estimating the coefficients of the output series expansion and this is achieved by collocation methods, like PCM, or regression methods. The authors discuss PCM for this purpose and, interestingly, they renounce it on the grounds that PCM becomes unwieldy when the number of input parameters is large. We have proposed a few techniques to address this issue, which is one of our major contributions of this study. The authors adopt a regression based collocation method for estimating the coefficients, which they address as regression based SRSM. It requires twice as many collocation points as there unknown coefficients,



for estimation. Moreover, in [15], the authors claim that SRSM maybe may be more useful in the case of complex nonlinear models.

The Stochastic Collocation Method (SCM) [17], used mainly in the field of fluid dynamics, transforms the physical random variables to an artificial stochastic space with known properties and then uses a collocation-based approach for modeling the relationship between the physical random variable and the output of interest.

Unlike PCM, the techniques mentioned above are quite complex to implement. PCM is appealing because it is simple and yet allows the evaluation of complicated output functions.

## **CONTRIBUTION**

Our main contributions towards PCM are

1. Extending the technique to handle systems with multiple correlated uncertainties. In [1], [2] one-dimensional PCM is discussed with respect to power systems analysis. Although the authors propose a version of PCM for handling multiple uncertain parameters, the technique is unwieldy and in some cases turns out to be inaccurate. In this thesis we have proposed our generalization of PCM for handling multiple correlated uncertain parameters, which is robust and accurate.

2. Devising a systematic heuristic algorithm for selecting the appropriate order for the PCM generated polynomial for a particular system.

Whenever use reduced-order modeling, one of the problems that immediately arise is deciding when to stop or, in the case of PCM, deciding what order of polynomial is sufficient for modeling the input-output mapping. We have developed an algorithm for order selection in single-dimensional PCM and also discussed its relevance to the multi-dimensional case.

3. Proposing techniques for reducing the number of uncertain input parameters

Motivated by the presence of strong correlation among power systems parameters we have proposed some Information theory based approaches for filtering out or reducing uncertain input parameters, i.e. using only a subset of the set of uncertain input parameters for developing the input-output stochastic mapping.

4. Applying PCM for optimization problems and extending the scope of PCM beyond power systems analysis.

PCM was first used for global weather change studies and then in the context of power systems. This indicates the underlying universal nature of the algorithm. We have attempted to extend PCM's scope by identifying possible applications in other fields. We do this by applying PCM for optimization related problems.

## THESIS ORGANIZATION

The remainder of this thesis is organized as follows.

- We review PCM and its underlying theory of Gaussian quadrature in Chapter 2.
- In Chapter 3 we describe our generalization of PCM for systems with multiple correlated uncertain parameters and also present an order selection algorithm for PCM, the previous chapter talks about PCM for single uncertain parameter analysis, this multidimensional generalization is one of our main contributions to this study. We have presented the theory with some interesting examples which illustrate our approach.
- Although we tout PCM as a very economic technique, the number of simulation points nevertheless grows exponentially with the number of uncertain parameters. Some Information theory based techniques that can be used to reduce the number of uncertain input parameters are discussed in Chapter 4.
- This thesis not only aims to present the effectiveness of PCM as a modeling technique for evaluating uncertainties in power systems but also attempts to enunciate PCM as a universal algorithm finding use in a variety of fields ranging from computer science to chemical engineering. The Optimization problem is an interesting and ubiquitous one, in Chapter 5 we discuss this problem and also

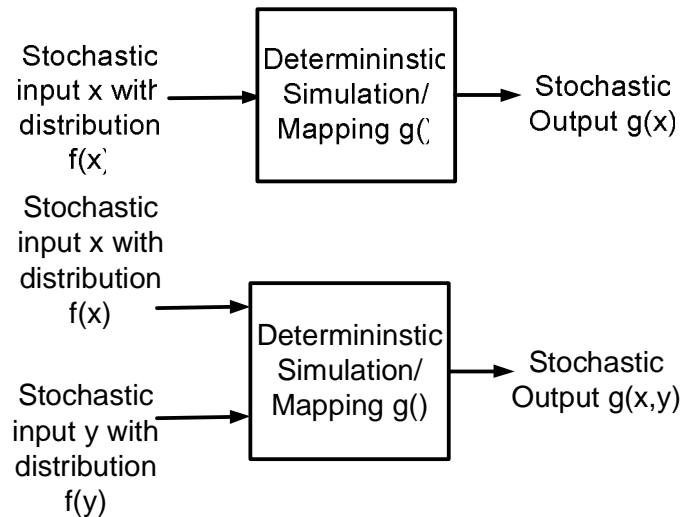
present the application of PCM to this problem to show the versatility of the technique.

- Although examples are provided in every chapter, a larger example that applies several of the concepts presented thus far is needed to demonstrate the applicability of the method. In Chapter 6, we apply our techniques to a larger power systems example in which the loads are classified as industrial and commercial, with loads of each type having strong interdependencies. Apart from this example, this chapter also discusses the application of PCM in other areas such as multicasting in wireless computer networks and modeling of chemical reaction dynamics.
- Chapter 7 concludes and summarizes the thesis and also discusses possible directions to move in the future.

## CHAPTER 2

### THEORY BEHIND THE PROBABLISTIC COLLOCATION METHOD

The **probabilistic collocation method** is a means for developing a parametric model for the deterministic mapping between a stochastic input and an output (Figure 2.1), using only a small number of simulations of the system. In particular, ***n*th-order PCM** seeks to represent the mapping using an *n*th-degree polynomial, whose coefficients are found by matching the model predictions with simulation outputs for a particular set of *n+1* input values. The *n+1* input values---henceforth called the **PCM points**---are specially chosen, in a manner that makes the fit robust to some possible errors in the model's parameterization. Specifically, the *n+1* PCM points are chosen so that the mean output predicted by the model is identical to the actual mean output, if in fact the mapping is a polynomial of any degree less than or equal to  $2n+1$ . Thus, PCM specifies a low-order mapping that approximates a much higher-order (in other words, more detailed) mapping, in the sense that the mean output predicted by both mappings is identical.



**Figure 2.1:** Mappings with one and two stochastic inputs, respectively, are shown. PCM can be used to characterize the mapping  $g()$  and the probability distribution of the output with a small number of simulations.

The specialty of PCM lies in its propensity for selecting simulation (or collocation) points in the high probability region of the input distribution, and it is this feature that makes PCM a cost effective modeling technique. The theory behind PCM is based on the concepts of Orthogonal polynomials and Gaussian quadrature. So, before elucidating the PCM mechanism, we find it necessary to throw some light on the above mentioned concepts.

## 2.1 ORTHOGONAL POLYNOMIALS AND GAUSSIAN QUADRATURE

Gaussian quadrature is a special kind of numerical integration technique. The advantage, which Gaussian quadrature provides over traditional numerical integration, is the freedom to select points at which the given function can be evaluated. This means that the number of points at which the function has to be evaluated is significantly reduced. Apart from the economy aspect, it has been claimed that the results obtained using Gaussian quadrature are more accurate compared to traditional numerical integration techniques like the Simpson's rule or the trapezoidal rule.

Gaussian quadrature uses orthogonal polynomials for the purpose of selecting points. The typical form of integrals in Gaussian quadrature is

$$\int_F f(x)g(x)dx \tag{2.1}$$

$g(x)$ , is an orthogonal polynomial.  $f(x)$ , is a non-negative weighing function defined in the connected space  $F$  and the above expression is known as an inner product.

Before getting into the theory behind Gaussian quadrature and the Probabilistic Collocation Method it is necessary to pen down some important details about orthogonal polynomials.

## Orthogonal Polynomials

Orthogonal polynomials are polynomials that are orthogonal to each other with respect to an inner product. We find it worthwhile to reproduce the definition of inner product and orthogonal polynomials from [2] and [8].

Given a real linear space of functions  $F$ , an inner product  $\langle f, g \rangle$  (we represent inner product by angled brackets) defined on  $F$  is a function of  $f, g \in F$  satisfying the following conditions

$$\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle \quad (2.2)$$

$$\langle \alpha f, g \rangle = \alpha \langle f, g \rangle = \langle f, \alpha g \rangle, \text{ where } \alpha \text{ is a scalar} \quad (2.3)$$

$$\langle f, g \rangle = \langle g, f \rangle \quad (2.4)$$

$$\langle f, f \rangle > 0, \text{ if } f \neq 0 \quad (2.5)$$

For example, consider two polynomials  $g(x)$  and  $h(x)$  if  $f(x)$  is any non-negative weighting function defined on the space, then  $\langle g(x), h(x) \rangle \triangleq \int_F f(x)g(x)h(x)dx$  is an inner product.

This expression is very important as it is the peculiar inner product that forms the basis for Gaussian quadrature integration and the Probabilistic Collocation Method. The polynomials  $g(x)$  and  $h(x)$  are said to be orthogonal if their inner product is zero.

### Orthonormal polynomials

A set of polynomials in the space  $H$  are said to be orthonormal if and only if the following relationship exists for all  $h_i(x)$  in  $H$ .

$$\langle h_i, h_j \rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (2.6)$$

The subscript of the polynomial indicates its degree i.e.  $h_i(x)$  has degree  $i$ . An important property of these orthonormal polynomials is that they are unique and they form a basis for the space for all polynomials. Another important property of these orthonormal polynomials is that their roots depend only upon the weighting function  $f(x)$ . Further, all the roots are contained in the space  $F$ , and each orthonormal polynomial  $h_i$  has exactly  $i$  roots. The roots of these polynomials form the collocation points for the Probabilistic Collocation Method. The set  $\{h_i\}$  of orthonormal polynomials of increasing degree form the backbone of PCM.



## GAUSSIAN QUADRATURE:

As mentioned before, Gaussian quadrature is a special numerical integration technique for integrals of the form

$$\int_F f(x)g(x)dx$$

In a nutshell, Gaussian quadrature seeks to obtain the best numerical estimate for the above integral and it does so by picking certain  $x$  values, evaluating  $g(x)$  at these points, and then computing the integral. The  $x$  values are the roots of the orthogonal polynomials, discussed in the previous section.

The result of Gaussian quadrature integration is the following formula

$$\int_F f(x)g(x)dx \approx \sum_{i=1}^n f_i g(x_i) \quad (2.7)$$

The coefficients  $f_i$  depend on the weighting function and the function  $g(x)$  is evaluated at abscissa values that are the roots of the  $n^{\text{th}}$  orthogonal/orthonormal polynomial calculated with respect to the weighting function  $f(x)$ .

The above integral is exactly correct when  $g(x)$  is a polynomial of degree  $(2n-1)$ .

Interestingly, the integral can be estimated using just  $n$  samples. This shows that the Gaussian quadrature has the ability to represent a higher order relationship using a lower order polynomial; PCM inherits this property from Gaussian quadrature.

## Proof for Gaussian Quadrature

We find it worthwhile to paraphrase the proof for Gaussian quadrature from [2]. The polynomials in  $H$  up to and including order  $i$  form an orthonormal basis for the space of all polynomials of degree less than or equal to  $i$ . Then, a polynomial of order  $(2n - 1)$  can be expressed as follows:

$$g(x) = h_n(x)(a_{n-1}h_{n-1}(x) + \dots + a_0h_0(x)) + b_{n-1}h_{n-1}(x) + \dots + b_0h_0(x) \quad (2.8)$$

Note,  $h_0(x)$  is a constant. Hence by orthogonality the Gaussian quadrature integral can be expressed as follows:

$$\int_F f(x)g(x)dx = b_0 \int_F f(x)h_0(x)dx \quad (2.9)$$

By evaluating the function  $g(x)$  at the  $n$  roots of the orthogonal polynomial  $h_n(x)$  we get the following set of linear equations:

$$\begin{bmatrix} g(x_1) \\ \vdots \\ g(x_n) \end{bmatrix} = \begin{bmatrix} h_{n-1}(x_1) \cdots h_0(x_1) \\ \vdots \\ h_{n-1}(x_n) \cdots h_0(x_n) \end{bmatrix} \begin{bmatrix} b_{n-1} \\ \vdots \\ b_0 \end{bmatrix} \quad (2.10)$$

To solve for  $b_0$ , we need to invert the above expression. If  $h_0(x)$  is chosen to be equal to 1, as it generally is then our desired result is  $b_0$ .

$$\begin{bmatrix} b_{n-1} \\ \vdots \\ b_0 \end{bmatrix} = \begin{bmatrix} h_{n-1}(x_1) \cdots h_0(x_1) \\ \vdots \\ h_{n-1}(x_n) \cdots h_0(x_n) \end{bmatrix}^{-1} \begin{bmatrix} g(x_1) \\ \vdots \\ g(x_n) \end{bmatrix} \quad (2.11)$$

$$b_0 = \int_F f(x)g(x)dx \approx \sum_{i=1}^n f_i g(x_i) \quad (2.12)$$

Where, the weights  $f_i$  are given by the last row of the matrix in (2.11).

## 2.2 ONE DIMENSIONAL PCM

Given an input random variable  $x$  with pdf  $f(x)$  and an output of interest, we seek to approximate the functional mapping  $g(x)$  that transforms the input to the output. Notice that the mean value of the output in this case is given by

$$E(x) = \int_F f(x)g(x)dx \quad (2.13)$$

Gaussian quadrature allows us to choose  $n+1$  points  $x_1, \dots, x_{n+1}$ , such that, for any  $g^*(x)$  that is a polynomial of order less than or equal to  $2n+1$  and for which  $g^*(x_1) = g(x_1), \dots, g^*(x_{n+1}) = g(x_{n+1})$ , the integral  $\int_F f(x)g^*(x)dx$  is the same. Thus, the mean value predicted by the degree- $(n+1)$  polynomial that passes through these points is the same as the mean predicted by any polynomial of degree less than or equal to  $2n+1$  that passes through the points. Equivalently, the degree- $(n+1)$  polynomial suffices to capture the mean output, if the mapping is indeed a polynomial of degree less than or

equal to  $2n+1$ . The Gaussian quadrature points (in this case the PCM points) are determined by computing the first  $n+1$  **orthogonal polynomials** with respect to  $f(x)$ .

Once the  $(n+1)$  PCM points are generated, the function under study is simulated at these points. The  $n^{\text{th}}$  order PCM polynomial will be of the form

$$\hat{g}(x) = a_0 + a_1x + \dots + a_nx^n \quad (2.14)$$

By substituting the  $(n+1)$  PCM points in the above equation we get  $n$  equations, and by solving these equations using the value of the function under study at these  $(n+1)$  PCM points, we can obtain the coefficients of the  $n^{\text{th}}$  order PCM polynomial.

### **An Example:**

To illustrate the Probabilistic collocation method, we have chosen an example from the field of physical chemistry.

The “Ideal Gas” law [22] is an equation that describes the physical behavior of an ideal gas. It combines three primitive gas laws viz. Boyle’s Law, Charles’s Law and Avogadro’s Law. The equation relates the pressure  $P$ , volume  $V$  and the temperature  $T$  of an ideal gas. In the same vein, an ideal gas is one whose physical properties satisfy the ideal gas equation.

The ideal gas law is stated as follows

$$PV = nRT \quad (2.15)$$

P is the pressure of the gas

V is the volume of the gas

n is the number of moles of the gas

R is the universal gas constant,  $R = 0.0821$

T is the temperature in kelvin

For our purpose this brief introduction to the “Ideal Gas” law would suffice.

Typical problems related to the “Ideal Gas” equation would be, finding the value of one of the entities given the rest.

Our first PCM example in this thesis is an attempt to model the “Ideal Gas” law. Having the actual relationship in hand helps as the PCM generated polynomial model can be compared with the actual analytical relationship. Another reason for the choice of this example is just to illustrate the prospect of PCM as an algorithm that can be used in several fields of study.

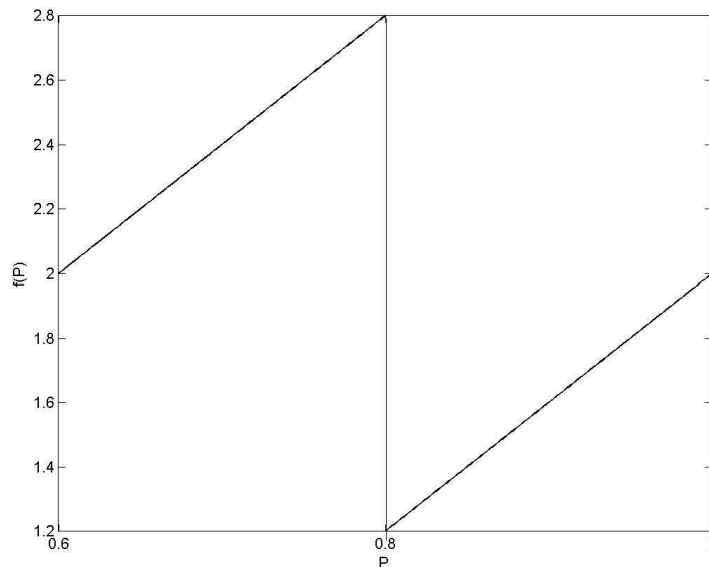
For a particular gas, we have made the following assumptions, with the temperature T at absolute zero (273 K) we want to find out the volume occupied by 1 mole of the gas in litres (1 litre = 0.264172051 gallon) when the pressure is randomly distributed between 0.6 and 1 atm.

We would like to remind the readers that this example is only for the purpose of illustrating PCM. Otherwise PCM, or any other uncertainty analysis technique for that matter, would be obviated for such an example, as the relationship between the uncertain parameter and the output of interest can trivially be computed analytically.

Just to make the analysis interesting and also to show that PCM can handle any kind of distribution, we have chosen an unconventional PDF for the pressure.

$$f(P) = \begin{cases} 4P - (2/5), & 0.6 < P < 0.8 \\ (4P - 1), & 0.8 < P < 1 \end{cases} \quad (2.16)$$

The PDF is depicted below in Figure 2.2



**Figure 2.2:** Plot of the distribution for the pressure, P

Using the above PDF the first few orthogonal (or orthonormal) polynomials are generated, the roots of which are the pressure values for which the Volume is calculated using the Ideal Gas equation.

The table below lists the orthogonal polynomials up to order 4, with their roots.

Orthogonal Polynomials
$h_0(P) = 1$
$h_1(P) = 8.69P - 7.03$ $P = 0.8933$
$h_2(P) = -84.34P^2 - 135.89P + 53.61$ $P = \{0.6905, 0.9207\}$
$h_3(P) = 829.75P^3 - 1997.69P^2 + 1583.37P - 412.96$ $P = \{0.6481, 0.8022, 0.9573\}$
$h_4(P) = -8240.93P^4 - 26450P^3 + 31553P^2 - 16576.4P + 3235.28$ $P = \{0.6293, 0.7354, 0.8714, 0.9734\}$

**Table 2.1:** Orthogonal polynomials and their roots.

The Volume is then calculated by substituting the roots of the orthogonal polynomials into the gas equation and the coefficients of the PCM polynomial are obtained by solving. For example, if we want the PCM quadratic polynomial, say  $g(P)$ , for the relation under study, we have to use the roots of the 3<sup>rd</sup> order orthogonal polynomial. In general, the roots of the order  $n$  orthogonal polynomial are used to generate the order  $(n-1)$  PCM polynomial.

The roots of the 3<sup>rd</sup> order orthogonal polynomial are 0.648104, 0.802175 and 0.957301 and the corresponding Volume values are

$$V_1 = 34.5661$$

$$V_2 = 27.9271$$

$$V_3 = 23.4016$$

We need a polynomial of the form  $aP^2 + bP + c$ . Using the values of the roots and the corresponding  $g()$  values, we obtain 3 equations solving which gives us the values of the coefficients.

Thus, the PCM quadratic for the relationship is as follows,

$$g(P) = 45.01P^2 - 108.37P + 85.894 \quad (2.12)$$

We would also like to present the linear and cubic PCM approximations

$$g(P) = -35.24P + 56.78 \quad (2.13)$$

$$g(P) = -57.048P^3 + 183.11P^2 - 218.45P + 114.77 \quad (2.14)$$



## ANALYSIS

We present plots comparing the polynomials generated by PCM with the actual function.

The actual function plot is obtained by exhaustively simulating the equation  $V = \frac{nRT}{P}$ .

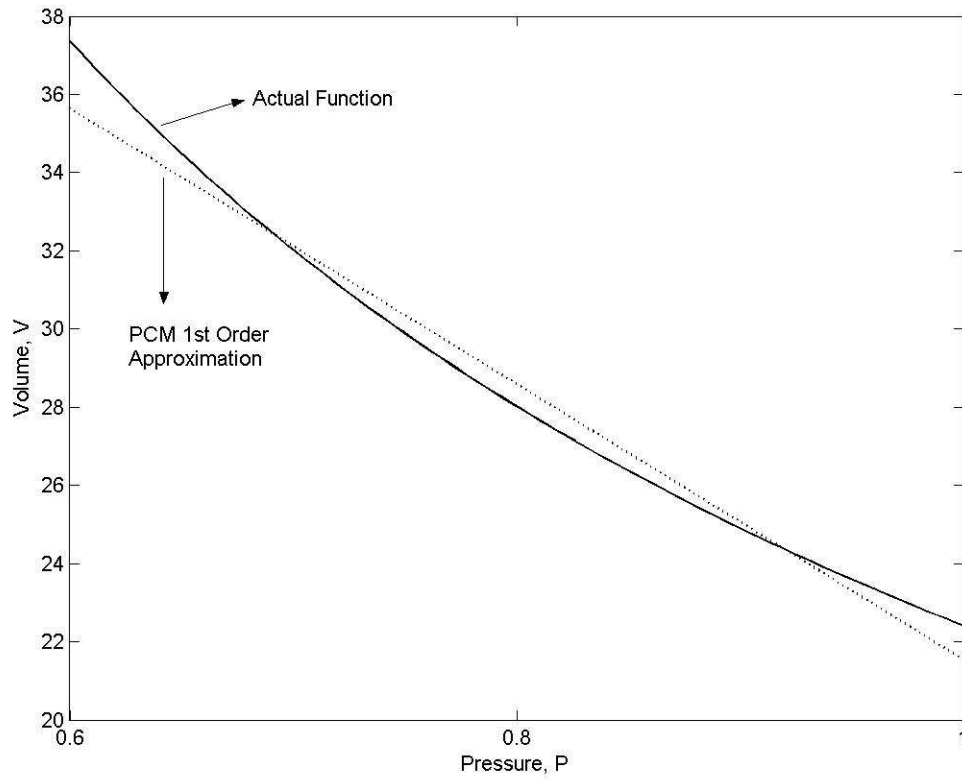
The plots show the accuracy of PCM, figure 2.4 reveals that the PCM quadratic is very close to the actual function, and from figure 2.5 it can be observed that the PCM cubic is so close to the actual function that it is hard to differentiate between them. The power of PCM is such that with just 4 simulations we are able to model the relationship between the Volume and pressure.

An important attribute of PCM mentioned before, is its penchant for getting the mean value of the output right. To illustrate this, table 2.2 presents the expected value and variance of the PCM polynomials along with those of the actual function.

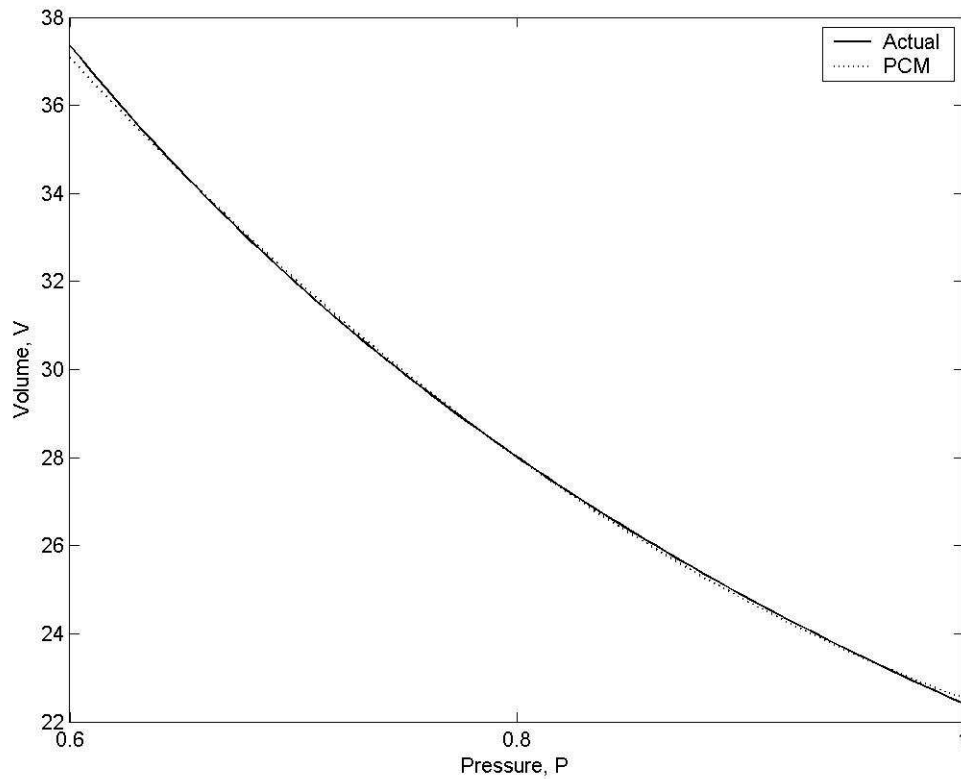
Function	Expected Value and Variance
PCM 1 <sup>st</sup> Order Polynomial	E = 28.2569, $\sigma^2 = 16.4515$
PCM 2 <sup>nd</sup> Order Polynomial	E = 28.2668, $\sigma^2 = 17.3091$
PCM 3 <sup>rd</sup> Order Polynomial	E = 28.267, $\sigma^2 = 17.333$
Actual Function	E = 28.2807, $\sigma^2 = 17.35$

**Table 2.2:** Comparison of mean and variance values of different order PCM predictions with that of the actual function.

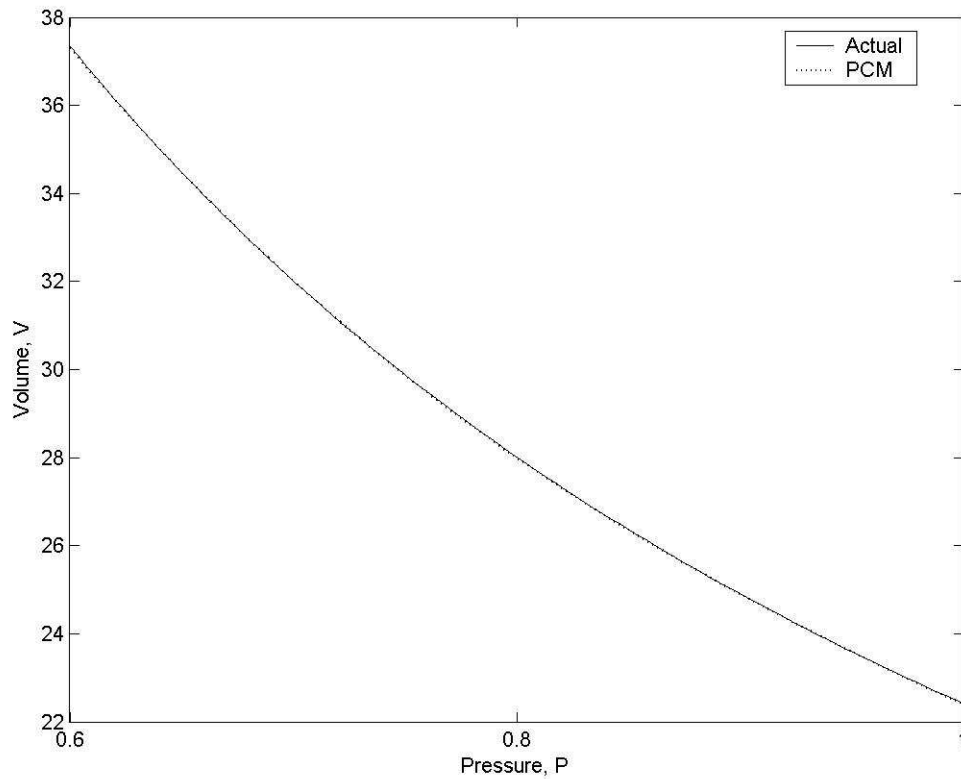
**PLOTS:**



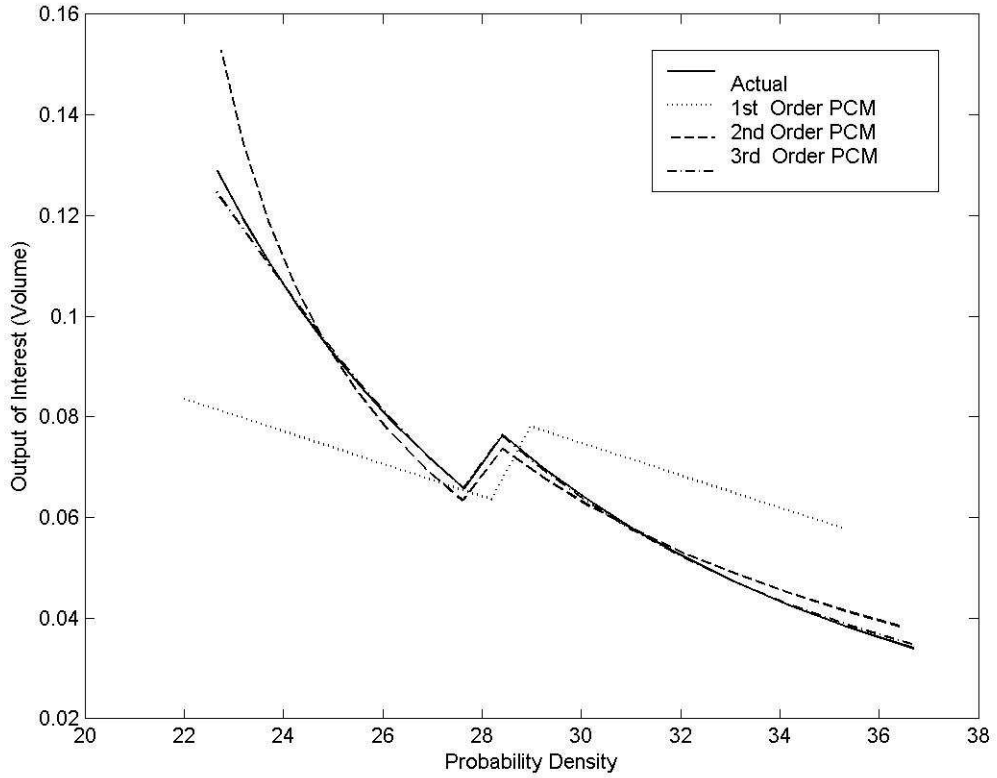
**Figure 2.3:** PCM linear approximation and the Actual function plotted through exhaustive simulation. The solid line represents the actual function and the dotted line is the PCM linear polynomial.



**Figure 2.4:** PCM 2<sup>nd</sup> Order and the Actual mapping. The solid line represents the actual mapping and the dotted line is the quadratic PCM. The two plots are identical except at the upper endpoint, where they disagree slightly.



**Figure 2.5:** 3<sup>rd</sup> Order PCM and the Actual mapping. The solid line represents the actual mapping and the dotted line is the cubic PCM polynomial. The two plots cannot be differentiated; this shows that the PCM cubic is an accurate approximation of the relationship between Volume and Pressure in this example.



**Figure 2.6:** Comparison of the output probability distribution when PCM is used with the actual output distribution. The solid line represents the actual function's distribution. The dotted, dash, and dot-dash lines represent the output distributions based on the PCM 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> order approximations respectively.

## 2.3 IMPROVING PCM

In the remainder of this chapter, we present two new results concerning PCM, namely the improvement of the PCM approximation using sensitivity information and the computation of error bounds for PCM approximations. These analyses follow naturally from well-known results in the Gaussian quadrature community, but have not heretofore been considered in the study of PCM. In presenting these results, we hope to briefly introduce the reader to relevant literature on quadrature, and to highlight some distinctions in analysis goals for quadrature and PCM, respectively.

### **Improving PCM Using Sensitivity Information**

Some simulation programs not only can compute the output for a particular parameter value, but also can determine the sensitivity of the output to the parameter. For instance, efficient means for characterizing sensitivities of power flow outputs to loading parameters have been developed, and when sensitivity information is available, we might expect to obtain a more detailed characterization of the input-output mapping. That is, since we have additional information of the input-output mapping in particular, knowledge of the derivative of this mapping at the simulation points we should be able to generate a more accurate approximation of the mapping.

In particular, let us assume that the output  $g(x)$  and the sensitivity of the output  $\frac{d(g(x))}{dx}$  to the parameter  $x$  have been found at the  $(n+1)$  PCM points  $x_1^*, \dots, x_{n+1}^*$ . Then we recommend fitting the mapping using a degree- $(2n+1)$  polynomial (as opposed to a degree  $n$  polynomial for PCM), which matches both the output data and output sensitivities. That is, we recommend approximating the mapping using the

degree- $(2n+1)$   $\hat{g}_d(x) = \alpha_1 x^{2n+1} + \dots + \alpha_{2n+1} x + \alpha_{2n+2}$  that satisfies the equality,

$$\hat{g}_d(x_i^*) = g(x_i^*), 1 \leq i \leq n+1$$

$$\frac{d \hat{g}_d(x_i^*)}{dx} = \frac{dg(x_i^*)}{dx}, 1 \leq i \leq n+1$$

It is easy to check that these  $(2n+2)$  equalities give  $(2n+2)$  independent linear relations for the parameters  $\alpha_1, \dots, \alpha_{2n+2}$ , and hence that  $\hat{g}_d(x)$  is determined uniquely from the known outputs and sensitivities. Let us refer to the approximation  $\hat{g}_d(x)$  as the  $n^{\text{th}}$  order PCM-with-sensitivity approximation.

It is worth making several observations about the PCM-with sensitivities approximation. First, we note that, if the actual mapping  $g(x)$  is a degree- $(2n+1)$  polynomial, the  $n^{\text{th}}$  PCM-with-sensitivity approximation is identical to the actual mapping. We should not be surprised that the mapping can be identified exactly, since we have available  $(2n+2)$  independent data points ( $n+1$  output values and  $n+1$  sensitivities). In fact, any set of  $(2n+2)$  independent measurements can be used to identify the mapping and the PCM-

with-sensitivities approach is only special in that only  $(n+1)$  simulations maybe needed (if sensitivities are automatically generated).

What is more surprising is that the PCM-with-sensitivities approximation  $\hat{g}_d(x)$  can be guaranteed to be close to the actual mapping, even when the mapping is not a polynomial of degree  $2n+1$ . It turns out that  $\hat{g}_d(x)$ , known in the quadrature community as a *Hermite polynomial*, is used in generating error bounds for Gaussian quadrature ([25], see also [26] for a succinct description of Markov analysis).

#### **CHAPTER SUMMARY:**

The results illustrate the following

- The accuracy of PCM in modeling a deterministic mapping between uncertain parameter and an output of interest.
- The economy of PCM

In the chapters to follow, we will delve deeper into these two points.



## CHAPTER 3

### MULTIPLE CORRELATED INPUTS: CONDITIONAL PCM

Our studies of One-dimensional PCM show that it is very economic compared to techniques like Monte Carlo simulation. Hence we present an extension of PCM for handling systems with multiple correlated uncertainties.

For convenience, let us first discuss our generalization of PCM to systems with *two* correlated, uncertain inputs (see Figure 2.1). We call this generalization **two-dimensional PCM**. We assume that the two uncertain inputs  $x$  and  $y$  are jointly distributed according to a density function  $f(x,y)$  that is non-zero over a finite, convex two-dimensional domain  $A$ . Our aim is to identify the mapping  $g(x,y)$  that specifies the output in terms of these inputs. We assume that this mapping can be approximated by a two-dimensional multinomial of the form

$$g^*(x, y) = \sum_{i=0}^n \sum_{j=0}^n a_{ij} x^i y^j \quad (3.1)$$

Henceforth, we refer to  $g^*(x,y)$  as a **generalized polynomial of degree  $n$** . We feel that a generalized polynomial representation for a two-dimensional mapping is appropriate, because (as in the one-dimensional case) higher-degree generalized polynomials provide more and more detailed representations of the mapping. More specifically, an order- $n$  generalized polynomial representation allows us to specify a set of polynomial mappings between each single input and the output, given the other input. To determine the coefficients in (2), we simulate the output for a particular set of  $(n+1)^2$  input pairs, which we again call PCM points. From the  $(n+1)^2$  outputs at the PCM points, we

determine the  $(n+1)^2$  coefficients by solving a system of linear equations. As in the one-dimensional case, the success of two-dimensional PCM depends strongly on appropriate choice of the PCM points.

### 3.1 THE TWO- DIMENSIONAL PCM ALGORITHM

We propose the following algorithm for choosing the PCM points:

1. We compute the marginal distribution for the input  $x$  as  $f(x) = \int_A f(x, y) dy$ . We then find the degree- $(n+1)$  orthogonal polynomial with respect to  $f(x)$ , and find the roots of this polynomial. Notice that these are the  $x$  values that we would choose as PCM points, if we were applying one-dimensional PCM of order  $n$  to find a mapping between  $x$  and an output. Let us label these points  $x_1, \dots, x_{n+1}$ .
2. We compute the conditional distributions  $f(y | x_i) = \frac{f(x_i, y)}{f(x_i)}$ . We then find the degree- $(n+1)$  orthogonal polynomials with respect to each distribution, and find the roots of these polynomials. Let us call the roots of the orthogonal polynomial with respect to  $f(y | x_i)$  as  $y_1(x_i), \dots, y_{n+1}(x_i)$ .
3. We use the  $(n+1)^2$  pairs of inputs  $[x_i, y_j(x_i)]$ ,  $1 \leq i \leq n+1, 1 \leq j \leq n+1$ , as the PCM points.

The following analytical results (presented without proof) can be deduced for two-dimensional PCM; these results motivate use of the method:

1. Given that the input  $x$  is any one of the values  $x_1, \dots, x_{n+1}$ , the mean output is correctly predicted by two-dimensional PCM whenever the actual mapping is a generalized polynomial of degree less than or equal to  $2n+1$ . Also, from continuity arguments, we can argue that the mean output predicted by PCM is nearly correct for inputs  $x$  that are close to one of the points  $x_1, \dots, x_{n+1}$ . Since the points  $x_1, \dots, x_{n+1}$  are chosen to reflect the high-probability domain for the input  $x$  (this is one of the benefits of one-dimensional PCM), two-dimensional PCM predicts the mean output correctly given likely values for  $x$ .
2. In the special case that  $x$  and  $y$  are in fact independent, the (unconditioned) mean value for the output is correctly predicted by PCM whenever the mapping is a generalized polynomial of degree less than or equal to  $2n+1$ . Further, in the more general case that  $x$  and  $y$  are not independent but the  $r$ th-conditional moment for  $y$  given  $x$  is an  $r$ th-order polynomial, PCM predicts the output mean whenever the actual mapping is a true two-dimensional polynomial of degree less than or equal to  $2n+1$  (i.e, a sum of monomial terms, each of which has total degree less than or equal to  $2n+1$ ).
3. The PCM points always fall within the region  $A$ , so that we should be able to simulate a meaningful output for each PCM point.

We note that PCM can easily be generalized to identify mappings between three or more uncertain inputs and an output. As in two-dimensional PCM, we can select PCM points for higher-dimensional PCM recursively from a sequence of marginal and conditional distributions. These higher-dimensional PCM algorithms are amenable to the same analyses as two-dimensional PCM.

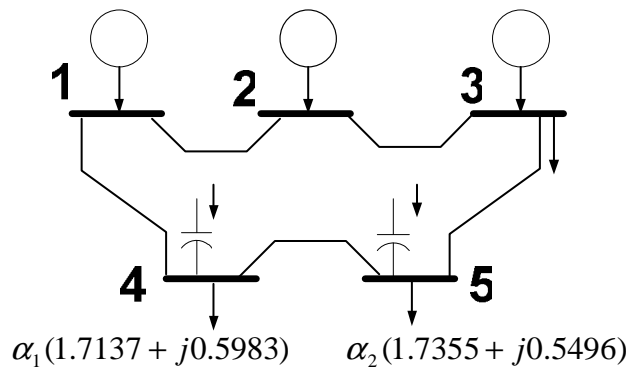
In the remainder of this chapter, we apply two-dimensional PCM to characterize the mapping between the two uncertain loads and the load flow voltage at a bus in a power system. Our study is in the context of a toy example obtained from [10], and is not meant to provide a comprehensive depiction of load flow uncertainties by any means. Our primary purpose is to illustrate two-dimensional PCM, and to explore some potential benefits and caveats of using PCM to characterize load flow solutions.

PCM-based characterization of load flow voltages falls within the broad class of probabilistic load flow (PLF) algorithms (see [10] for a summary of some work on PLF). These are methods for computing uncertainties on load flow solution parameters (e.g., bus voltages or line loadings), given uncertainty distributions on load powers and other system parameters. A full study of the literature on PLF is beyond the scope of this article, but we present a few general concepts. As discussed in, e.g., [7] and [11], PLF algorithms are either based on Monte Carlo simulation techniques, on exact analysis, or on some combination of these. Very often, analytical methods assume a load flow model that is linearized around one or multiple equilibria, and require some structure (e.g., Gaussianity) in the parameter distributions. Monte Carlo techniques account for the

nonlinearities in the load flow solution and allow for general input parameter distributions, but are computationally intensive. As an alternative to PLF, load flows for systems with uncertain inputs have also been characterized by identifying limits on the output variables given limits or distributions on the inputs (e.g., [7], [12]). These methods, called boundary load flow algorithms, have recently been combined with techniques that provide fuzzy-set descriptions of output variables, given fuzzy descriptions of input variables [7].

### 3.2 EXAMPLES

We believe that PCM can contribute to PLF analysis, by providing an intelligent simulation strategy and also by providing a method for meshing probabilistic and boundary methods.



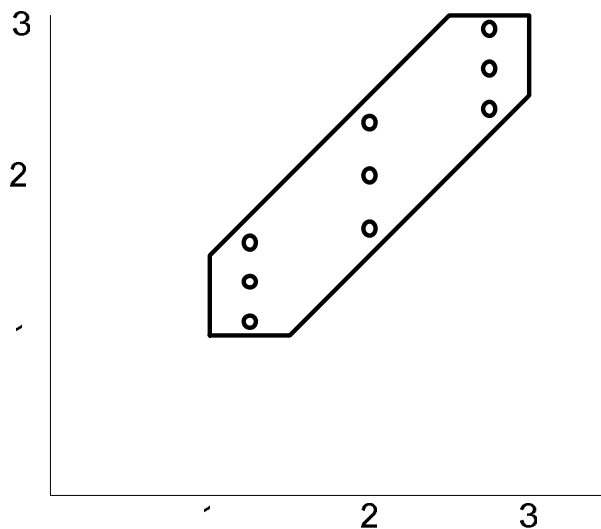
**Figure 3.1:** Load flow example .We apply PCM to characterize the voltage at bus 4, given that the loads at buses 4 and 5 are uncertain.

We apply PCM to find the PLF solution in the small power system example shown in Figure 3.1. In this example, we assume that the scaling parameters (inputs)  $x$  and  $y$  (see

Figure 3.1), which specify the load power magnitudes as buses 4 and 5, are jointly distributed as shown in Figure 3.2. The positive correlation is meant to reflect that load requirements tend to be correlated with external parameters (e.g., temperature), which are roughly constant over a set of loads. Our output variable is the magnitude of the voltage at bus 4. Application of PCM to this example first requires computation of the PCM points; the nine points for second-order PCM are shown on Figure 3.2. Using the PCM collocation points, we characterize the mapping between the inputs and output. The second-order generalized representation for the mapping found using PCM is the following:

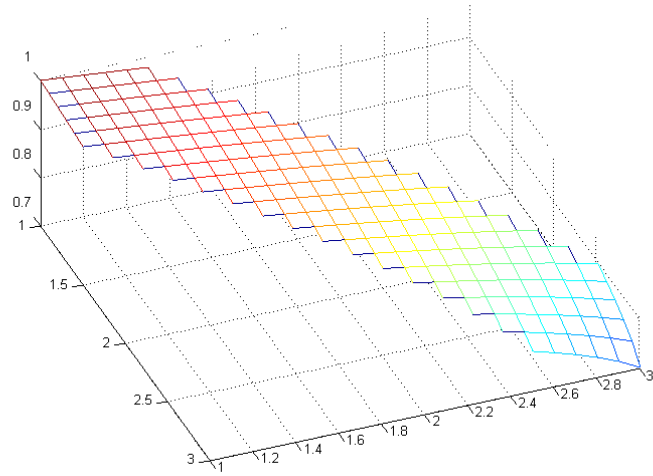
$$g^x(x, y) = -0.041x^2y^2 + 0.2x^2y - 0.24x^2 + 0.12xy^2 - 0.61xy + 0.72x - 0.1y^2 + 0.45y + 0.48$$

This predicted mapping is compared to the actual mapping (generated through exhaustive simulation) in Figure 3.3. Finally, we numerically determine the distribution for the output variable and compare it to the actual output distribution in Figure 3.4.

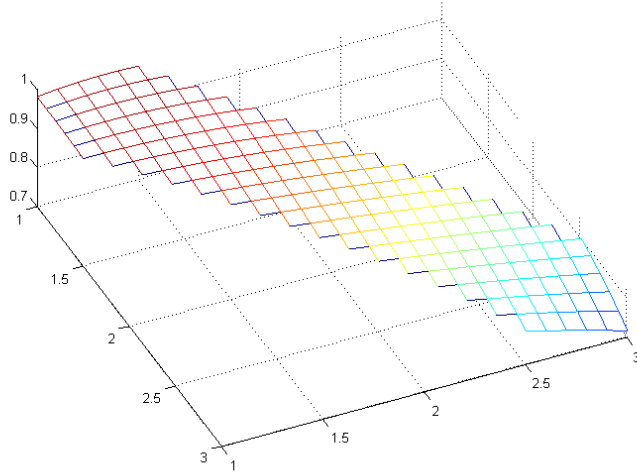


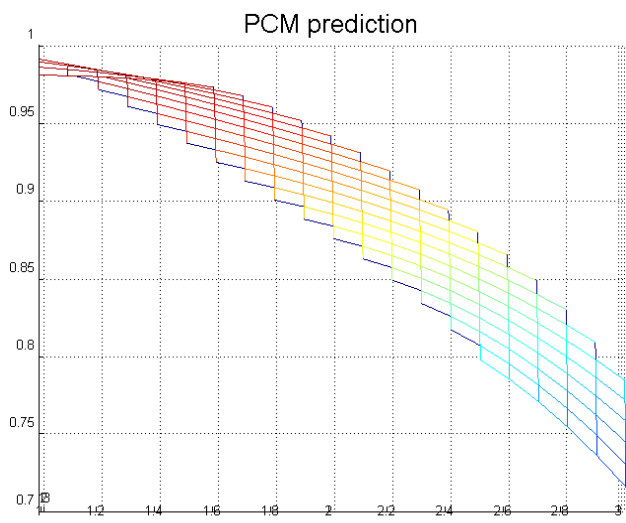
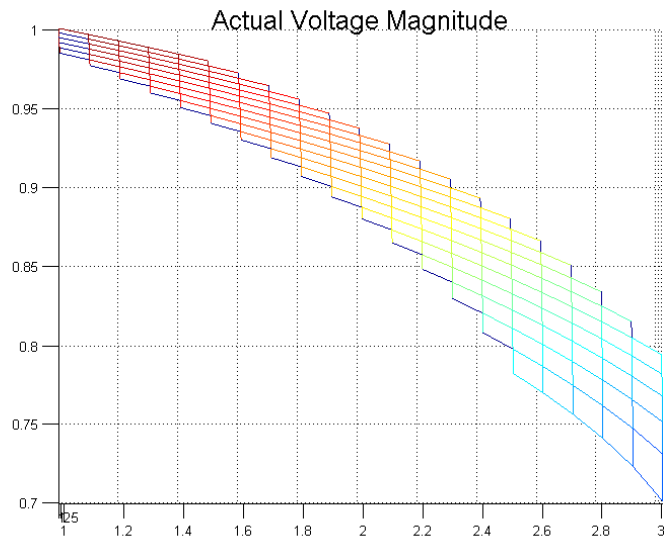
**Figure 3.2:** Input distribution and PCM points. The parameters (inputs)  $x$  and  $y$  are distributed uniformly over the polygonal region shown. The PCM points are also illustrated.

Actual Voltage Magnitude



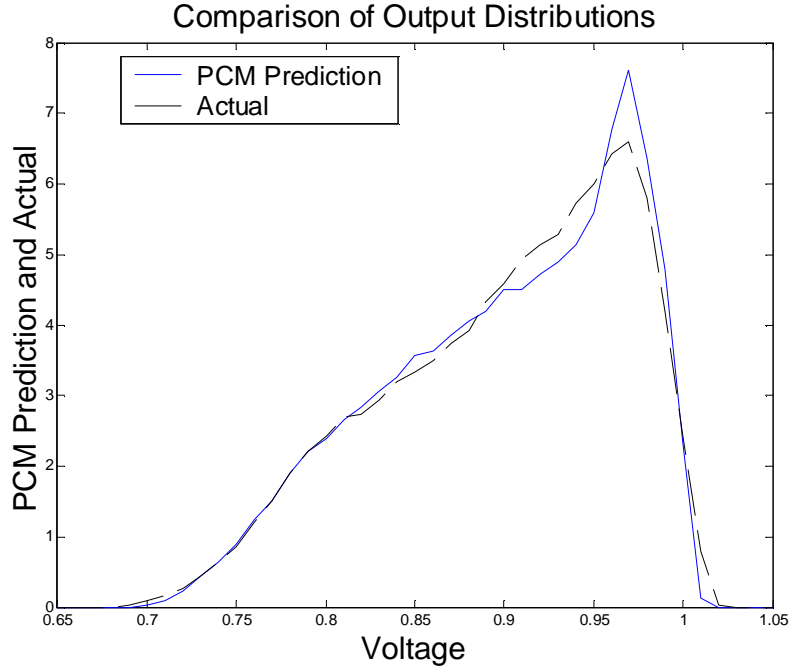
PCM prediction





**Figure 3.3:** Output plots. The mapping between the two input parameters and the voltage output that is predicted by PCM is compared with the actual mapping, found through exhaustive simulation. Each three-dimensional mapping is shown from two viewpoints to better illustrate their comparison.





**Figure 3.4:** The output distribution (i.e., the distribution of the voltage at bus 4) computed from the PCM-based mapping is compared with the actual output distribution. For this simple example, PCM characterizes both the input-output mapping and the output distribution well. Our solution highlights that the mapping between input and output over the domain of the uncertain loads is non-linear, especially because the correlation between the two loads makes heavy loading conditions frequent. PCM is able to capture this non-linearity, while (in this example) requiring only nine carefully-chosen simulation points to develop a good quadratic mapping. This ability to capture non-linear mappings using only a small number of simulations suggests that PCM holds promise as a PLF algorithm. We note that the PCM prediction, which requires only nine simulation points, is essentially indistinguishable from the mapping generated through brute-force simulation, which we construct using 400 simulations. Thus, our example highlights the significant computational savings that can be obtained through use of PCM. Finally, we numerically determine the distribution for the output variable and compare it to the actual output distribution in Figure 3.4. We note that PCM is also advantageous in this example,

in that we could allow uncertainties with arbitrary joint distributions on the input parameters.

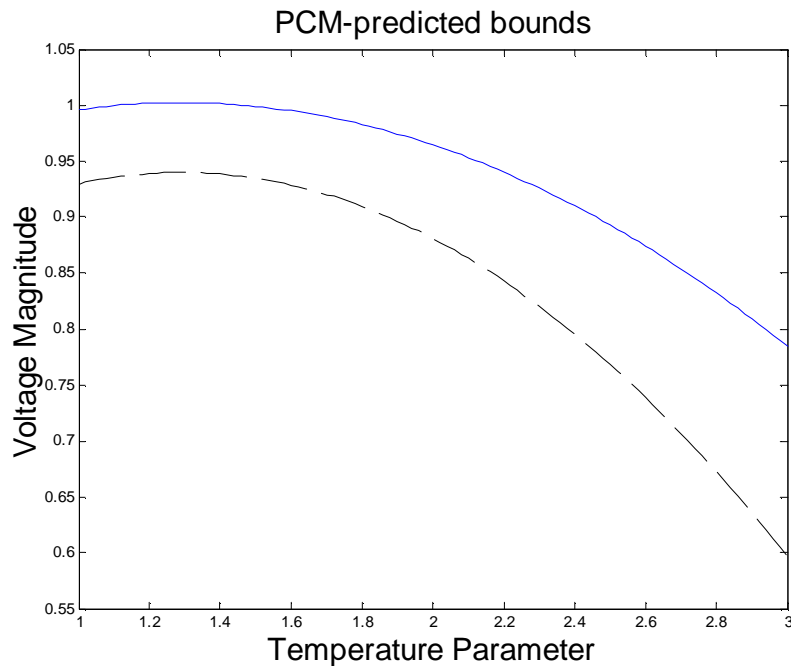
### 3.3 MESHING PCM WITH BOUNDARY LOAD FLOW

One difficulty in applying PCM is that the number of required PCM points typically grows exponentially in the number of uncertain parameters. When the number of uncertain parameters becomes large, we note that meshing PCM with a boundary load flow algorithm can provide a tractable solution. In particular, we can select PCM points for a few significant or important uncertain parameters; for each PCM point, we can apply a boundary load flow algorithm with respect to the other uncertain parameters, to find the largest and smallest possible output. Using these extremal outputs, we can develop a pair of mappings from the significant inputs to the output using PCM, which serve as bounds on the actual mapping. Such a meshed algorithm is best illustrated with an example. A plausible alternate description for the load scaling parameters in Figure 2 is that these parameters have a strong dependence on a single uncertain input parameter (e.g., temperature) with small, independent deviations from this predicted dependence.

For instance, the two parameters could have the form  $x = T + \varepsilon_1$  and  $y = T + \varepsilon_2$ , where  $T$  is a significant random parameter, and  $\varepsilon_1$  and  $\varepsilon_2$  are small, independent random parameters. While we could apply three-dimensional PCM to such a system, a less computationally intensive approach is the following. We can choose PCM points as if we are applying one-dimensional PCM, in which the uncertain parameter is  $T$ ; for each of these PCM points, we can compute the extremal output values over the domain of  $\varepsilon_1$  and

$\mathcal{E}_2$  (see [7] for an efficient means for doing so). We can then develop one-dimensional polynomial representations for both sets of extrema. An example of such “boundary mappings” is shown in Figure 3.5

As far as we know, PCM is the only non-Monte Carlo PLF method that is applicable when inputs are correlated according to arbitrary joint distributions. There is some literature on PLF when the inputs are jointly Gaussian, but we have not seen PLF algorithms for more general correlations among inputs.

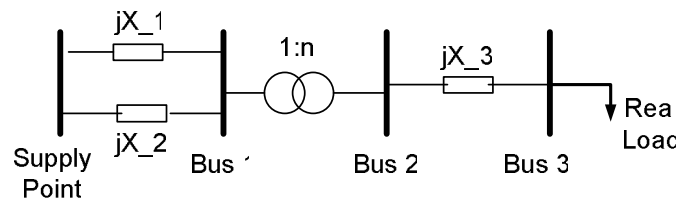


**Figure 3.5:** Boundary mappings. PCM is meshed with a boundary load flow algorithm to find bounds on the mapping between a significant uncertain parameter and the output voltage. The solid line is the upper bound and the dashed line is the lower bound.

### Example: Dynamic simulation of a disturbance

In the context of power systems, PCM was originally advanced as a tool for evaluating uncertainties in time-step simulations of transient dynamics ([1] and [2]). PCM is potentially valuable for evaluation of uncertainties in transients, because it can reduce the number of simulations (which are very often computationally intensive) required for uncertainty analysis. PCM also has the advantage that it can be implemented without significant modification of the time-step simulation programs for transients, since it only requires measurement of output values for various inputs.

Here, we apply two-dimensional PCM to characterize a small power system's transient response to a disturbance. The example that we use is drawn from [6], where it is also used to illustrate characterization of transient-simulation uncertainties, using trajectory-sensitivity methods. Our explorations of this example illustrate how PCM compares with, and complements, the trajectory-sensitivity based methods.



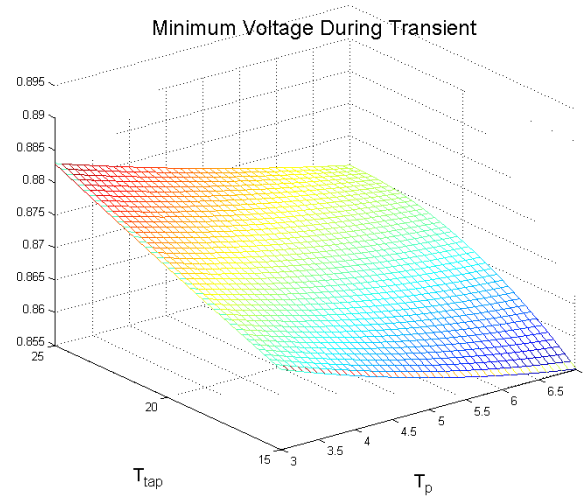
**Figure 3.6:** Dynamic simulation of disturbance example.

We consider the response of this power system to a disturbance, in particular tripping of the line with admittance  $X_1$ . The uncertain parameters in this example are the load recovery time constant and the tap-changing interval of the transformer.

The small system shown in Figure 3.6 is disturbed through tripping of one of the lines between the supply point and bus 1. We consider the transient response of the voltage magnitude at bus 3. This transient response is modulated by the recovery dynamics of the load, as well as the logic of the tap-changing transformer. It is in the parameters of these recovery dynamics that we assume some uncertainty (in accordance with [6]). In particular, we assume that the load time constant  $T_p$  and the interval between tap changes  $T_{tap}$  are uniformly and independently distributed, over the intervals [3, 7] and [15, 25], respectively.

We apply PCM to characterize the mapping between the inputs  $T_p$  and  $T_{tap}$  and an output of interest, which we choose to be the minimum voltage on bus 3 during the duration of the simulation. We find that a second-order generalized polynomial model is sufficient to specify the mapping (Figure 10). Thus, with only nine simulations, we are able to extract the mapping between the inputs and the output, and further to expose that this mapping is not linear. A compelling feature of PCM is that, using these nine simulations, we can in fact characterize many different output features (e.g., the output voltage at specific times, or various flows in the power network). We note that our analysis compares favorably with the trajectory sensitivity analysis in that we simulate the actual power system rather than a linear approximation thereof. We caution, however, that each simulation of the actual power system may be very expensive compared to a trajectory sensitivity-based simulation; it is only because so few points are required for PCM that our analysis is feasible. Finally, we mention that one further possible application of PCM to power system dynamic simulations is to identify whether

linear relationships between input and output variables hold, and hence to evaluate whether trajectory sensitivity analyses can be used.



**Figure 3.7:** PCM generated mapping. The PCM-generated mapping between two uncertain parameters and the minimum voltage reached by bus 3 during a transient simulation is shown

### 3.4 ORDER SELECTION ALGORITHM

In [2], the authors mention the necessity for a good order selection algorithm for practical applications of PCM. A good order selection algorithm can prove to be cost effective as a new set of simulations is required for each order of PCM polynomial selected.

We observed from our studies on PCM that the order selection can be done with mere visual inspection in certain cases. Such cases usually involve curves with multiple extrema. But in the case of curves with none or a single extreme we find the need for a proper order selection algorithm.

Before getting into the order selection algorithm we find it worthwhile to define the term **Kullback-Leibler Distance**.

The KL distance gives the distance between two PDFs in our case the KL distance can be used to compare the distance between successive PDFs (Output Distributions), and then we can go ahead and select the PDF when the KL distance becomes sufficiently small.

The **Kullback-Leibler** [13] distance is a measure of the difference between two probability density functions P and Q is given by.

$$D(P \parallel Q) = \int P(x) \text{Log} \left( \frac{P(x)}{Q(x)} \right) \quad (3.2)$$

The above integral is finite if and only if P is contained by Q.

#### **THE ALGORITHM:**

Our studies suggest the following heuristic order selection algorithm for one dimensional PCM (applicable to either case mentioned above) followed by the justification of its relevance to higher dimensional PCM.

1. We apply PCM of successive orders (beginning with first-order PCM), until visual inspection suggests that the predicted mapping has not changed between two successive applications.

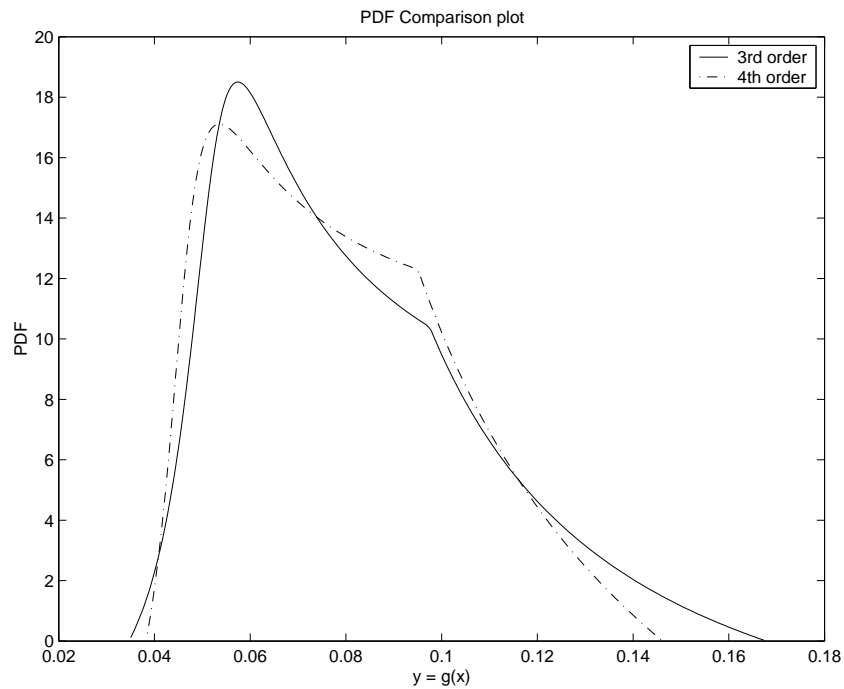
2. If the mapping predicted by the second-highest-order PCM applied in step 1 has at least two extrema, the visually-determined PCM fit is in our experience the proper one. (When the mapping has several extrema, we find that the PCM fit converges dramatically to the correct mapping beyond a certain order, so that visual inspection is sufficient to identify the proper fit. Order-selection is illustrated for a mapping with three extrema in figure 3.9.
  
3. If the second-to-last PCM prediction from the first step has fewer than two extrema, we require an analytical comparison measure to determine whether or not a sufficient order has been chosen. In particular, we numerically compute the output distribution using the mapping of each order. We then compute the **Kullback-Leibler (KL) distance** between successive pairs of distribution (see Table 3.1); if the KL distance between the highest two-order PCM output distributions is sufficiently small (i.e., drastically smaller than the KL distances between lower-order fits), then sufficiently high-order PCM has been used. Otherwise, a higher-order PCM algorithm should be applied, until a sufficiently small KL distance is obtained. We note that, if we desire a completely automatic algorithm for order-selection, we can use comparisons of KL distances regardless of the number of extrema.

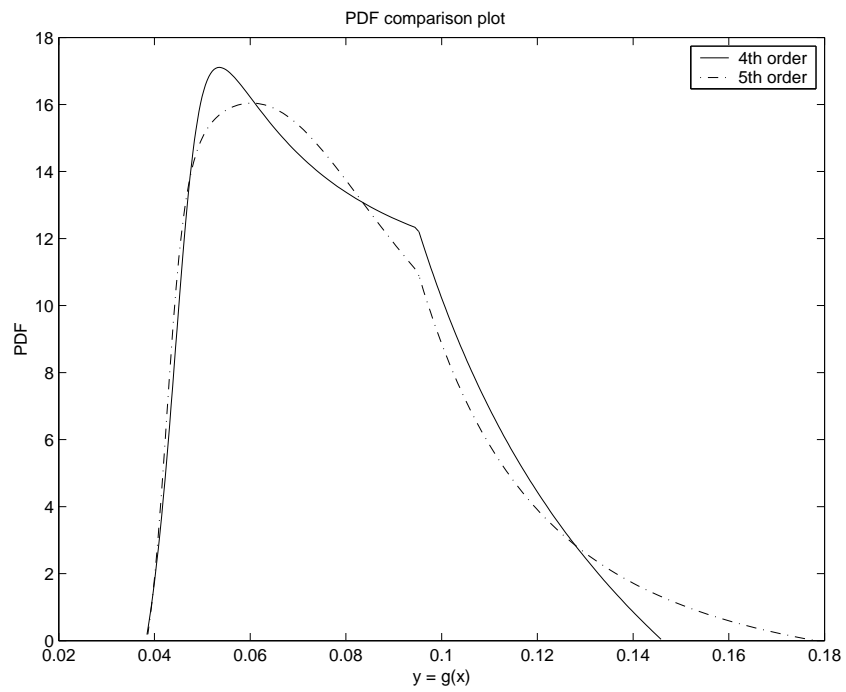
We applied the order selection to a series RC circuit example from [2], and the results are presented below. The results indicate that the appropriate order is 5.

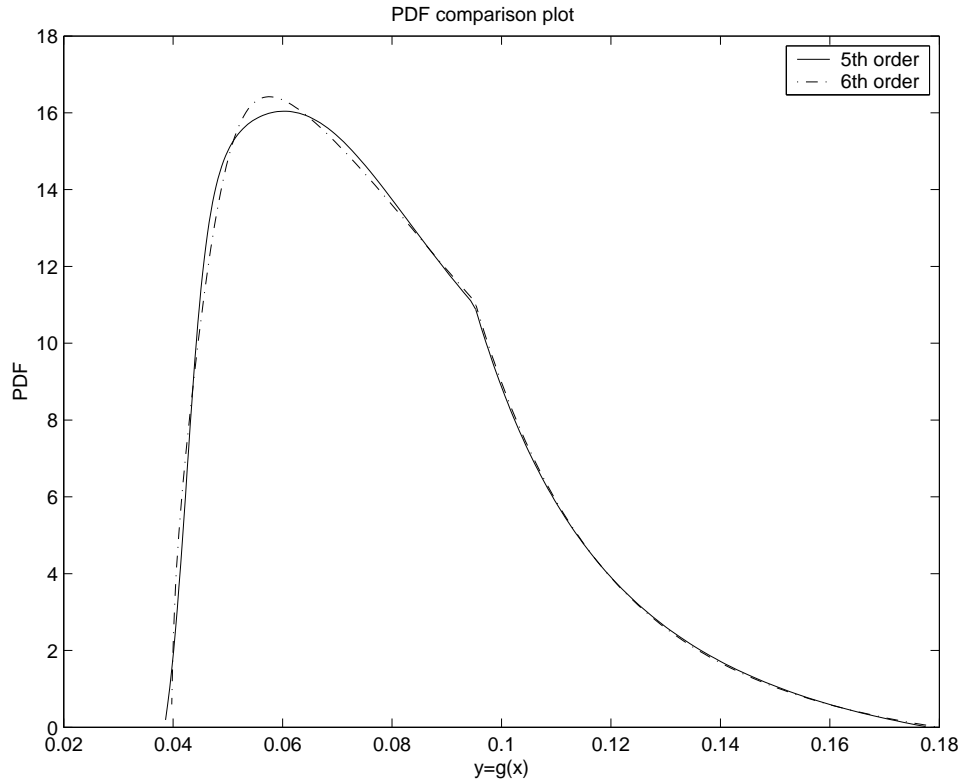


PDF comparison	KL distance
PCM 2 <sup>nd</sup> Vs. PCM 3 <sup>rd</sup>	0.1332
PCM 3 <sup>rd</sup> Vs. PCM 4 <sup>th</sup>	0.1134
PCM 4 <sup>th</sup> Vs. PCM 5 <sup>th</sup>	0.0977
PCM 5 <sup>th</sup> Vs. PCM 6 <sup>th</sup>	0.0033

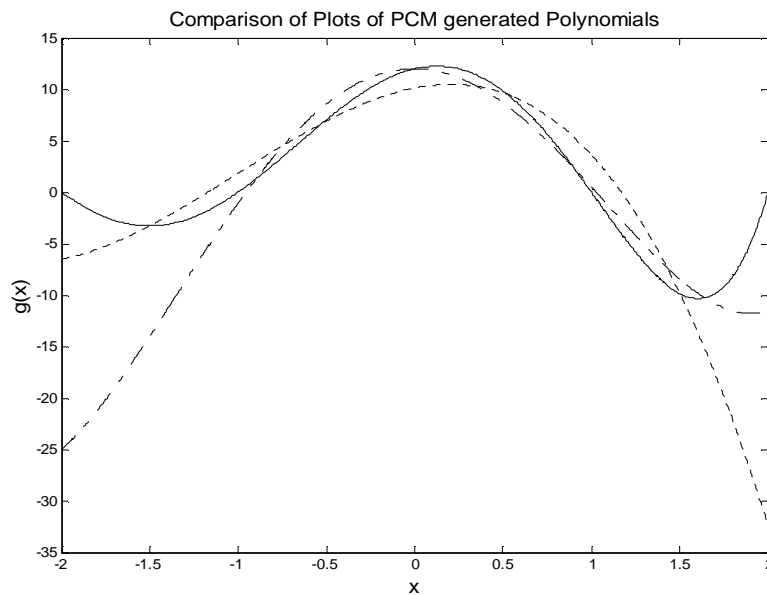
**Table 3.1:** KL distance comparison







**Figure 3.8:** Comparison of the output distribution plots of PCM generated polynomials of successive order.

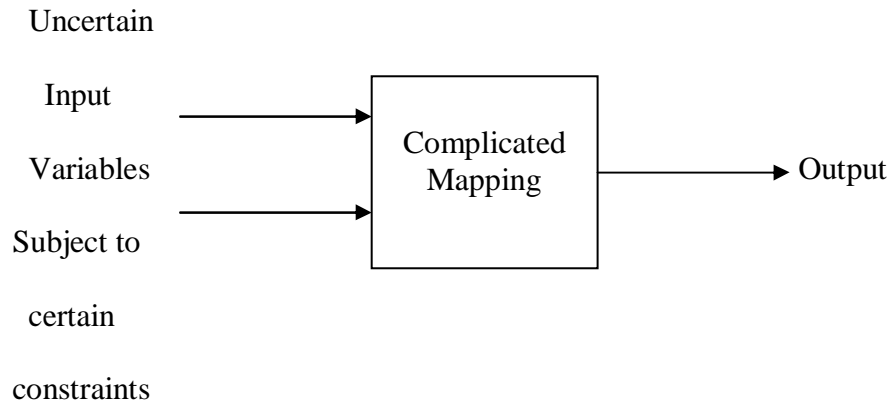


**Figure 3.9:** Comparison of plots of PCM generated polynomials for the case where the output function has two extrema. The appropriate order can be identified by visual inspection. In this case the 5<sup>th</sup> and higher order polynomials have two extrema and are quite different in terms of shape from the lower order polynomials.

Our studies show that this algorithm is very effective for single dimensional PCM. In our two-dimensional PCM algorithm we generate the PCM collocation points for any one of the uncertain parameters and on the basis of these points, generate the collocation points for the other uncertain parameter in the system under examination. The order selection algorithm can be applied when the PCM collocation points for the first parameter are generated. Once the appropriate order is selected it can be labeled as “the order of the system” and for the second uncertain parameter, we can generate collocation points based on “the order of the system”.

## CHAPTER 4

### OPTIMIZATION



**Figure 4.1:** Pictorial depiction of the optimization problem

An optimization problem is concerned with finding the minimum or maximum of a function, with respect to its arguments, which are in many cases constrained to a bounded set. There is a wide literature on optimization. We request that the readers refer to, e.g. [18], for basic notions in optimization.

In this section, we consider the problem of optimizing a function over the domain of an uncertain parameter, in the case where function evaluations are time-consuming expensive. In particular, we discuss in an exploratory manner, the possibility of applying PCM for solving optimization problems.

In general, if the function under study is a black box and if the uncertain input (parameter) variables are continuous and bounded, PCM can be used to approximate the

maximum or minimum of the function over the interval. In particular, PCM generates a polynomial approximation for the black box function, which can then be optimized.

Broadly, there are two viewpoints on using PCM to find a maximum/minimum

1. We can address the standard optimization problem of minimizing/maximizing a function over a bounded domain. In this case, it is reasonable for us to assume a uniform distribution for the uncertain parameter in generating the PCM fit.
2. We can view the parameters over which the optimization is done as being uncertain, and find the maximum/minimum in a manner that reflects the distribution of these parameters. That is, by using the distribution in PCM, we can search more carefully for the optimum over high-probability parameter values.

We can come up with a different PCM mapping between the input variables and the output of interest in either case. PCM generates a polynomial mapping from which the minimum or maximum value of the function can be found directly.

An optimization problem in the power systems domain could be maximizing an output voltage at a particular bus over the domain of the input parameters (say, loads).

**Example:**

Let us explore this possible application of PCM through an example.

The example used is the 5 Bus loadflow example from [10]. Two loads  $\alpha_1$  and  $\alpha_2$  are uncertain, and the output of interest is the voltage at bus 4. The constraint here is that the sum of the two loads ( $\alpha_1, \alpha_2$ ) should be equal to 1.75 and  $\alpha_1$  is distributed in the range (0, 1). The optimization problem in this case would be to find the distribution for the loads that maximizes the voltage at bus 4.

**Distribution 1:**

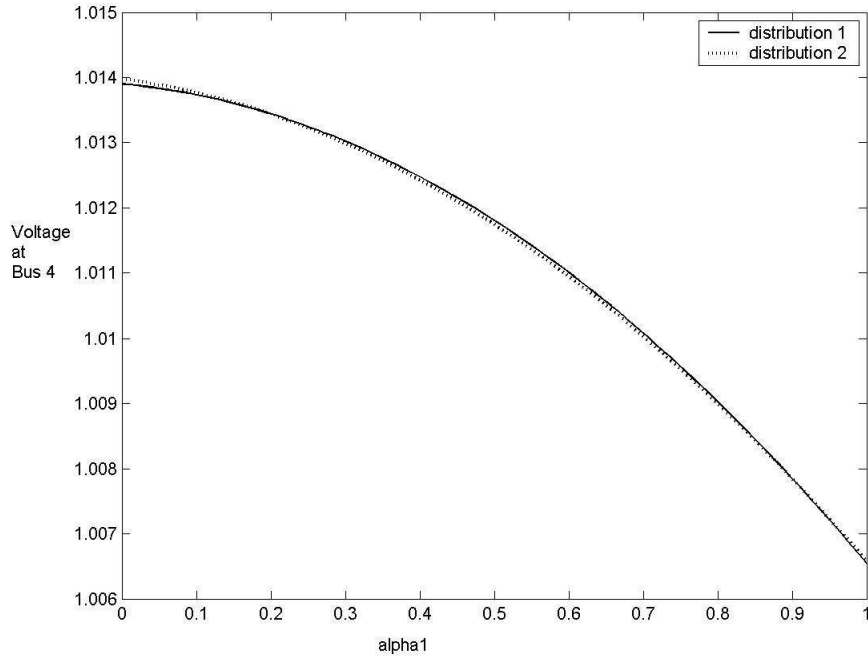
The PCM mapping was generated by first assuming uniform distribution  $\alpha_1$  for in the range (0, 1) and  $\alpha_2 = 1.75 - \alpha_1$ .

**Distribution 2:**

A different distribution was assumed for  $\alpha_1$ , viz.  $f(\alpha_1) = \frac{(3\alpha_1 + 1/2)}{2}$ ,

$\alpha_2 = 1.75 - \alpha_1$ , and the corresponding PCM polynomial was generated.

The plots of the two polynomials are depicted in the figure 4.2



**Figure 4.2:** Comparison of PCM polynomials generated using the two different input distributions.

The maximum value of Voltage for distribution-1 is 1.0139 at  $\alpha_1 = 0$

The maximum value of Voltage for the distribution-2 is 1.0140 at  $\alpha_1 = 0$

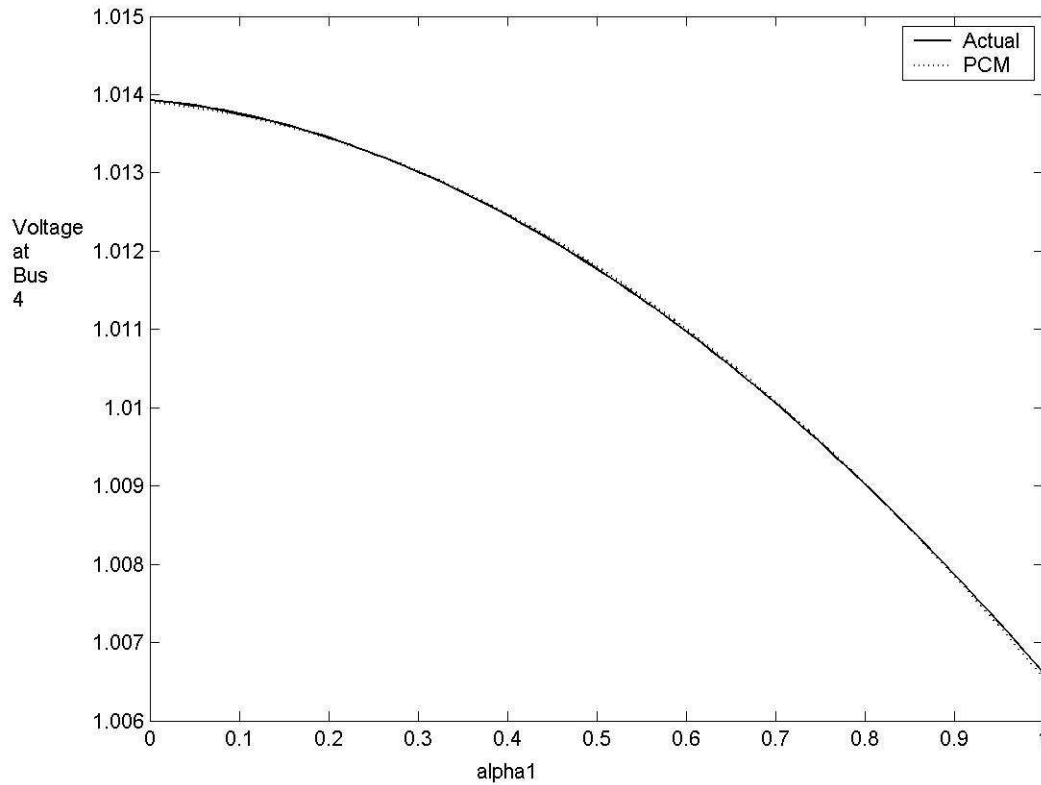
From this we can infer that a higher maximum voltage can be achieved if the loads are distributed as in distribution-2. The constraint in this problem is that the sum of the loads must equal 1.75 we could also alter the constraints and generate the PCM polynomial for the problem. In either case, we claim that PCM can be a handy tool as it is economic in terms of simulations and also once the PCM polynomial is generated the optimization problem reduces to the task of finding out the maximum/minimum of the PCM polynomial generated.



## ACCURACY OF PCM

How accurate is PCM in capturing the maximum/minimum value of the function?

The following example illustrates the accuracy of PCM. For distribution-1 of the previous example the PCM second order fit is compared with the actual output fit generated via exhaustive simulation.



**Figure 4.3:** Comparison of the actual function generated by exhaustive simulation with the PCM generated polynomial using the 1<sup>st</sup> distribution.

The two plots cannot be differentiated; this shows that PCM is quite accurate in modeling the function and eventually the minimum/maximum of the function. The maximum value is captured accurately up to 4 decimal places. It is the same for both the fits viz., 1.0139.

#### 4.1 ADVANTAGE OF USING PCM

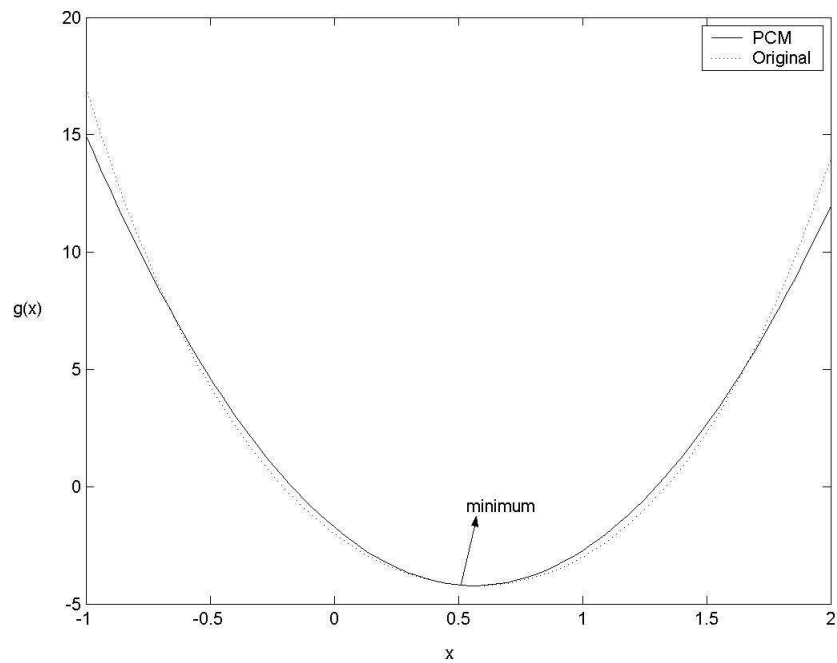
1. The main advantage again is economy. PCM requires a small number of simulations for generating the polynomial mapping.
2. PCM specifies a low-order mapping that approximates a much higher-order relationship. For instance if the original relationship is a quartic, a quadratic PCM polynomial provides good approximation in many cases. The example depicted by the figure below illustrates this fact

##### Example:

The parameters are  $x_1$  and  $x_2$ . The distribution and constraints are as follows,

$-1 < x_1 < 2$ , Uniform distribution.

$x_2 = 2 - x_1$ . The output function is  $g(x_1, x_2) = x_1^4 - 2x_1x_2^2 + x_2 + x_1 - 4$



**Figure 4.4:** The minimum value of a quartic function captured by a PCM quadratic polynomial.

The PCM mapping shown in figure 4.4 for this relationship is a quadratic whereas the actual relationship is a quartic, yet it captures the minimum accurately. In case there is more than one minimum/maximum the lower order PCM mapping could possibly capture one of them.

The capability of PCM to solve optimization problems adds a new dimension to the algorithm. We have shown that PCM could be effectively used for maximizing voltage in power systems. This is just a rudimentary attempt at using PCM to solve optimization problems, and to tout PCM as an optimization algorithm in general is not appropriate. However, the results obtained so far are promising and we hope that future work in PCM will be concentrated in this area.

## 4.2 COMPARING PCM WITH A TRADITIONAL MINIMIZATION TECHNIQUE

It is interesting to see how PCM fares when pitted against traditional function minimization techniques such as the steepest descent, gradient descent or the Newton-Raphson method.

Among the above mentioned techniques, Newton-Raphson has the fastest convergence rate. Newton-Raphson is an iterative process for minimizing a function with respect to one or more variables. The Newton-Raphson formula for minimizing a single variable (one dimensional) function  $f(x)$  is

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

The iteration is started by guessing an initial value  $x_0$ .

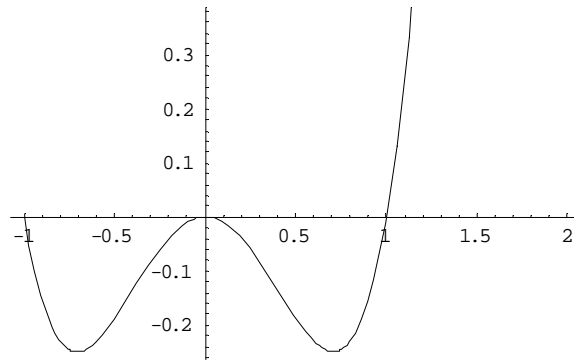
The method iteratively tries to locate the minimum of the function, and the accuracy of the technique increases with the number of iterations. Of course, the initial guess must be intelligent otherwise this technique may not converge.

When comparing an iterative minimization technique like Newton-Raphson with PCM, we must first identify a yardstick for the comparison. Comparing the number of iterations that Newton-Raphson takes to converge to the minimum value of the function with the Order of PCM that produces the polynomial with the correct minimum value appears sensible. For each Newton-Raphson iteration, we need to calculate the value of the first

and second derivative of the function under study at the current estimate for  $x_{n+1}$ . Thus we need to simulate the function iteratively; as a matter of fact we would be performing more than one simulation per iteration as we need to calculate both the first and second derivative of the function each time.

### EXAMPLE

Minimize  $g(x) = x^2(x^2 - 1), -1 < x < 2$



The above function can be minimized analytically. The purpose of choosing such an example is that it makes the task of comparison easier. We can compare the actual minima calculated analytically with those computed using Newton-Raphson and PCM.

The function has a local minimum at  $x = 0$  and global minima at  $x = \left\{ \frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\}$ . The minimum value of the function is -0.25.

We applied Newton-Raphson to minimize this function and also used PCM for the same purpose. Uniform distribution was assumed for the variable  $x$ . The results are presented in table 4.1.

Newton-Raphson Iterations	PCM order
<p>1<sup>st</sup> Iteration</p> $x_{\min} = 0.5, g(x_{\min}) = -0.1875$	<p>1<sup>st</sup> Order</p> $x_{\min} = -1, g^*_{\min} = - - 0.750011$
<p>2<sup>nd</sup> Iteration</p> $x_{\min} = 1, g(x_{\min}) = 0$	<p>2<sup>nd</sup> Order</p> $x_{\min} = -1.07427, g^*_{\min} = -0.578827$
<p>3<sup>rd</sup> Iteration</p> $x_{\min} = 0.8, g(x_{\min}) = -0.2304$	<p>3<sup>rd</sup> Order</p> $x_{\min} = \left\{ 0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\}, g^*_{\min} = \{0, -0.25, -0.25\}$
<p>4<sup>th</sup> Iteration</p> $x_{\min} = 0.721127, g(x_{\min}) = -0.249599$	<p>4<sup>th</sup> Order</p> $x_{\min} = \left\{ 0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\}, g^*_{\min} = \{0, -0.25, -0.25\}$
<p>5<sup>h</sup> Iteration</p> $x_{\min} = 0.707505, g(x_{\min}) = - - 0.24999968$	
<p>6<sup>h</sup> Iteration</p> $x_{\min} = 0.707107, g(x_{\min}) = -0.25$	

**Table 4.1:** Comparison of PCM with Newton-Raphson minimization

## OBSERVATIONS

Newton-Raphson took 6 iterations to converge to the minimum this is partly due to the initial guess being slightly off the hook. If we had taken  $x_0 = 1$ , we would have got a convergence in 5 iterations.

Although the function under study is a quartic, the PCM cubic polynomial captures the minimum accurately. We could have stopped with 5 iterations for Newton-Raphson and at the 3<sup>rd</sup> order for PCM, the extra iteration (and order) are just to check for convergence.

Assuming we had taken  $x_0 = 1$  and do not consider the extra iteration to check for convergence, Newton-Raphson estimates the minimum in 4 iterations. In the same vein, not considering the extra PCM order, the 3<sup>rd</sup> order PCM polynomial captures the minimum of the function and it took 4 simulations for generating the 3<sup>rd</sup> order polynomial.

Hence, in this example PCM performs as well as Newton-Raphson, but it may not always be the case. Our purpose here was to illustrate that PCM could be used as a tool for minimization and not to claim or try to prove that it works better than existing minimization techniques. PCM has a long way to go in this aspect and this chapter is just a prefatory to the study of using PCM to solve optimization problems.

## **CHAPTER 5**

### **INFORMATION-THEORETIC APPROACH FOR USING FEWER**

#### **PARAMETERS IN PCM**

The advantage of PCM over traditional Monte Carlo simulation techniques is that PCM requires very few simulations to identify the mapping between the uncertain input(s) and the output of interest. Although PCM is economic, PCM too requires an exponential amount of simulations as the number of inputs increases. For instance if  $k$  is the number of system uncertainties, it would take  $(n+1)^k$  simulations to generate a polynomial of order  $n$ . Though this number of simulations is typically small compared to the number needed for traditional Monte Carlo techniques, it is necessary to come up with variable reduction techniques to make the process of modeling the system uncertainties less cumbersome.

When the multiple PCM inputs are strongly correlated, the input variables potentially carry a lot of redundant information. In such cases it may be possible to model the mapping using only a subset of the input variables or a lower order basis for them. In order to do so, some mechanism for measuring dependencies between the input variables is required. Some interesting information theoretic concepts, including Entropy and Mutual Information, can be used to measure dependencies between the variables. In particular, we use Mutual Information as a good measure of dependency between jointly distributed random variables using which variable reduction in PCM can be achieved. The remainder of this chapter discusses this information-theoretic approach for reducing the number of input variables in PCM.



## 5.1 DEPENDENCY MEASUREMENT

Before describing its application to PCM we review certain Information theoretic concepts useful for the study.

*Mutual Information* [13] is an information-theoretic concept, which can be used as an indicator for the degree of dependency between jointly distributed random variables. Another useful measure is the correlation coefficient, which is a degree of correlation between two random variables. Correlation is the degree to which two or more quantities are linearly associated, [24].

The **Differential Entropy  $H(X)$**  [13] of a continuous random variable  $X$  with a density  $f(x)$  is defined as

$$H(X) = -\int_S f(x) \log_2 f(x) dx \quad (5.1)$$

where,  $S$  is the support set of the random variable. Support set is the set of  $x$  for which  $f(x) > 0$  is called the support set of  $x$ .

*Entropy* [13] is a measure of randomness of a random variable, and in the discrete domain it represents the shortest description length (in bits) for the variable.

*Differential Entropy* [13] is also related to the shortest description length. One caveat here is that we can get negative values for differential entropy. Hence an appropriate measure for description length is the volume of the support set of the random variable given by  $2^{h(X)}$ , which is obviously non-negative.

The **Joint Entropy  $H(\mathbf{X}; \mathbf{Y})$**  [13] of jointly distributed continuous random variables  $x$  and  $y$  with joint density  $f(x, y)$  is defined as

$$H(X;Y) = -\int f(x, y) \log_2 f(x, y) dx dy \quad (5.2)$$

The Joint Entropy again is a measure of randomness or description length the difference here is that we are considering a vector of random variables instead of a single random variable.

### **Mutual Information:**

The **Mutual Information  $I(\mathbf{X}; \mathbf{Y})$**  [13] between two jointly distributed continuous random variables  $x$  and  $y$  with joint density  $f(x, y)$  is defined as

$$I(X;Y) = \int f(x, y) \log_2 \left( \frac{f(x, y)}{f(x)f(y)} \right) dx dy \quad (5.3)$$

The Mutual Information between two jointly distributed random variables is the amount of information one random variable contains about another. In a sense, it is the reduction in uncertainty of the random variable  $X$  due to the knowledge of  $Y$  and vice versa. It is an estimate of the strength of association between jointly distributed random variables.

### Correlation Coefficient:

The **Correlation Coefficient** [13] is a numeric measure of the strength of linear relationship between two random variables

$$\rho(x, y) = \frac{\text{cov}(x, y)}{\sqrt{\sigma^2(x)\sigma^2(y)}} \quad (5.4)$$

Where  $\text{cov}(x, y)$  is the covariance defined as

$$\text{cov}(x, y) = E[(x - E(x))(y - E(y))] = 0, \quad \text{if } x \text{ and } y \text{ are independent}$$

$E$  = Expected value

$\sigma^2$  = Variance

The correlation coefficient lies between -1 and 1. It is -1 if  $x$  and  $y$  are perfectly anti correlated and 1 if  $x$  and  $y$  are perfectly correlated.

## 5.2 REDUCING THE NUMBER OF INPUT VARIABLES

Reducing or filtering input random variables is the process of eliminating certain variables considered containing redundant information, and using the remnant variables for generating the PCM mapping between the inputs and the output of interest.

For instance, if  $x$  and  $y$  are the input random variables, the 2-D PCM fit for the system will be of the form

$$g(x, y) = \sum_{i=0}^n \sum_{j=0}^n a_{ij} x^i y^j \quad (5.5)$$

Assuming that  $y$  has redundant information and can be eliminated, the process of developing the PCM mapping for the system degenerates to a 1-D PCM problem with

only one input random variable viz.,  $x$ . The PCM fit after variable reduction will be of the form

$$g^*(x) = \sum_{i=0}^n a_i x^i \quad (5.6)$$

Though we discard the random variable  $y$ , it must be noted that we need both the random variables for running simulations of the black box system under analysis. We suggest the usage of the conditional mean for the redundant random variable instead of its PCM values.

The caveat to be kept in mind is that the reduction shouldn't result in loss information, in other words  $g^*(x)$  should approximate  $g(x, y)$ . The moments of both the polynomials and their output distribution plots are good comparative measures that can be used to check the accuracy of the reduction process.

The combination of Mutual Information and Correlation Coefficient values can be used as a tool for deciding when to reduce input random variables.

We distinguish between two cases in which we can eliminate input random variables, we to call them Case I and Case II.

### **Case I**

The variables have high mutual information between them. Experimental results suggest that a reasonable cutoff value for the above measure is 3.5 and greater. The variables are also strongly correlated (or anti-correlated) i.e., their correlation coefficient value is close to 1 or  $-1$ . Experimental evidence suggests that a good cutoff value is 0.9 and greater.

### **Case II**

The variables don't have a very high mutual information value, but the entropy of one of the input variables is very small compared to the entropy of the other random variable and also to the mutual information. The correlation coefficient isn't significant in this case but the variance of the individual random variables can be used in lieu, for comparison. In the following section the two cases are described via examples.

Note, in either case, it is not necessary that for the reduced fit to be of the same order as the original. We recommend going using a higher order for the reduced fit, for the sake of accuracy.

## **EXAMPLES**

### **Case I**

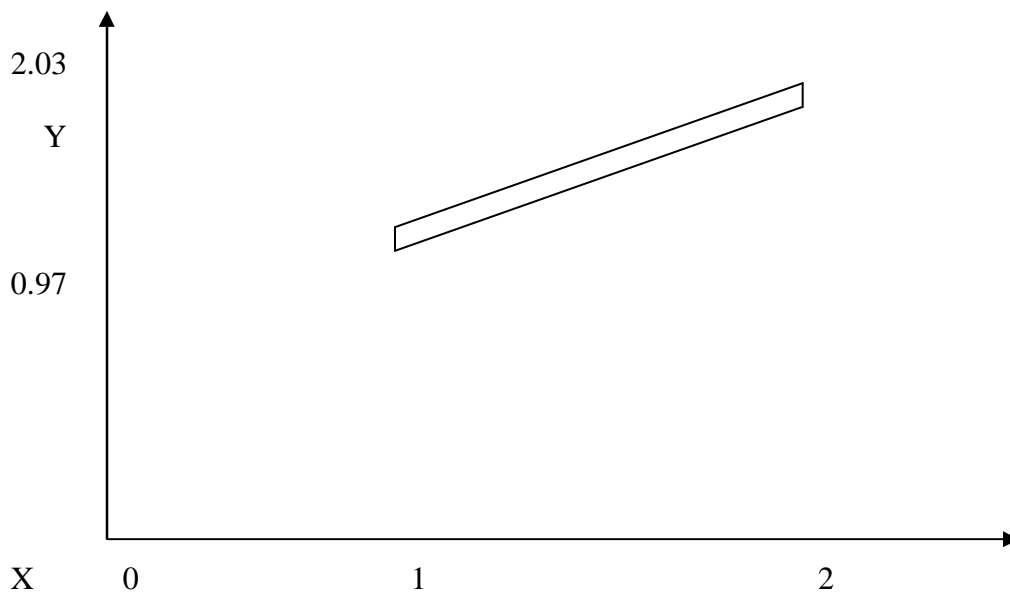
One of the random variables is uniformly distributed and the distribution for the second random variable exhibits a very strong dependence on the first. The motivation for selecting an example where the variables exhibit strong correlation is to show that the correlation coefficient is related to the notion of mutual information.

## DETAILS

$$1 < x < 2$$

$$x - 0.03 < y < x + 0.03$$

$$f(x, y) = (1/0.06)$$



**Figure 5.1:** Input distribution region.

The Mutual Information, Joint Entropy and Correlation Coefficient values are as follows

$$I(X;Y) = 4.24261$$

$$H(X;Y) = -0.246373$$

$$\rho(x, y) = 0.998205$$

We assume that the two random variables are two load scaling parameters chosen for the 5-Bus loadflow example from [10]. For details; please refer to earlier chapters in the thesis. As before, loads 4 and 5 are considered uncertain and the output of interest is the voltage at Bus number 4. First the 2-D PCM Quadratic polynomial mapping was generated for this example. The 1-D PCM Quadratic polynomial was then developed by using PCM points for only one of the input random variables viz.,  $x$ , while for the second variable  $y$ , the conditional mean at each  $x$  value was used. The 2-D and 1-D polynomials and their distributions are as follows

$$g(x, y) = 2.86 \times 10^{-13} x^2 y^2 - 9.22 \times 10^{-13} xy^2 - 0.0014x^2 y + 0.015xy + 6.07 \times 10^{-13} y^2 - 0.029y + 0.026x^2 - 0.082x + 1.09$$

$$E(g(x, y)) = 0.960256$$

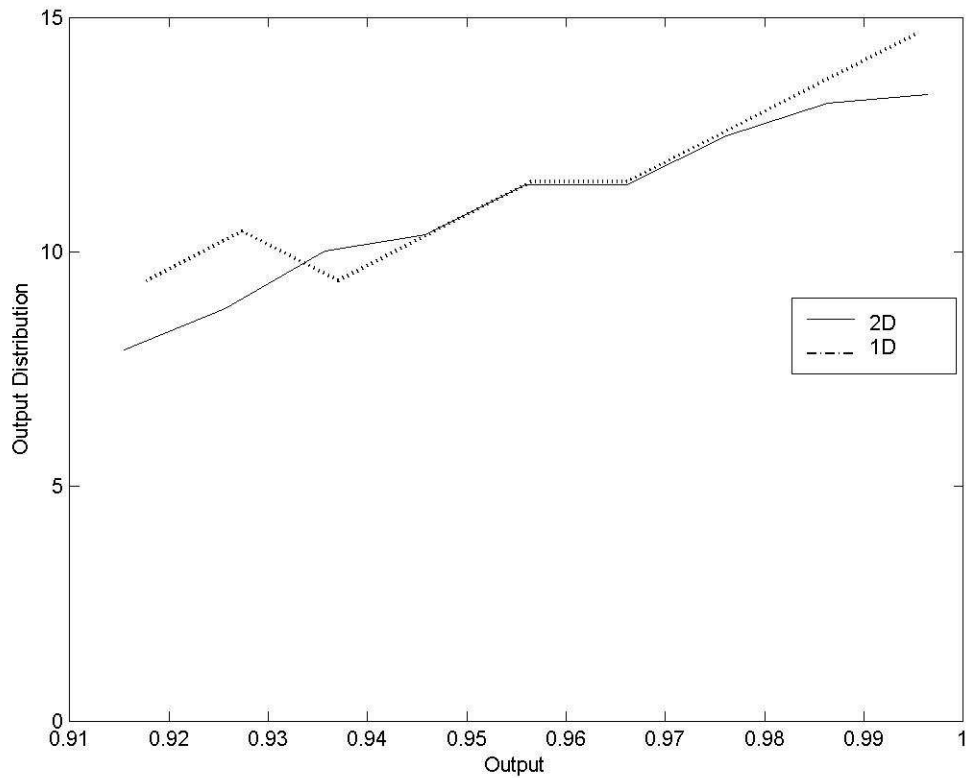
$$Variance = 0.000634529$$

$$g(x) = -0.0233x^2 - 0.017x + 1.04$$

$$E(g(x)) = 0.960256$$

$$Variance = 0.000633957$$

Thus, we see that the Expected Values and Variances agree strongly. Next, the output distribution of the 1-D and the 2-D PCM polynomials are compared in figure 5.2.



**Figure 5.2:** Comparison of output distributions, case I. The solid line represents the output distribution based on the 2D PCM approximation and the dotted line represents the output distribution based on the 1D PCM approximation.

Figure 5.2 depicts the plots for the output distribution corresponding to the 2-D and 1-D PCM. The Output distribution plots are quite similar, corroborating the statistical results presented before. The results show that when the two random variables have high mutual information value and are strongly correlated it is sufficient to use just one of the random variables for characterizing the input-output mapping Table 5.1 shows how the Mutual information increases as the size of the distribution is reduced in the y-direction.



Distribution	I(X; Y)
$x - 0.1 < y < x + 0.1$	2.31276
$x - 0.09 < y < x + 0.09$	2.40428
$x - 0.08 < y < x + 0.08$	2.52115
$x - 0.07 < y < x + 0.07$	2.67316
$x - 0.06 < y < x + 0.06$	2.87623
$x - 0.05 < y < x + 0.05$	3.15829
$x - 0.04 < y < x + 0.04$	3.57336
$x - 0.03 < y < x + 0.03$	4.24261
$x - 0.02 < y < x + 0.02$	5.51107
$x - 0.01 < y < x + 0.01$	8.97349

**Table 5.1:** Mutual Information values

It indicates that the correlation gets stronger, as one would expect, because the dependence of  $y$  on  $x$  becomes stronger. Mutual Information captures this phenomenon effectively.

## Case II:

### Example:

$$f(x, y) = \left(\frac{x}{3}\right) - \left(\frac{y}{9}\right), 1 < x < 3, 1 < y < 2$$

$$I(X;Y) = 0.000431909$$

$$H(X) = 0.615262$$

$$H(Y) = -0.00206016$$

$$\text{Variance}(x) = 0.283951$$

$$\text{Variance}(y) = 0.0829904$$

Again the same 5-Bus load flow example was used with the voltage at bus 4 as the output of interest. Both 2-D and 1-D PCM fits were generated in the same fashion as in Case-I.

The results are as follows

#### 2-D

$$g(x, y) = -0.0017x^2y^2 + 0.004xy^2 + 0.0022x^2y - 0.011xy - 0.011y^2 - 0.0001y - 0.007x^2 - 0.016x + 1.04$$

$$\text{Mean} = 0.922204$$

$$\text{Variance} = 0.0011$$

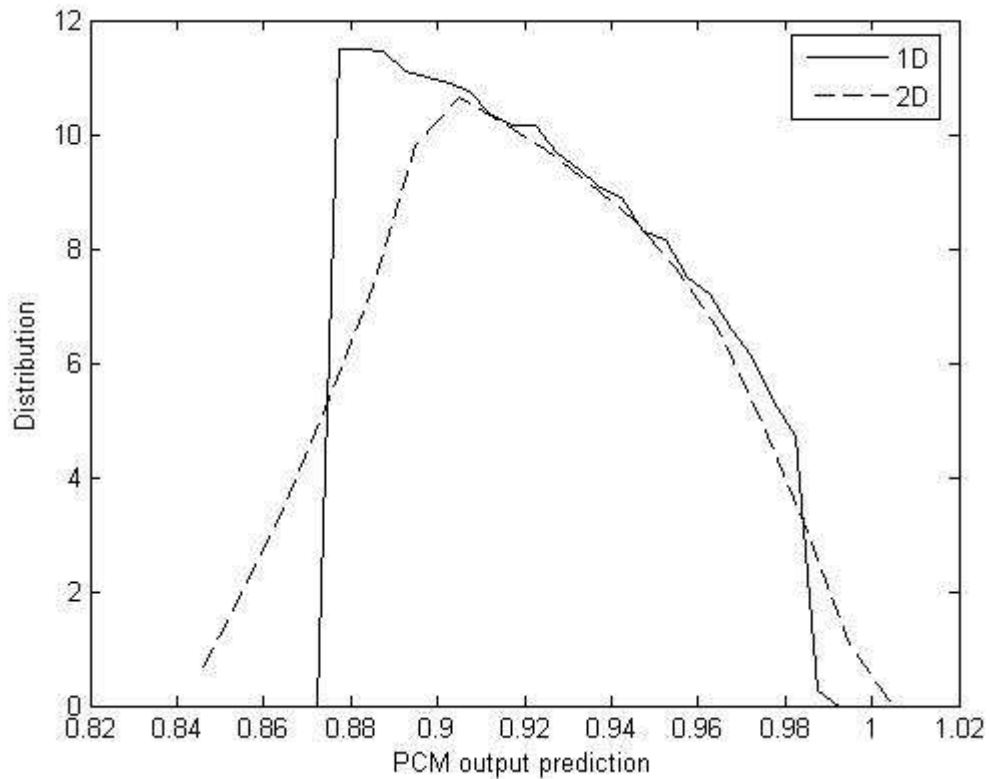
#### 1-D

$$g(x) = -0.00776x^2 - 0.0024x + 1.017$$

$$\text{Mean} = 0.923196$$

$$\text{Variance} = 0.00092$$

### Distribution Plots:



**Figure 5.3:** Comparison of output distributions, case II. The solid line represents the output distribution based on the 2D PCM approximation and the dotted line represents the output distribution based on the 1D PCM approximation.

The expected values, variances agree well and the distribution plots are similar. Thus the results support the claim that we can use just one of the random variables instead of two to model the input to output mapping of the system, when the mutual information and entropy values of the input variables are as described under the conditions for Case II.

### 5.3 JUSTIFICATION

In our development, we have suggested using Mutual Information as the primary criterion for eliminating redundant random variables, and have mentioned that the correlation coefficient can provide a second criterion. A brief comparison of the two criteria is valuable for identifying the advantages and limitations of each. In this section, we provide a conceptual comparison of the two.

Broadly, our motivation for invoking information theoretic concepts rather than only using the correlation coefficient is that less restrictive criteria for parameter reduction can be developed. Specifically, by eliminating parameters with high mutual information, we permit elimination of parameters that are nearly deterministically but non-linearly related. For instance, consider the following, which is a limiting case, in that one parameter is a deterministic function of the other:

Example:

Consider a system with a pair of uncertain inputs  $X_1$  and  $X_2$ ,  $X_2 = X_1^2$  and  $X_1$  is uniformly distributed between -1 and 1. Since  $X_2$  is a deterministic function of  $X_1$ , the mapping between  $X_1$  and the system output is a deterministic one. Hence, we can identify the mapping between  $X_1$  and the output using PCM, albeit perhaps with a higher-degree polynomial than if the output is expressed in terms of both  $X_1$  and  $X_2$ . Hence, our criterion should eliminate  $X_2$  (or alternatively  $X_1$ ) in this case. Since the conditional entropies of each variable given the other are arbitrarily negative for this pair

of random variables, the information-theoretic condition indeed indicates that one of the parameters can be eliminated. However the correlation coefficient for  $X_1$  and  $X_2$  is 0, and hence a correlation-based test would not indicate that a parameter could be eliminated.

To summarize, information-theoretic concepts allow us to eliminate parameters that are strongly-interdependent in non-linear ways, while correlation coefficients only allow us to identify linear dependencies. Since the applicability of PCM is based on whether or not the mapping from the parameters to the output is deterministic rather than on its linearity the less restrictive information-theoretic condition should be the primary one. It is worth noting that a high correlation coefficient yields a stronger result, in that it indicates not just the possibility for parameter reduction but the possibility for using a lower-order PCM fit of the same degree.

### **The Criterion: $I(\mathbf{X};\mathbf{Y})$ , $\text{var}(\mathbf{X})$ , or $I(\mathbf{X};\mathbf{Y})-H(\mathbf{X},\mathbf{Y})$**

In the above development, we have distinguished between two cases – one in which high mutual information permits elimination of variables, and another in which low spread (variance) of one of the variables permits its elimination. The reader may wonder why these measures cannot be combined into a single one (e.g., why mutual information by itself cannot be used to eliminate variables), and hence some further discussion of the criteria is needed. In fact, the underlying difference between these two cases brings up a

more general concern about what the proper criterion is, and suggests yet another measure for variable reduction.

Perhaps the best way to explain the distinction between the two cases is to note that the mutual information and correlation coefficient are unitless quantities, while entropies and variances have units. That is, simple scaling of the random input variables does not change their mutual information, but does change the variance and entropy of each variable. Thus, the mutual information (or correlation coefficient) identifies the reduction in uncertainty in one variable through knowledge of the other, but does not identify the actual randomness in these variables. Thus, when we use the mutual information-based criterion, we are considering the reduction in one variable's uncertainty due to knowledge of the other, in a *scale-free* way. In contrast, when we choose to eliminate random variables with small variances, we make the assumption that the two variables are defined on the same scale, also that variation in the output of interest over equally-sized domains of each variable are on the same order. Such an assumption is reasonable, for instance, in the power-flow example, in which the random parameters are scaling factors for loads of nearly the same magnitude and hence also have comparable impact on the output voltage.

More generally, when the absolute scaling of the system is well-understood, we note that the mutual information-based criterion can be modified to take into account the absolute statistics of the inputs. One way of doing so is to use a measure such as  $I(X;Y)-H(X,Y)=-H(X/Y)-H(Y/X)$ . When this quantity is sufficiently positive, the relative entropy of  $X$

given  $Y$  and/or the relative entropy of  $Y$  given  $X$  are small, and reduction of one of the variables is possible. We note that this measure accounts for the absolute uncertainty in each variable conditioned on the other, rather than using only the change in uncertainty.

More work is needed to better delineate when each criterion should be used. We hope to complete this aspect of the study in the near future.

## **CHAPTER 6**

### **EXAMPLES**

#### **6.1 A LARGER EXAMPLE**

In this chapter we will apply several of the concepts discussed thus far to a larger electric power systems example. Although stylized examples have been provided in each chapter, a larger realistic example is necessary to demonstrate the applicability of the technique. Power systems loads are generally classified as industrial, commercial or residential based on the usage sector. Traditionally loads classified under the same category have interdependencies and if these interdependencies are strong enough, we can use the techniques discussed in this thesis for generating a reduced-order PCM polynomial for the system.

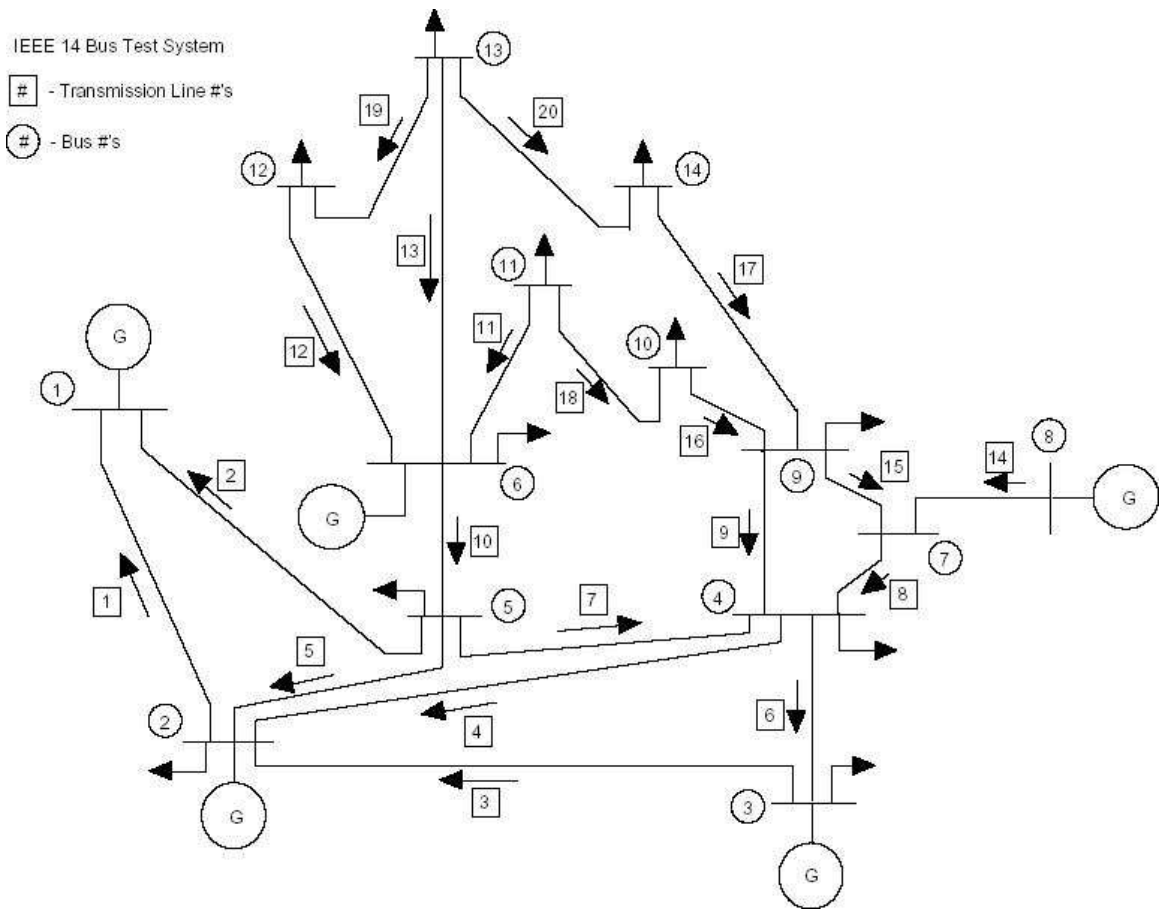
Figure 6.1 represents an IEEE 14 Bus Test System. The numbers inside the squares represent the transmission line numbers and the bus numbers are encircled. For our purpose, we have assumed that the loads at 6 of the buses uncertain, and we divide the uncertain loads into two categories viz. industrial and commercial. To be specific, Loads at bus #4, 5 and 9 are categorized as industrial whereas the loads at bus #12, 13 and 14 are considered commercial. We have two sets of 3 uncertain parameters and our output of interest is the magnitude of the voltage at bus #4. We will approach the example as follows; we will first generate a PCM linear fit for the mapping between the uncertain loads and the voltage at bus #4. As there are 6 uncertain parameters, to come up with a PCM linear polynomial we would require 64 system simulations. Then, by applying the information theoretic techniques discussed earlier, we attempt reduce the number of



uncertain loads viz. from 6 uncertain loads to just two uncertain loads. The rationale here is that under each load type we have assumed two of the loads to be strongly dependent on one predominant load i.e. load at bus #4 in the case of industrial and the load at bus # 12 for commercial, as a result they would have sufficiently high mutual information and correlation coefficient values for us to reduce the number of input uncertain parameters. After uncertain parameters reduction, we attempt to model the input-output mapping using only the above mentioned two predominant loads. The details of the load distributions are given below in table 6.1. For ease of mathematical representation, we label the loads at buses 4, 5, 9, 12, 13, 14 as a, b, c and x, y, z respectively.

<b>Distribution details</b>
<p><b>Industrial Loads</b></p> <p><math>0.1 &lt; a &lt; 0.6</math> Uniformly distributed</p> <p><math>a - 0.03 &lt; b &lt; a + 0.03</math> Uniformly distributed</p> <p><math>a - 0.02 &lt; c &lt; a + 0.02</math> Uniformly distributed</p> <p><math>I(A; B; C) \approx 17.7564, H(A, B, C) \approx -0.103064, Cov(a, b) \approx 0.998847, Cov(a, c) \approx 0.999418</math></p>
<p><b>Commercial Loads</b></p> <p><math>0.25 &lt; x &lt; 0.7</math> Uniformly distributed</p> <p><math>x - 0.04 &lt; y &lt; x + 0.04</math> Uniformly distributed</p> <p><math>x - 0.03 &lt; z &lt; x + 0.03</math> Uniformly distributed</p> <p><math>I(X; Y; Z) \approx 9.44264, H(X, Y, Z) \approx -0.112934, Cov(x, y) \approx 0.99, Cov(x, z) \approx 0.99</math></p>

**Table 6.1:** Distribution details, Mutual information and Correlation coefficient values.



**Figure 6.1:** IEEE 14 bus test system

## RESULTS

With 6 uncertain loads	With 2 uncertain loads
Mean = 0.9984	Mean = 0.997118
Variance = 0.00016829	Variance = 0.000166101

The mean and variance values agree with each other. The numerical results show that the reduction is successful.

In the remainder of this chapter we will discuss a possible application of PCM in Computer Science.

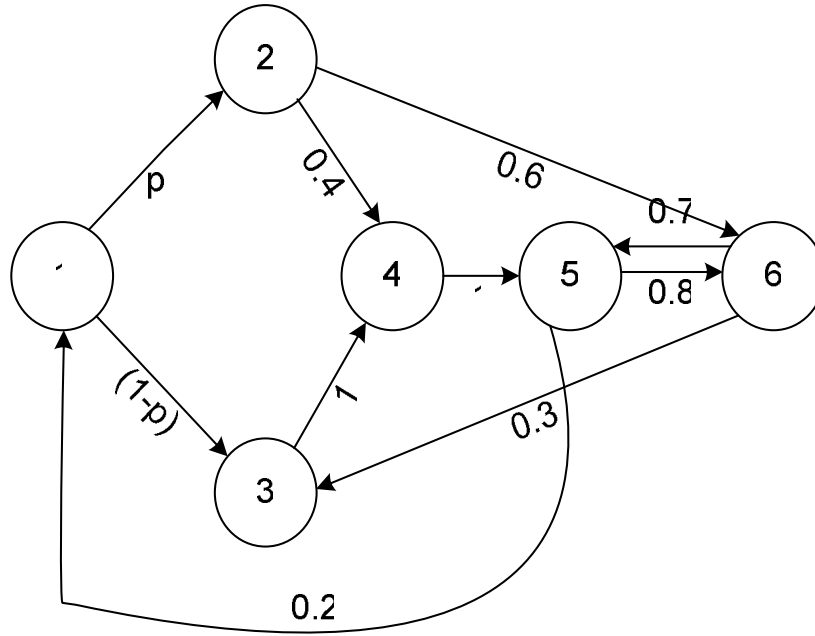
## **6.2 COMPUTER SCIENCE APPLICATION**

Increasingly, evaluation and optimization of communication/computer networks requires intensive simulation. For instance, the problem of optimal routing/resource allocation for multicasting in ad-hoc networks (e.g., [28]) often must be solved using computationally-intensive heuristics or exhaustive search algorithms. When, further, parameters in these networks (e.g., demands, channel capacities, buffer sizes) are uncertain or variable, characterization of the network's parameters over the range of parameter values is difficult. The toy example described in this section exposes that PCM can potentially be used to reduce simulation in performance-evaluation of communication/computer networks with variable parameters.

Specifically, in this example, we consider performance-characterization of an optimized Jackson (queueing) network model operating over a range of possible demands (inflow rates). The analysis of queueing networks with variable and uncertain parameters has been considered in [27]; however, this article does not consider any optimization of the network model, and hence analytical results on the network performance can be obtained.

Unfortunately, even in the highly abstracted realm of queueing-network modeling, optimization and transient-analysis problems typically must be solved through simulation (see, e.g., [29]). The complexity of optimizing queueing-network performance is a primary motivation for using PCM, when the network must operate over a range of parameter values.

The example Jackson network that we consider is shown in Figure 6.2; we refer the reader to [30] for further details on Jackson networks. Given a set of demands (input flow rates), the routing probability of the jobs leaving Queue 1 can be designed to optimize network performance. Specifically, given particular demands, we assume that the routing parameter  $p \in [0,1]$  is designed to minimize the *bottleneck expected queue length* (i.e., the largest expected queue length in the network). We note that optimizations of this sort are in general computationally intensive; in our case, we have implemented the optimization simply by numerically computing the bottleneck expected queue length as a function of  $p$  and choosing the minimizing  $p$ .



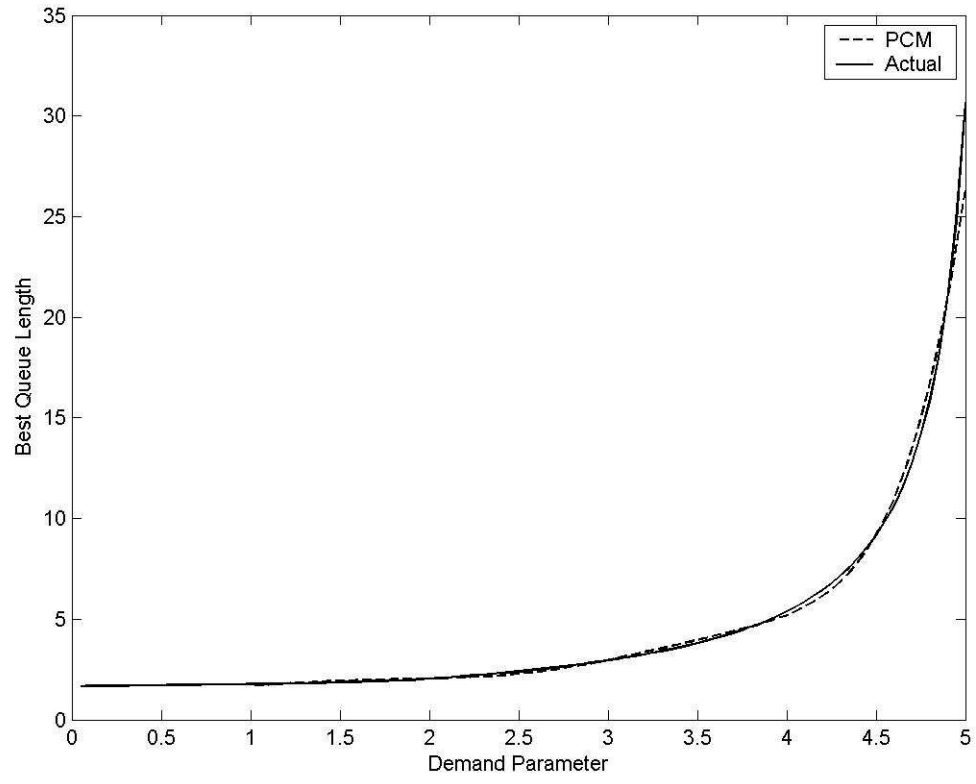
**Figure 6.2:** Jackson Network

As the system designers, we expect the demands to vary with time. For each set of demands, the system can be re-designed for optimal performance (minimum expected queue length at the bottleneck). However, we wish to make sure that this optimal solution meets system requirements such as queue length, delay, and server busyness requirements, over the domain of demand values. We contend that PCM is a valuable tool for characterizing these outputs in terms of the uncertain demand parameters.

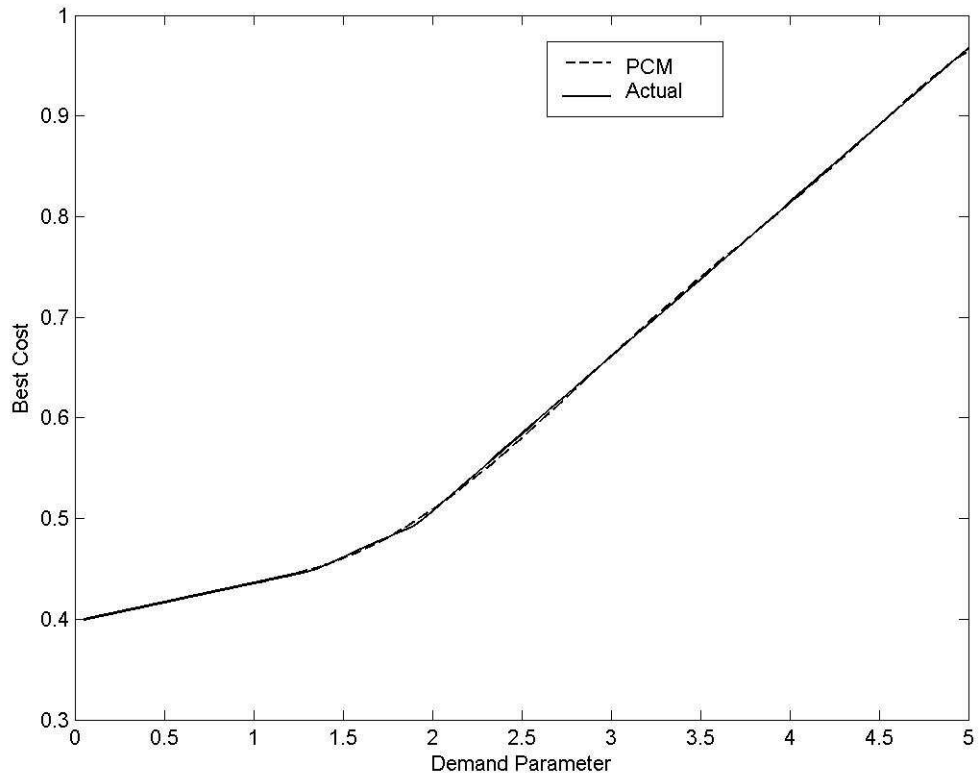
In this illustrative example, we assume that the demand at Queue 2 is variable, with any input rate between 0 and 5 equally likely. Hence, we apply one-dimensional PCM with the uncertain parameter given by the demand at Queue 2, which we assume to be uniformly distributed on the interval  $[0, 5]$ . In particular, we identify the mapping between this uncertain parameter and two outputs of interest, namely the bottleneck

expected queue length and the fraction of time the bottleneck queue is busy. Figures 6.3 and 6.4 compare the PCM-generated mappings for these outputs with mappings generated by exhaustive simulation. The output distribution for the above mentioned outputs of interest is depicted in figures 6.5 and 6.6. These plots suggest that the 7<sup>th</sup> Order PCM polynomials generated are good approximations of the actual relations.

This is an encouraging result and we hope to delve deeper in to the field of Computer Science in the future, and identify more PCM applications.

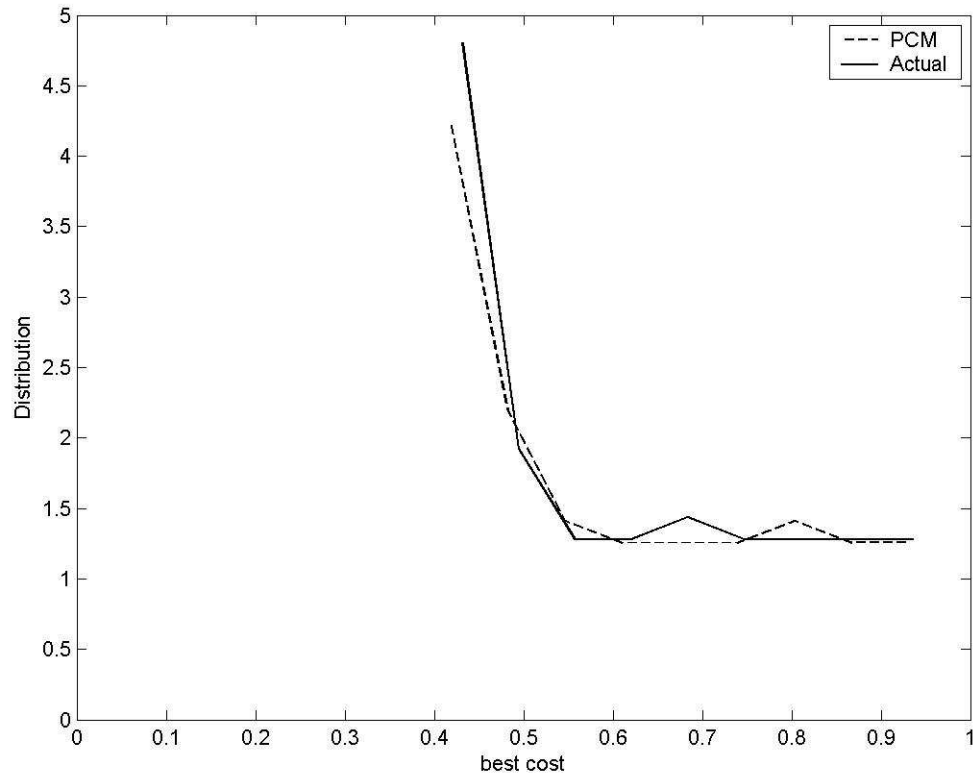


**Figure 6.3** Expected Queue Length plot

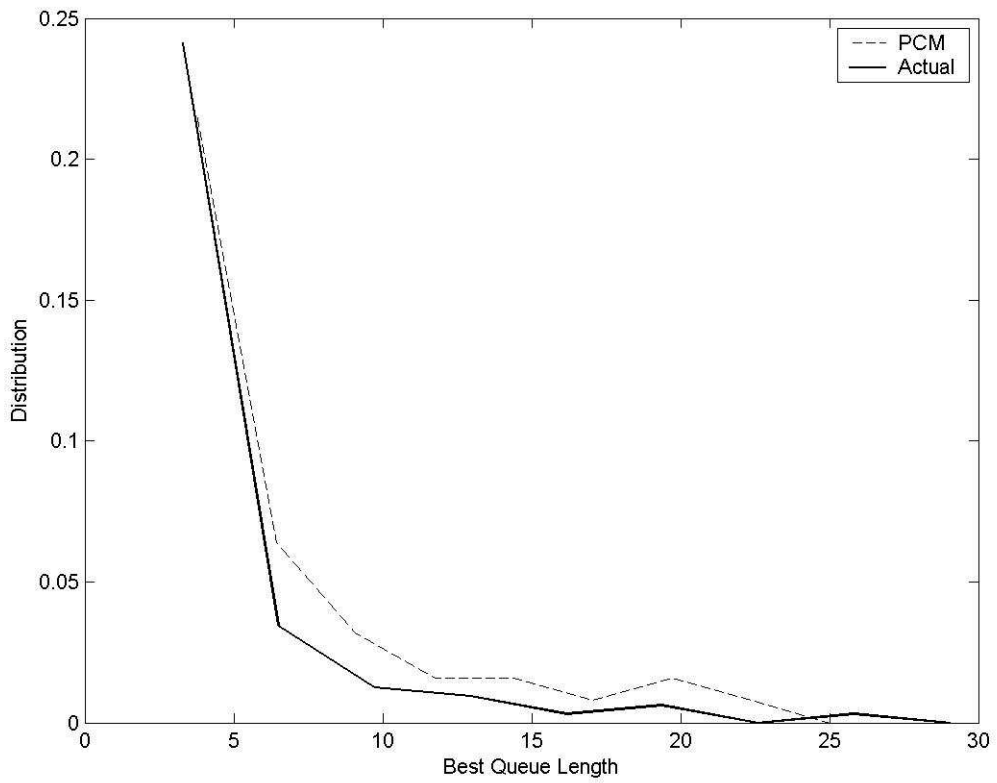


**Figure 6.4:** Busyness plot





**Figure 6.5** Comparison of distributions for busyness



**Figure 6.6** Comparison of distributions for queue length

## **CHAPTER 7**

### **SUMMARY AND FUTURE WORK**

In this thesis we have attempted to show that the Probabilistic Collocation Method (PCM) is an effective technique for simulating and modeling outputs of complex processes.

In the first chapter we motivated the need for PCM and talked briefly about some prevalent uncertainty analysis techniques. In chapter 2 we introduced the one-dimensional PCM, discussed its underlying theory of Gaussian quadrature and orthogonal polynomials. We illustrated one dimensional PCM with the help of an example from physical chemistry and wrapped up the chapter by discussing some refinements to the algorithm viz. PCM with sensitivity information, and error bounds on PCM. In the succeeding chapter we provided our generalization of PCM to handle multiple correlated uncertain parameters, we also proposed a way to mesh PCM with boundary load flow algorithms for filtering out some of the uncertain variables and an order selection algorithm for selecting the appropriate PCM order, we provided examples to illustrate each idea. In chapter 4 we talked about optimization problems and discussed the possibility of using PCM for solving optimization problems. This aspect of PCM is in the incipient stage, but the results look promising.

Information theoretic approaches for reducing the number of input uncertain variables were discussed in chapter 6; two cases were identified and the approach for each was discussed with illustrative examples. In chapter 7, we applied PCM to model the input-

output relationship for a 14 Bus IEEE test system and then with the aid of the input variable reduction techniques discussed earlier we developed another mapping for the same system with far less input variables. The results obtained were promising.

## **FUTURE DIRECTIONS**

- *Optimization:* As mentioned earlier, our attempt at applying PCM for solving optimization problems is rudimentary. In the future, we would like to attune PCM for handling problems in this domain.
- *Justification:* We have corroborated most of our results with analytical proofs. In the future we would like to make refinements, if necessary, and attempt to publish our results in an applied math context.
- *Software Package:* Develop a package for PCM. We do have Mathematica modules for PCM and most of the results presented in this thesis were generated using them. However, we would want a version of PCM coded in a programming language like C/C++.
- *Areas of Application:* This thesis shows a glimpse of PCM's versatility. In the future, we would like to identify PCM applications in areas that have not been discussed in this thesis.

## BIBLIOGRAPHY

- [1] J. R. Hockenberry and B. C. Lesieutre, "Evaluation of uncertainty in dynamic simulations of power systems models: the probabilistic collocation method," to appear in the *IEEE Transactions on Power Systems*.
- [2] J. R. Hockenberry, *Evaluation of Uncertainties in Dynamic, Reduced-Order Power System Models*, Ph.D. Thesis, Massachusetts Institute of Technology, Sep., 2000.
- [3] S. Roy, D. Ramamurthy, and B.C. Lesieutre, "Studies on the probabilistic collocation method and its application to power systems analysis," *36th Annual North American Power Symposium, Moscow, Idaho, Aug. 2004*
- [4] R. Billinton, M. Fotuhi-Firuzabad, and L. Bertling, "Bibliography on the application of probability methods in power system reliability evaluation 1996-1999," *IEEE Transactions on Power Systems*, Vol. 16, No. 4, Nov. 2001.
- [5] M. Webster, M.A. Tatang, and G. J. McRae, "Application of the probabilistic collocation method for an uncertainty analysis of a simple ocean model," Joint Program on the Science and Policy of Global Change, MIT, Cambridge, MA, Tech. Rep. 4, Jan. 1996.

- [6] T. B. Nguyen, M. A. Pai, and I. A. Hiskens, "Bounding uncertainty in power system dynamic simulations", *Proceedings of the IEEE PES 2000 Winter Meeting*, Singapore, January 2000.
- [7] A. Dimitrovski and K. Tomsovic, "Boundary load flow solutions", *IEEE Transactions on Power Systems*, Vol. 19, No. 1, Feb. 2004, pp. 348-355.
- [8] P.J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, 2<sup>nd</sup> Edition, Academic Press, New York, 1984.
- [9] G. Akrivis, "The Error Norm of Certain Gaussian Quadrature Formulae", *Mathematics of Computation*, Vol. 45, No. 172, Oct. 1985, pp. 513-519.
- [10] A. R. Bergen and V. Vittal, *Power Systems Analysis (2<sup>nd</sup> Edition)*, Prentice Hall, Upper Saddle River New Jersey, 2000.
- [11] A. M. L. da Silva and V. L. Arienti, "Probabilistic load flow by a multilinear simulation algorithm", *Proceedings of the IEE C*, Vol. 137, No. 4, pp. 276-282, July 1990.
- [12] R. N. Allan and A. M. L. da Silva, "Probabilistic load flow using multilinearizations", *Proceedings of the IEE C*, Vol. 128, No. 5, pp. 280-287, Sep. 1981.

- [13] T. M. Cover and J. Thomas, *Elements of Information Theory*, Wiley & Sons, New York, 1991.
- [14] S.S. Isukapalli and P. G. Georgopoulos, “Uncertainty analysis of transport transformation models,” PhD. Thesis, Rutgers State University, Jan., 1999.
- [15] S. S. Isukapalli, A. Roy, and P. G. Georgopoulos, “Stochastic Response Surface Methods (SRSMs) for Uncertainty Propagation: Application to Environmental and Biological Systems,” *Risk Analysis*, 18(3):351-363, 1998.
- [16] R. Y. Rubenstein, *Simulation and the Monte Carlo method*, Wiley series in probability and statistics, 1981.
- [17] L. Mathelin and M.Y. Hussaini, “ A Stochastic collocation algorithm for uncertainty analysis,”
- [18] D. P. Bertsekas, *Nonlinear Programming, 2<sup>nd</sup> Edition*, Athena Scientific, Belmont, 1999.
- [19] L. Mathelin and M.Y. Hussaini, “A Stochastic Collocation algorithm for uncertainty analysis,” NASA/CR-2003-212153, 2003.

- [20] E. Toth and A. Brath, "Flood forecasting using artificial neural networks in black-box and conceptual rainfall-runoff modeling," *Proceedings of the 1st Biennial Meeting of the iEMSs 2002*
- [21] I. Rivals and L. Personnaz, "Black-box modeling with state-space neural networks," *Neural Adaptive Control Technology*, R. Zbikowski and K. J. Hunt eds., World Scientific 1996, pp. 237-264.
- [22] [http://en.wikipedia.org/wiki/Ideal\\_gas](http://en.wikipedia.org/wiki/Ideal_gas), *Wikipedia*, the free encyclopedia.
- [23] [http://mathews.ecs.fullerton.edu/n2003/gaussianquad/GaussianQuadBib/Links/GaussianQuadBib\\_Ink\\_3.html](http://mathews.ecs.fullerton.edu/n2003/gaussianquad/GaussianQuadBib/Links/GaussianQuadBib_Ink_3.html), Bibliography for Gauss-Legendre quadrature.
- [24] <http://mathworld.wolfram.com/CorrelationCoefficient.html>, *Mathworld*, Wolfram Research.
- [25] A. Markoff, "Sur la methode de Gauss pur le Calcul approche des integrals," *Math. Ann.*, 25, pp. 427-432, 1885.
- [26] M. A. Bokhari, H. Taufiq, "An extension of Gauss quadrature formula," *Technical Report 319*, Dept. of Mathematical Sciences, King Fahd University, May 2004.



- [27] J. Luthi, G. Haring, “Bottleneck Analysis for Computer and Communication Systems with Workload Variabilities & Uncertainties,” in *Proceedings of the Second International Symposium on Mathematical Modeling*, I. Troch and F. Breitenecker, ed., pp 525-534, Vienna, Austria, Feb. 1997.
- [28] D. O’Neill and Y. Li, “Optimal routes and flows in multicasting over ad hoc networks,” in *Proceedings of the IEEE Vehicular Technologies Conference*, May 2004, Milan, Italy.
- [29] D. Gross and C. M. Harris, *Fundamentals of Queueing Theory, 2nd ed.* John Wiley and Sons, 1985.
- [30] F. P. Kelly, *Reversibility and Stochastic Networks, 3rd ed.* John Wiley and Sons, 1994. Available at [www.statslab.cam.ac.uk/~frank.rsn.html](http://www.statslab.cam.ac.uk/~frank.rsn.html).