Efficient Sampling for Non-Intrusive Polynomial Chaos Applications with Multiple Uncertain Input Variables

Serhat Hosder
Missouri University of Science and Technology, hosders@mst.edu

Robert W. Walters

Michael Balch

48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference
 15th
23 - 26 April 2007, Honolulu, Hawaii

AIAA 2007-1939

Efficient Sampling for Non-Intrusive Polynomial Chaos Applications with Multiple Uncertain Input Variables

Serhat Hosder*, Robert W. Walters† and Michael Balch ‡

Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061, USA

The accuracy and the computational efficiency of a Point-Collocation Non-Intrusive Polynomial Chaos (NIPC) method applied to stochastic problems with multiple uncertain input variables has been investigated. Two stochastic model problems with multiple uniform random variables were studied to determine the effect of different sampling methods (Random, Latin Hypercube, and Hammersley) for the selection of the collocation points. The effect of the number of collocation points on the accuracy of polynomial chaos expansions were also investigated. The results of the stochastic model problems show that all three sampling methods exhibit a similar performance in terms of the the accuracy and the computational efficiency of the chaos expansions. It has been observed that using a number of collocation points that is twice more than the minimum number required gives a better approximation to the statistics at each polynomial degree. This improvement can be related to the increase of the accuracy of the polynomial coefficients due to the use of more information in their calculation. The results of the stochastic model problems also indicate that for problems with multiple random variables, improving the accuracy of polynomial chaos coefficients in NIPC approaches may reduce the computational expense by achieving the same accuracy level with a lower order polynomial expansion. To demonstrate the application of Point-Collocation NIPC to an aerospace problem with multiple uncertain input variables, a stochastic computational aerodynamics problem which includes the numerical simulation of steady, inviscid, transonic flow over a three-dimensional wing with an uncertain free-stream Mach number and angle of attack has been studied. For this study, a 5^{th} degree Point-Collocation NIPC expansion obtained with Hammersley sampling was capable of estimating the statistics at an accuracy level of 1000 Latin Hypercube Monte Carlo simulations with a significantly lower computational cost.

I. Introduction

In the stochastic modeling of most large scale aerospace problems with multiple uncertain input variables and parameters, computational efficiency becomes an important factor in the selection of the method to be used due to the expensive nature of computational simulations. In addition to the efficiency, a certain level of accuracy (confidence) is also desired in the solution of the stochastic problems. This paper will address both efficiency and the accuracy of a Point-Collocation Non-Intrusive Polynomial Chaos (NIPC) approach applied to stochastic problems with multiple uncertain variables. The effect of the number of collocation points and different sampling techniques for the selection of collocation points will be investigated in terms of the accuracy of the polynomial chaos expansions and the computational efficiency.

Polynomial Chaos (PC) method is among several methods to model and propagate uncertainty in stochastic computational simulations. Several researchers have studied and implemented the PC approach for a wide range of problems. Ghanem and Spanos[1] (1990) and Ghanem[2,3] (1999) applied the PC method to several problems of interest to the structures community. Mathelin et al.[4] studied uncertainty propagation for a turbulent, compressible nozzle flow by this technique. Xiu and Karniadakis[5] analyzed the flow past a circular cylinder and incompressible channel flow by the PC method and extended the method beyond the

*Postdoctoral Associate, Aerospace and Ocean Engineering (AOE) Department, AIAA Member.
†Professor, AOE Department, Interim Associate Vice President for Research, AIAA Associate Fellow.
‡Graduate Student, AOE Department, AIAA Student Member.

original formulation of Wiener[6] to include a variety of basis functions. In 2003, Walters[7] applied the PC method to a two-dimensional steady-state heat conduction problem for representing geometric uncertainty.

The Polynomial chaos (PC) method for the propagation of uncertainty in computational simulations involves the substitution of uncertain variables and parameters in the governing equations with the polynomial expansions. In general, an intrusive approach will calculate the unknown polynomial coefficients by projecting the resulting equations onto basis functions (orthogonal polynomials) for different modes. As its name suggests, the intrusive approach requires the modification of the deterministic code and this may be difficult, expensive, and time consuming for many complex computational problems such as the full Navier-Stokes simulation of 3-D, viscous, turbulent flows around realistic aerospace vehicles, chemically reacting flows, numerical modeling of planetary atmospheres, or multi-system level simulations which include the interaction of many different codes from different disciplines. To overcome such inconveniences associated with the intrusive approach, non-intrusive polynomial chaos formulations (NIPC) have been developed for uncertainty propagation. Most of the NIPC approaches in the literature are based on *sampling* (Debusschere et al.,[8] Reagan et. al.,[9] and Isukapalli[10]) or *quadrature* methods (Debusschere et al.[8] and Mathelin et al.[11]) to determine the projected polynomial coefficients. Recently, Hosder and Walters[12] applied a Point-Collocation NIPC to selected stochastic Computational Fluid Dynamics (CFD) problems and Loeven et al[13] introduced a non-intrusive Probabilistic Collocation approach for efficient propagation of arbitrarily distributed parametric uncertainties.

In the following section, a description of the non-intrusive polynomial chaos approaches including the Point-Collocation NIPC method is given. The sampling techniques investigated for the selection of collocation points are outlined in Section III. Section IV includes the results obtained from the investigation of two stochastic model problems with multiple uniform random variables. The results of a stochastic computational aerodynamics problem which includes the numerical simulation of steady, inviscid, transonic flow over a three-dimensional wing with an uncertain free-stream Mach number and angle of attack will also be presented in Section IV to demonstrate the application of the Point-Collocation NIPC to a stochastic aerospace problem with multiple uncertain input variables. The conclusions will be given in Section V.

II. Non-Intrusive Polynomial Chaos Approaches

The polynomial chaos is a stochastic method, which is based on the spectral representation of the uncertainty. An important aspect of spectral representation of uncertainty is that one may decompose a random function (or variable) into separable deterministic and stochastic components. For example, for any random variable (*i.e.*,α) such as velocity, density or pressure in a stochastic fluid dynamics problem, we can write,

$$\alpha\,(\vec{x}\vec{\xi}) = \sum_{i=0}^{P} \alpha_i(\vec{x})\Psi_i(\vec{\xi}) \tag{1}$$

where $\alpha_i(\vec{x})$ is the deterministic component and $\Psi_i(\vec{\xi})$ is the random basis function corresponding to the i^{th} mode. Here we assume α to be a function of deterministic independent variable vector $\vec{x}$ and the n-dimensional random variable vector $\vec{\xi} = (\xi_1, ..., \xi_n)$ which has a specific probability distribution. The discrete sum is taken over the number of output modes ,

$$P + 1 = \frac{(n+p)!}{n!p!} \tag{2}$$

which is a function of the order of polynomial chaos (p) and the number of random dimensions (n). The basis function ideally takes the form of multi-dimensional Hermite Polynomial to span the n-dimensional random space when the input uncertainty is Gaussian (unbounded), which was first used by Wiener[6, 14] in his original work of polynomial chaos. Legendre (Jacobi) and Laguerre polynomials are optimal basis functions for bounded (uniform) and semi-bounded (exponential) input uncertainty distributions respectively in terms of the convergence of the statistics. Different basis functions can be used with different input uncertainty distributions (See Xiu and Karniadakis[5] for a detailed description), however the convergence may be affected depending on the basis function used.[15] In all the stochastic model problems studied in this paper, we model the input uncertainties as uniform random variables that have bounded probability distributions. Therefore, in our Point-Collocation NIPC method described below we use multi-dimensional Legendre polynomials that are orthogonal in the interval [-1,1] for each random dimension. The detailed information about polynomial chaos expansions can be found in Walters and Huyse[16] and Hosder et al.[12]

To model the uncertainty propagation in computational simulations via polynomial chaos with the intrusive approach, all dependent variables and random parameters in the governing equations are replaced with their polynomial chaos expansions. Taking the inner product of the equations, (or projecting each equation onto k^{th} basis) yield $P+1$ times the number of deterministic equations which can be solved by the same numerical methods applied to the original deterministic system. Although straightforward in theory, an intrusive formulation for complex problems can be relatively difficult, expensive, and time consuming to implement. To overcome such inconveniences associated with the intrusive approach, non-intrusive polynomial chaos formulations have been considered for uncertainty propagation.

The objective of the non-intrusive polynomial chaos methods is to obtain approximations of the polynomial coefficients without making any modifications to the deterministic code. Main approaches for non-intrusive polynomial chaos are sampling based, collocation based, and quadrature methods. To find the polynomial coefficients $\alpha_k = \alpha_k(\vec{x})$, $(k = 0, 1, .., P)$ in Equation 1 using sampling based and quadrature methods, the equation is projected onto k^{th} basis:

$$\left\langle \alpha\ (\vec{x}\vec{\xi}), \Psi_k(\vec{\xi}) \right\rangle = \sum_{i=0}^{P} \alpha_i \Psi_i(\vec{\xi}), \Psi_k(\vec{\xi}) \tag{3}$$

where the inner product of two functions $f(\vec{\xi})$ and $g(\vec{\xi})$ is defined by

$$\left\langle f(\vec{\xi}) g(\vec{\xi}) \right\rangle = {}_R\, f(\vec{\xi}) g(\vec{\xi}) p_N(\vec{\xi}) d\vec{\xi} \tag{4}$$

Here the weight function $p_N(\vec{\xi})$ is the probability density function of $\vec{\xi}$ and the integral is evaluated on the support (R) region of this weight function. Using the orthogonality of the basis functions,

$$\left\langle \alpha\ (\vec{x}\vec{\xi}), \Psi_k(\vec{\xi}) \right\rangle = \alpha_k \left\langle \Psi_k^2(\vec{\xi}) \right\rangle \tag{5}$$

we can obtain

$$\alpha_k = \frac{\left\langle \alpha\ (\vec{x}\vec{\xi}), \Psi_k(\vec{\xi}) \right\rangle}{\left\langle \Psi_k^2(\vec{\xi}) \right\rangle} \tag{6}$$

In sampling based methods, the main strategy is to compute $\alpha\ (\vec{x}\vec{\xi})\Psi_k(\vec{\xi})$ for a number of samples ($\vec{\xi}_i$ values) and perform averaging to determine the estimate of the inner product $\left\langle \alpha\ (\vec{x}\vec{\xi}), \Psi_k(\vec{\xi}) \right\rangle$. Quadrature methods calculate the same term, which is an integral over the support of the weight function $p_N(\vec{\xi})$, using numerical quadrature. Once this term is evaluated, both methods (sampling based and quadrature) use Equation 6 to estimate the projected polynomial coefficients for each mode.

A. Point-Collocation Non-Intrusive Polynomial Chaos

The collocation based NIPC method starts with replacing the uncertain variables of interest with their polynomial expansions given by Equation 1. Then, $P+1$ vectors $(\vec{\xi}_i = \{\xi_1, \xi_2, ..., \xi_n\}_i,\ i = 0, 1, 2, ..., P)$ are chosen in random space for a given PC expansion with $P+1$ modes and the deterministic code is evaluated at these points. With the left hand side of Equation (1) known from the solutions of deterministic evaluations at the chosen random points, a linear system of equations can be obtained:

$$\begin{matrix} \Psi_0(\vec{\xi}_0) & \Psi_1(\vec{\xi}_0) & \cdots & \Psi_P(\vec{\xi}_0) \\ \Psi_0(\vec{\xi}_1) & \Psi_1(\vec{\xi}_1) & \cdots & \Psi_P(\vec{\xi}_1) \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_0(\vec{\xi}_P) & \Psi_1(\vec{\xi}_P) & \cdots & \Psi_P(\vec{\xi}_P) \end{matrix} \quad \begin{matrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_P \end{matrix} \quad = \quad \begin{matrix} \alpha\ (\vec{\xi}_0) \\ \alpha\ (\vec{\xi}_1) \\ \vdots \\ \alpha\ (\vec{\xi}_P) \end{matrix} \tag{7}$$

The spectral modes (α_k) of the random variable, α , are obtained by solving the linear system of equations given above. Using these, mean (μ_{α^*}) and the variance $(\sigma^2_{\alpha^*})$ of the solution can be obtained by

$$\mu_{\alpha^*} = \alpha_0 \tag{8}$$

$$\sigma_{\alpha^*}^2 = \sum_{i=1}^{P} \alpha_i^2 \left\langle \Psi_i^2(\vec{\xi}) \right\rangle \tag{9}$$

The Point-Collocation approach was first introduced by Walters[7] to approximate the polynomial chaos coefficients of the metric terms which are required as input stochastic variables for the intrusive polynomial chaos representation of a stochastic heat transfer problem with geometric uncertainty. Recently Hosder et. al.[12] applied this Point-Collocation NIPC method to stochastic fluid dynamics problems with geometric uncertainty. They demonstrated the efficiency and the accuracy of the NIPC method in terms of modeling and propagation of an input uncertainty and the quantification of the variation in an output variable. That study included a single random input variable, and the collocation locations were equally spaced in the the random space. In this paper, multiple uncertain input variables are considered and different sampling methods are investigated to determine if any particular technique gives optimum collocation points for efficiency and accuracy.

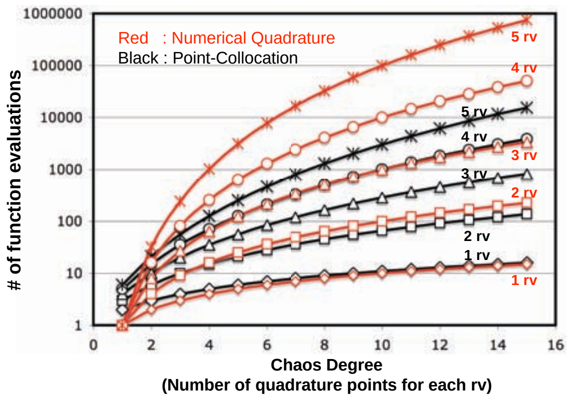


Figure 1. The number of function evaluations for the Point-Collocation ($n_p = 1$) and the numerical quadrature NIPC for different number of random variables.

The solution of linear problem given by Equation 7 requires $P+1$ deterministic function evaluations. If more than $P+1$ samples are chosen, then the over-determined system of equations can be solved using the Least Squares method. In this paper, we investigate this option by increasing the number of collocation points in a systematic way through the introduction of a parameter n_p defined as

$$n_p = \frac{\textit{number of samples}}{P+1} \tag{10}$$

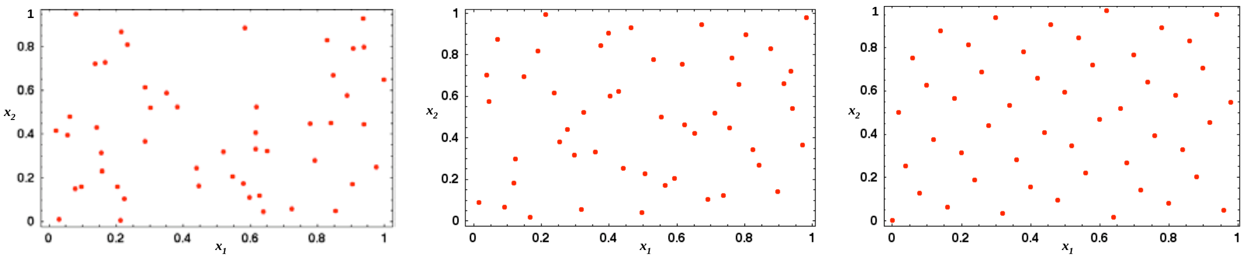
In the solution of stochastic model problems with multiple uncertain variables, we have used $n_p = 1, 2, 3,$ and 4 to study the effect of the number of collocation points (samples) on the accuracy of the polynomial chaos expansions.

Figure 1 shows the computational cost associated with the Point-Collocation and the numerical quadrature NIPC. For the point collocation, the number of function evaluations is $n_p \times (P+1)$ where P+1 is the number of output modes for a given polynomial degree and number of random variables (n) (See Equation 2). The number of function evaluations for the numerical quadrature is n_q^n where n_q is the number of quadrature points in each random dimension. For a single random variable, the number of function evaluations for each method is comparable. However, as the number of random variables increase, the computational cost of the numerical quadrature grows significantly. One may think of using an optimum number of quadrature points to reduce the cost, but for a general stochastic function or problem, considerable number of quadrature points may be required to evaluate the integration with desired accuracy as shown by Huyse et al.[15] For both methods, the computational cost becomes formidable with the increase of the polynomial degree and the number of random variables. It should be noted that in Figure 1 the limits of the number of function evaluations are extended to very large numbers to show the general trend. In reality, especially for large scale stochastic computations, one can afford only a certain number of deterministic runs to the produce the output values at the selected collocation or quadrature points. This again emphasizes the necessity of the implementation of innovative methods in large scale stochastic problems to model and propagate multiple input uncertainties with desired accuracy and efficiency.

III. Sampling Methods for the Point-Collocation NIPC approach

In this study, we investigate three widely-used sampling techniques in stochastic computations for the selection of the collocation points:

- Random Sampling (RS)
- Latin Hypercube Sampling (LHS)



(a) Random Sampling (RS) (b) Latin Hypercube Sampling (LHS) (c) Hammersley Sampling

Figure 2. Samples obtained with different methods in a two-dimensional random space

- Hammersley Sampling (HS)

These techniques are also used as Modern Design of Experiments (DOE) techniques in computational simulations for engineering design and optimization and are discussed in detail by Giunta *et al.*[17] In the stochastic model problems studied in this paper, we evaluate the NIPC expansions using the collocation points obtained from each technique and determine the performance for estimating the statistics (mean and standard deviation) in terms of the accuracy and computational cost. Figure 2 shows an example of 50 samples produced by each method in a two dimensional random space. Both variables (x_1 and x_2) are uniform random variables defined for the interval [0,1]. Compared to the RS (can also be referred as crude Monte Carlo sampling), LHS and HS exhibit a better coverage of the design space. Since the Hammersley points are calculated with a deterministic algorithm[17,18] for a given number of samples and random variables, they are scattered in a more organized pattern in the random space.

IV. Results

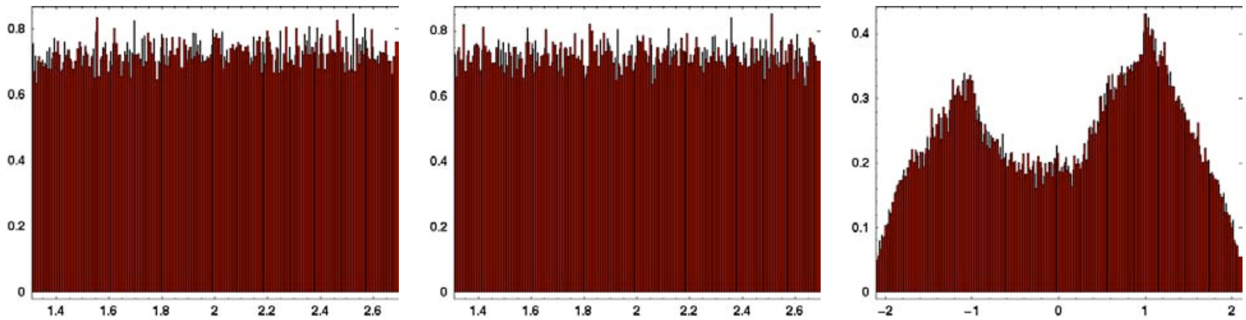
A. Stochastic Model Problem with Two Random Variables

The first stochastic model problem tested in this study is based on the equation

$$f(x_1, x_2) = ln \;\; 1 + x_1^2 \;\; Sin\,(5x_2)\,. \tag{11}$$

Both $x_1 = x_1(\xi_1)$ and $x_2 = x_2(\xi_2)$ are modeled as uniform random variables with a mean value of 2.0 and a coefficient of variation (CoV) of 20% which gives a constant probability density function (PDF) of 0.722 for each variable. Here, $\xi_{1,2}$ are uniform random variables defined in the interval $[-1, 1]$, which have a constant PDF of 0.5. This stochastic model problem is chosen to evaluate the performance of Point-Collocation NIPC method with different sampling techniques (RS, LHS, and HS) and number of collocation points (n_p) in estimating the statistics of $f(x_1, x_2) = f\,(x_1(\xi_1), x_2(\xi_2))$ which is a stochastic output variable with highly non-linear dependence on x_1 and x_2.

The statistics and the histogram of $f(x_1, x_2)$ were obtained with crude Monte Carlo (MC) simulations using 100,000 samples from the corresponding probability distributions of x_1 and x_2 (Figure 3(a) and 3(b)). The MC histogram of $f(x_1, x_2)$ is shown in Figure 3(c). The distribution is bounded, but significantly different than the input uniform distributions due to the non-linear nature of the stochastic problem. To compare the statistics of the MC simulations to the NIPC results, 95% confidence intervals (CI) for the MC mean and standard deviation were constructed using the Bootstrap Method. The advantage of the Bootstrap Method is that it is not restricted to a specific distribution, e.g. a Gaussian. It is easy and efficient to implement, and can be completely automated to any estimator, such as the mean or the variance. In practice, one takes 25–200 bootstrap samples to obtain a standard error estimate. In our computations, we used 250 bootstrap samples. Each sample consisted of 100,000 observations selected randomly from the original Monte Carlo simulations by giving equal probability $(1/100,000)$ to each observation. Table 1 presents the MC statistics and the associated 95% confidence intervals. For this stochastic model problem, the exact values of the statistics were also calculated. Note that the exact mean and the standard deviation fall within the 95% CI of MC results which validates the confidence interval estimates obtained by the Bootstrap method.



(a) Histogram of $x_1(\xi_1)$ (b) Histogram of $x_2(\xi_2)$ (c) Histogram of $f = f(x_1(\xi_1), x_2(\xi_2))$

Figure 3. Histograms of the input variables (x_1 and x_2) and the output variable ($f(x_1, x_2)$) obtained with 100,000 crude MC simulations.

Table 1. The exact and the Monte Carlo (MC) statistics obtained for the stochastic model problem with two uniform random variables. The 95% confidence intervals for the MC statistics were calculated using the Bootstrap method.

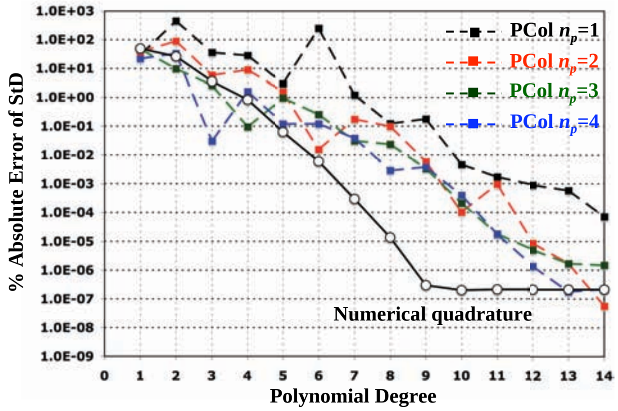
Statistics of $f(x_1, x_2)$	Exact	MC	95% CI of MC
Mean	0.079169123	0.07783876	[0.07084801, 0.08527571]
Standard Deviation	1.12413615	1.12605069	[1.1229222, 1.12894006]

For this model problem, systematic studies on the convergence of Point-Collocation NIPC statistics were performed to determine the effect of the number of collocation points (n_p = 1, 2, 3, and 4) and the effect of different sampling techniques (RS, LHS, and HS) for the selection of the collocation points. Figure 4 and 5 give a comprehensive summary on the convergence characteristics of the standard deviation (StD) of $f(x_1, x_2)$. In Figure 4 and 5, the first column of graphs give the convergence characteristics in terms of the absolute error which is defined as

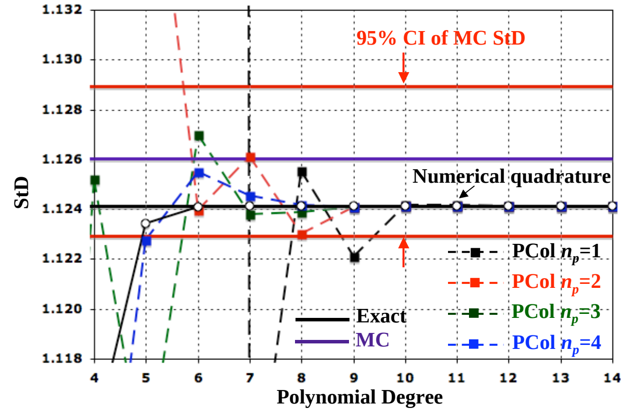
$$\% \text{ Absolute Error} = \frac{[StD_{exact} - StD_{NIPC}]}{StD_{exact}} \times 100 \tag{12}$$

where StD_{exact} stands for the exact value of the standard deviation and StD_{NIPC} is the NIPC approximation to the same statistics. In the second column of plots, the convergence behavior is analyzed with comparison to the exact solution and MC results which include the standard deviation and the associated 95% CI. All plots also include the convergence history of the standard deviation obtained with the numerical quadrature NIPC using 14 quadrature points in both random dimensions. The plots on each row of Figure 4 give the results for a particular sampling technique, while the plots on each row of Figure 5 present the results for a particular value of n_p.

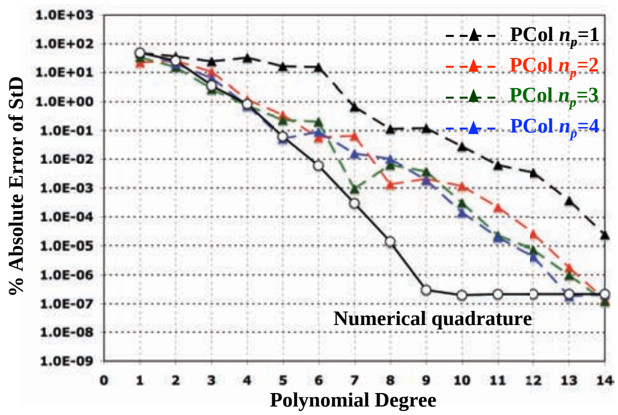
The effect of the number of collocation points on the convergence of the standard deviation can be analyzed by examining Figure 4. For all sampling methods, using a number of collocation points that is twice more than the minimum number required ($n_p = 2$) gives a better approximation to the statistics at each polynomial degree. This improvement can be related to the increase of the accuracy of the polynomial coefficients due to the use of more information (collocation points) in their calculation. For $n_p = 2$, the standard deviation of the Point-Collocation NIPC tends to fall within the 95% CI of the MC statistics with a lower degree polynomial compared to the case obtained with $n_p = 1$. This behavior is more significant for the collocation points obtained with LHS. For this sampling technique, the standard deviation of the Point-Collocation NIPC falls within the 95% CI of MC at a polynomial degree of 8 with $n_p = 1$, whereas it enters the CI at a polynomial degree of 6 for $n_p = 2$. The number of function evaluations is 45 for the first case and 56 for the second. In this specific example, the computational work required for $n_p = 2$ exceeds the computational work required for $n_p = 1$ to obtain the statistics at the same accuracy level. However, as will be shown in the next model problem, for stochastic problems with more random variables, using $n_p = 2$ can be computationally more efficient than using higher degree polynomials with $n_p = 1$ to achieve the same



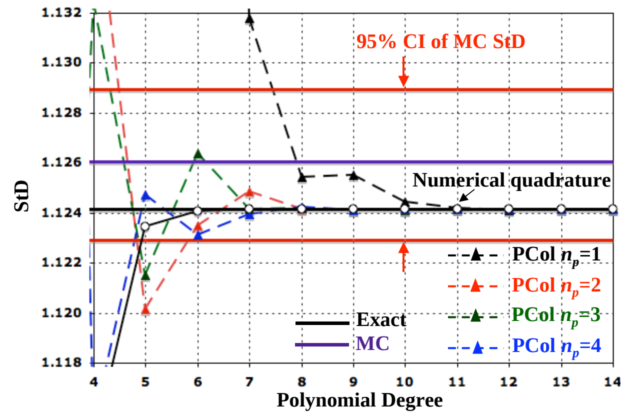
(a) Absolute error (RS)



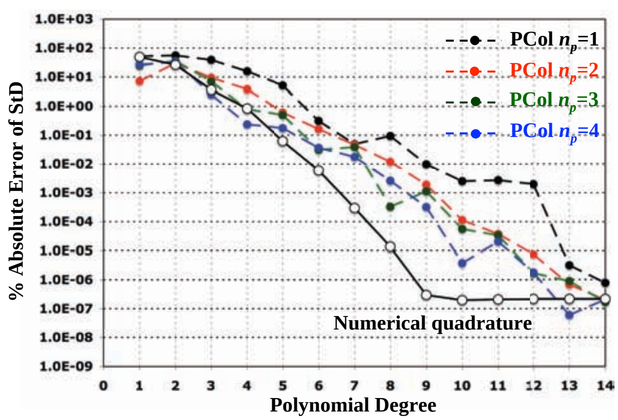
(b) Convergence (RS)



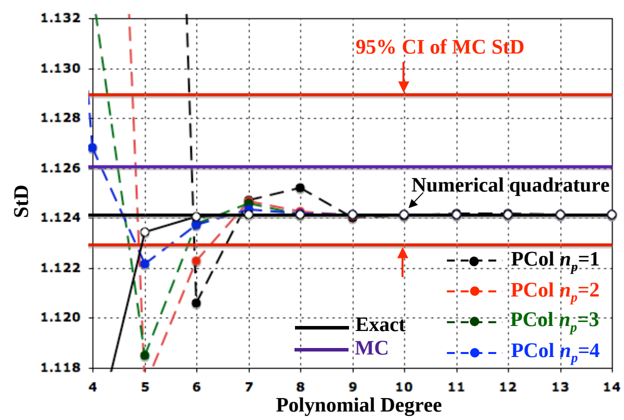
(c) Absolute error (LHS)



(d) Convergence (LHS)



(e) Absolute error (HS)

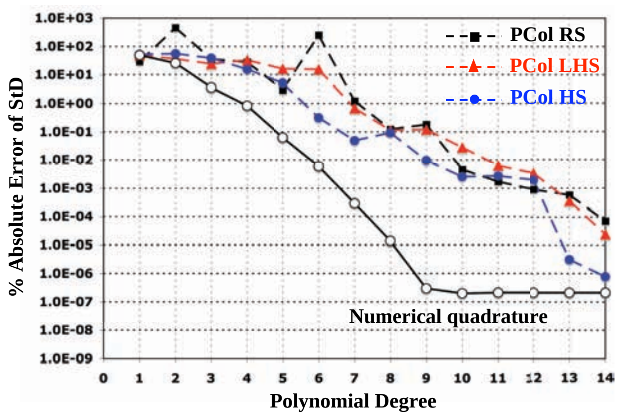


(f) Convergence (HS)

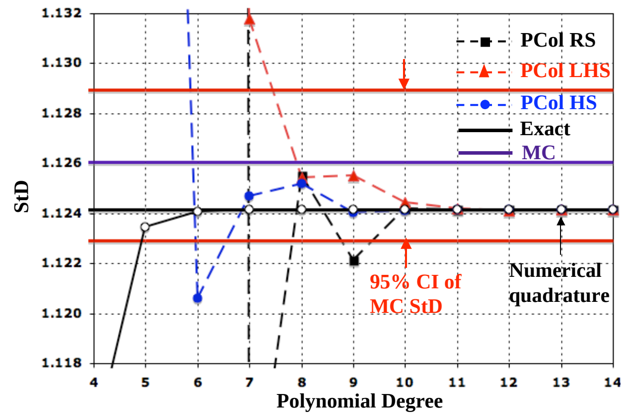
Figure 4. Absolute error and convergence of the standard deviation (StD) obtained with Point-Collocation (PCol) NIPC for the stochastic model problem with two uniform random variables. Each figure shows the results obtained with a specific sampling technique (RS, LHS, and HS).

accuracy level, since the required number of function evaluations grows significantly with the increase of the number of random variables. For this model problem, further increase of n_p beyond 2 does not give a significant accuracy improvement.

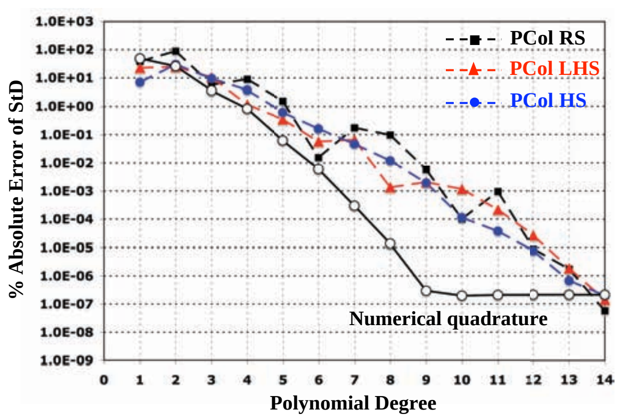
Figure 5 shows that for a particular value of n_p, none of the sampling methods have a significant advantage over each other in terms of accuracy. However, the convergence of the Point-Collocation NIPC statistics obtained with HS and $n_p = 2$ exhibits a much smoother (monotonic) convergence compared to the cases



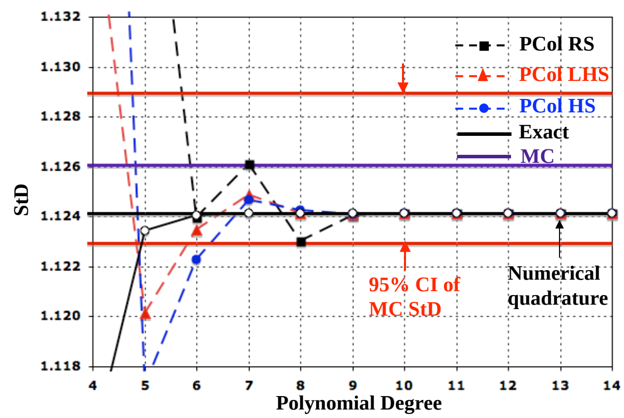
(a) Absolute error ($n_p = 1$)



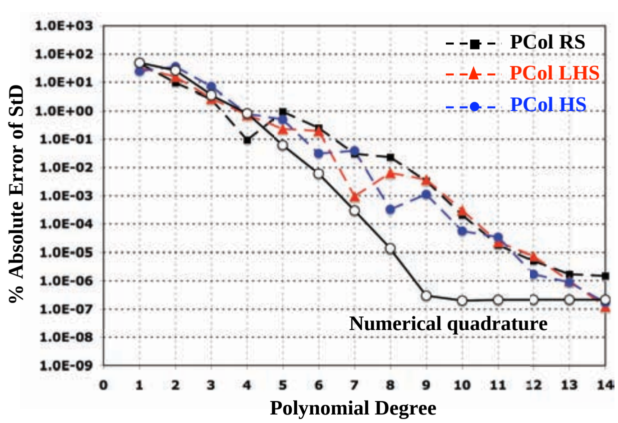
(b) Convergence ($n_p = 1$)



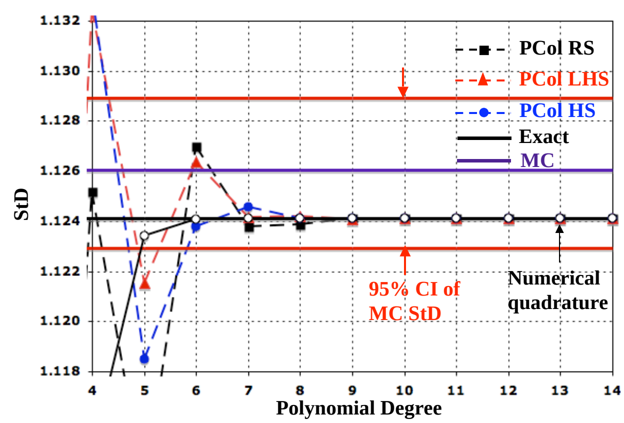
(c) Absolute error ($n_p = 2$)



(d) Convergence ($n_p = 2$)



(e) Absolute error ($n_p = 3$)

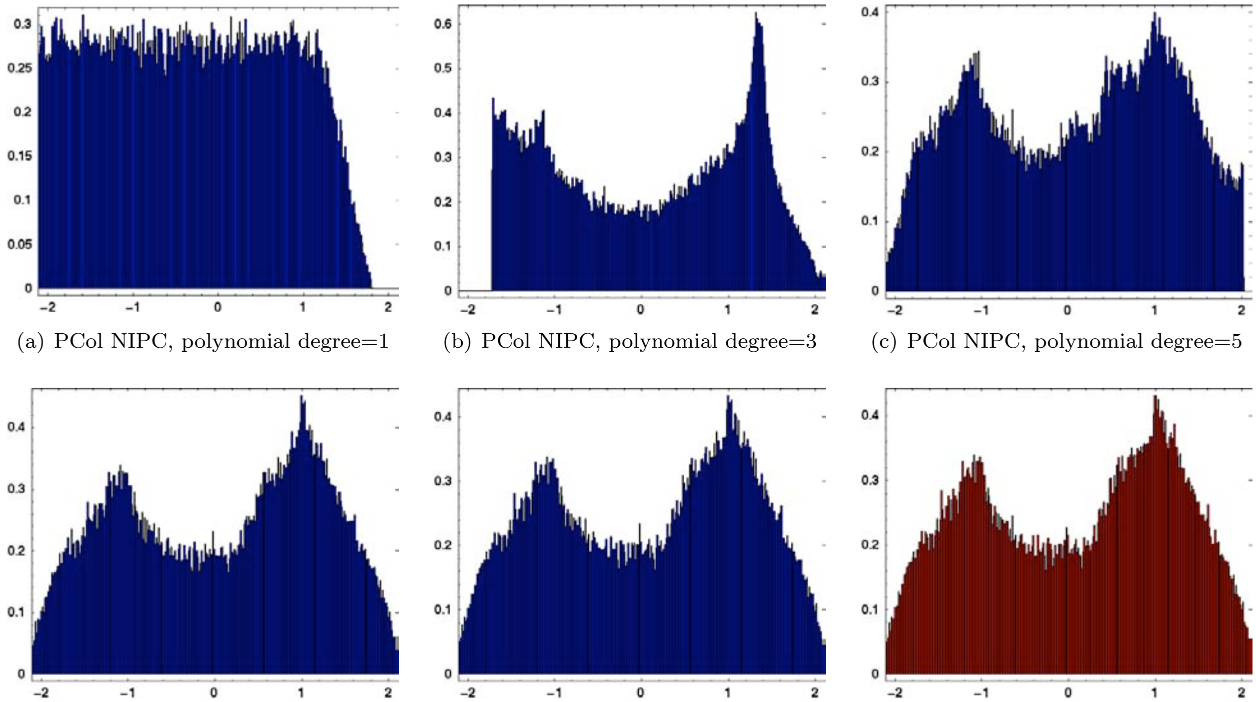


(f) Convergence ($n_p = 3$)

Figure 5. Absolute error and convergence of the standard deviation (StD) obtained with Point-Collocation (PCol) NIPC for the stochastic model problem with two uniform random variables. Each figure shows the results obtained with a specific value of n_p ($n_p = 1$, 2 and 3).

obtained with other sampling techniques. The standard deviation values of the numerical quadrature NIPC are more accurate than the Point-Collocation results for all polynomial degrees. The absolute error of the numerical quadrature converges to a value of approximately 10^{-7}% at a polynomial degree of 9 and the Point-Collocation statistics obtained with $n_p > 1$ reach this accuracy level at a polynomial degree of 14.

The histograms of $f(x_1, x_2)$ obtained with Point-Collocation NIPC (based on HS and $n_p = 2$) for various polynomial degrees are shown in Figure 6. The probability distribution of the polynomial chaos expansion



(a) PCol NIPC, polynomial degree=1 (b) PCol NIPC, polynomial degree=3 (c) PCol NIPC, polynomial degree=5

(d) PCol NIPC, polynomial degree=7 (e) PCol NIPC, polynomial degree=9 (f) Monte Carlo

Figure 6. The histogram of $f(x_1, x_2)$ obtained with the Point-Collocation (PCol) NIPC (HS and $n_p = 2$) for various polynomial degrees. Monte Carlo histogram is included for comparison.

converges to the Monte Carlo distribution at a polynomial degree of 7. Note that this is also the polynomial degree that the standard deviation of the same NIPC method falls within the 95% CI of the MC statistics (Figure 5).

B. Stochastic Model Problem with Four Random Variables

The second stochastic model problem is described by the exponential function

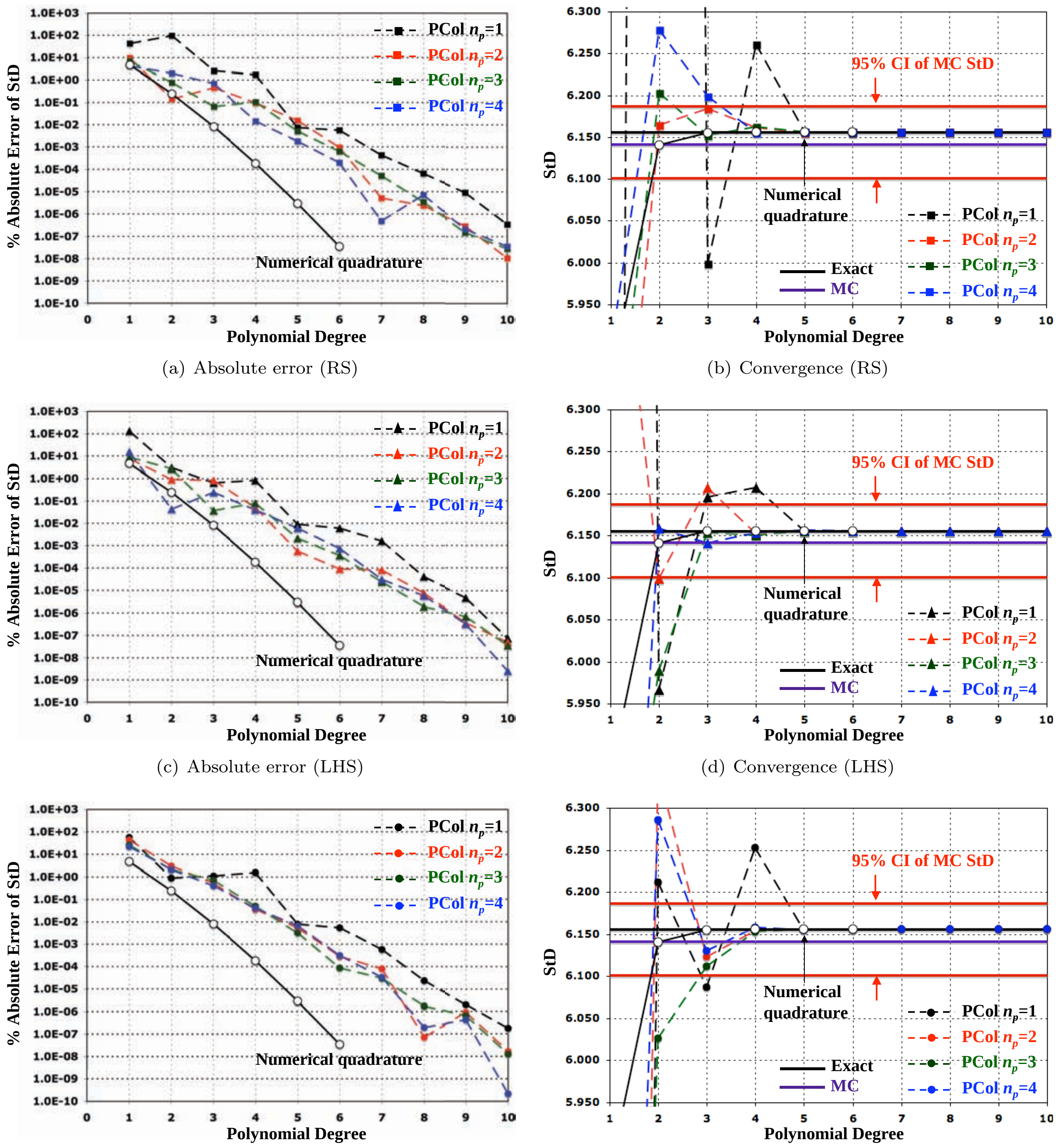
$$f(x_1, x_2, x_3, x_4) = e^{1.5(x_1+x_2+x_3+x_4)} \tag{13}$$

For this model problem, all $x_i = x_i(\xi_i)$ $(i = 1, ..4)$ are taken as uniform random variables with a mean value of 0.4 and a coefficient of variation (CoV) of 40% which gives a constant probability density function (PDF) of 1.804 for each x_i. Here, ξ_i are again uniform random variables defined in the interval $[-1, 1]$, which have a constant PDF of 0.5. This stochastic model problem with four random variables is studied to evaluate the accuracy and the computational efficiency of Point-Collocation NIPC method with different sampling techniques and number of collocation points in estimating the statistics of $f(x_i) = f\left(x_i(\xi_i)\right)$.

Table 2. The exact and the Monte Carlo (MC) statistics obtained for the stochastic model problem with four uniform random variables. The 95% confidence intervals for the MC statistics were calculated using the Bootstrap method.

Statistics of $f(x_1, x_2, x_3, x_4)$	Exact	MC	95% CI of MC
Mean	12.36096638	12.34363325	[12.30993751, 12.37955764]
Standard Deviation	6.155660136	6.141888533	[6.101064788, 6.187226973]

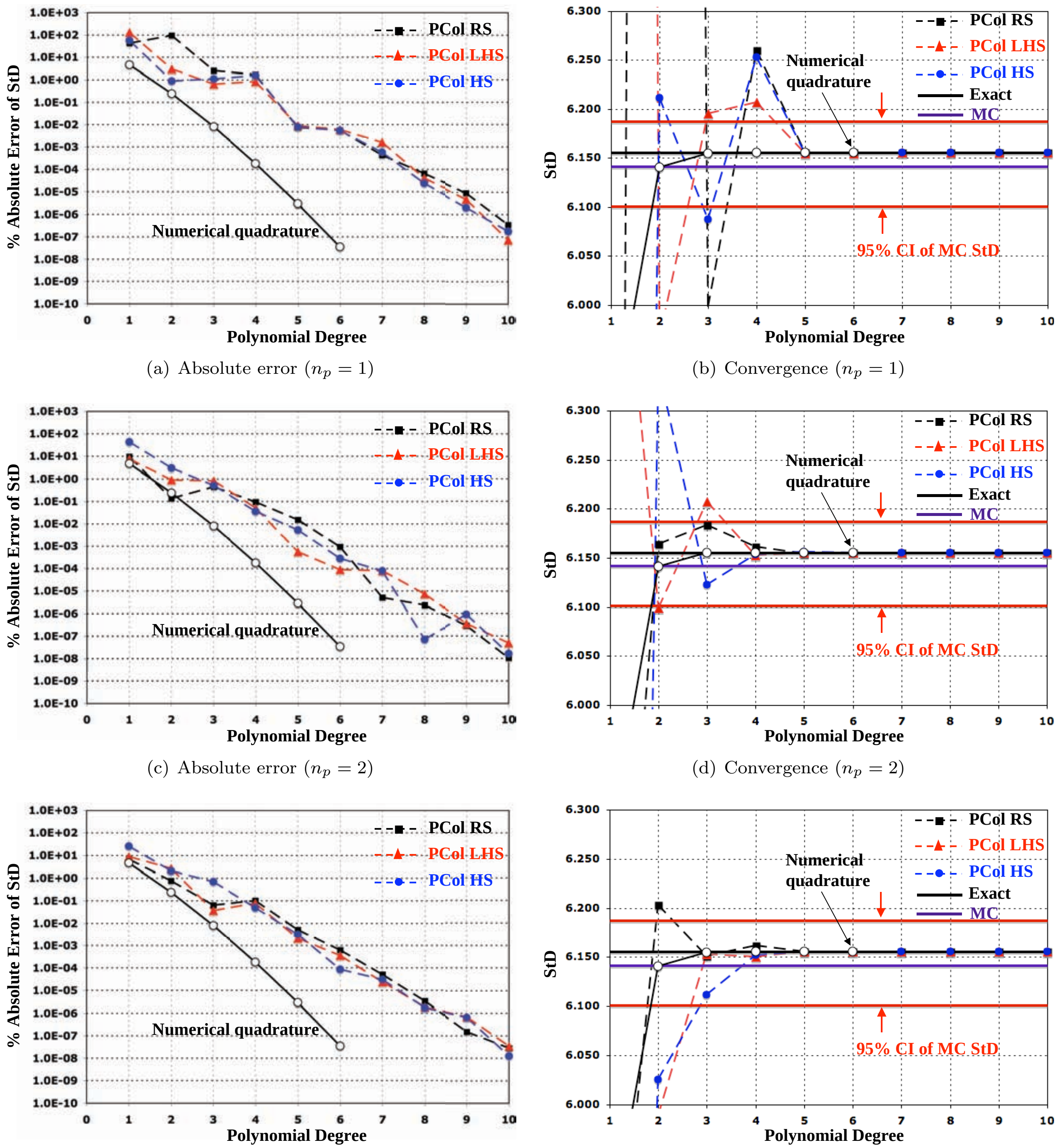
The statistics of this model problem were also calculated using crude MC simulations with 100,000 samples. Table 2 gives the MC statistics, the associated 95% confidence intervals obtained with the Bootstrap



(a) Absolute error (RS)
(b) Convergence (RS)
(c) Absolute error (LHS)
(d) Convergence (LHS)
(e) Absolute error (HS)
(f) Convergence (HS)

Figure 7. Absolute error and convergence of the standard deviation (StD) obtained with Point-Collocation (PCol) NIPC for the stochastic model problem with four uniform random variables. Each figure shows the results obtained with a specific sampling technique (RS, LHS, and HS).

method, and the exact solution to the statistics. The histogram of $f(x_i)$ can be seen in Figure 9(c). The distribution is bounded, skewed to right, and different than the input uniform distributions due to the exponential functional dependency. Compared to the output histogram obtained for the first model problem, the shape of the current output distribution can be considered as less complex. In a similar approach performed for the first model problem, Figures 7 and 8 may be used for this problem to analyze the convergence of standard deviation as a function of the chaos order for various number of collocation points



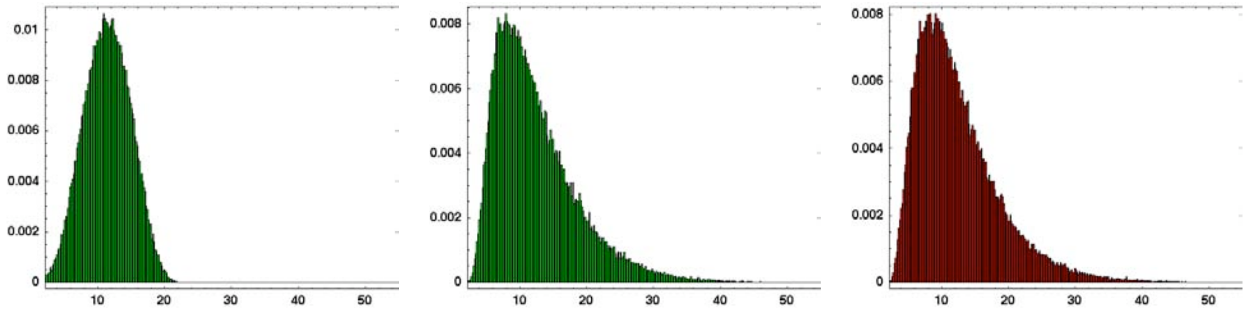
(a) Absolute error ($n_p = 1$) (b) Convergence ($n_p = 1$)

(c) Absolute error ($n_p = 2$) (d) Convergence ($n_p = 2$)

(e) Absolute error ($n_p = 3$) (f) Convergence ($n_p = 3$)

Figure 8. Absolute error and convergence of the standard deviation (StD) obtained with Point-Collocation (PCol) NIPC for the stochastic model problem with two uniform random variables. Each figure shows the results obtained with a specific value of n_p ($n_p = 1$, 2 and 3).

and sampling techniques. As also observed in the first model problem, the selection of a specific sampling technique does not seem to change the accuracy of the polynomial chaos expansions. However, the NIPC expansions obtained with the LHS and HS exhibit more smooth (monotonic) convergence compared to the cases with random collocation points, especially for $n_p > 1$. Again, using a number of collocation points that is twice more than the minimum number required ($n_p = 2$) gives a better approximation to the statistics at each polynomial degree. For this stochastic model problem with four random variables, a reduction in the



(a) PCol NIPC, polynomial degree=1 (b) PCol NIPC, polynomial degree=3 (c) Monte Carlo with 100,000 samples

Figure 9. The histograms of $f(x_i)$ $(i = 1, .., 4)$ obtained with Monte Carlo simulations and with the Point-Collocation (PCol) NIPC (HS and $n_p = 2$) for chaos orders of 1 and 3.

number of function evaluations can be observed when $n_p > 1$ is chosen over $n_p = 1$ (Table 3). For example, the standard deviation of the Point-Collocation NIPC obtained with HS falls within the 95% CI of MC at a polynomial degree of 5 for $n_p = 1$, while it enters the CI at a polynomial degree of 3 for $n_p = 2$. The corresponding number of function evaluations is 126 and 70, which indicates a significant improvement in computational efficiency. This example points out the fact that for problems with multiple random variables, improving the accuracy of polynomial chaos coefficients in NIPC approaches may reduce the computational expense via achieving the same accuracy level with a lower order polynomial expansion.

In Figures 7 and 8, the standard deviation values calculated with the numerical quadrature up to a polynomial degree of 6 are also included. For this problem, 6 quadrature points were used in each random dimension. The absolute error of the numerical quadrature is again lower compared to Point-Collocation methods for each polynomial degree, however the associated computational cost (number of function evaluations= 6^4) is large due to the increase in random dimensions.

The histograms of $f(x_i)$ obtained with Point-Collocation NIPC (based on HS and $n_p = 2$) for a polynomial degree of 1 and 3 are shown in Figure 9. The probability distribution of the polynomial chaos expansion converges to the Monte Carlo distribution with a polynomial degree of 3 at which the standard deviation of the same NIPC method falls within the 95% CI of the MC statistics (Figure 8).

Table 3. The degree of the polynomial at which the standard deviation of Point-Collocation NIPC falls within the 95% CI of MC for different values of n_p and sampling techniques (LHS and HS). The number of function evaluations are given in parentheses for each case.

	$n_p = 1$	$n_p = 2$	$n_p = 3$	$n_p = 4$
LHS	5 (126)	4 (140)	3 (105)	2 (60)
HS	5 (126)	3 (70)	3 (105)	3 (140)

C. Stochastic Aerodynamics Problem with a Three-Dimensional Wing

To demonstrate the application of Point-Collocation NIPC to a aerospace problem with multiple uncertain input variables, a stochastic computational aerodynamics problem which includes the numerical simulation of steady, inviscid, transonic flow over a three-dimensional wing has been selected. The wing geometry is the AGARD 445.6 Aeroelastic Wing[19] (Figure 10) which has been extensively used to validate computational aeroelasticity tools especially in the determination of flutter boundary at various transonic Mach numbers. The wing has a quarter-chord sweep angle of 45 deg., a panel aspect ratio of 1.65, a taper ratio of 0.66, and a NACA 65A004 airfoil section. The current study includes the aerodynamic simulations with the rigid wing assumption which can be thought as a preliminary investigation before the application of the NIPC to a stochastic computational aeroelasticity problem involving the same geometry.

For the CFD simulations, we have used CFL3D code of NASA Langley Research Center to solve the steady Euler equations numerically. CFL3D is a three-dimensional, finite-volume Navier-Stokes code capable of solving steady or time-dependent aerodynamic flows ranging from subsonic to supersonic speeds.[20] The computational grid has a C-H topology having $193 \times 65 \times 42$ points.

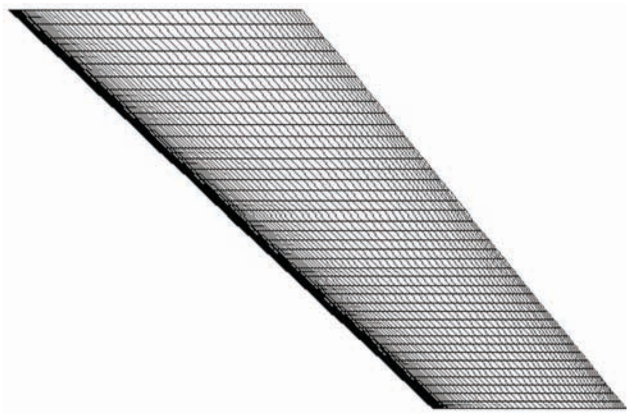
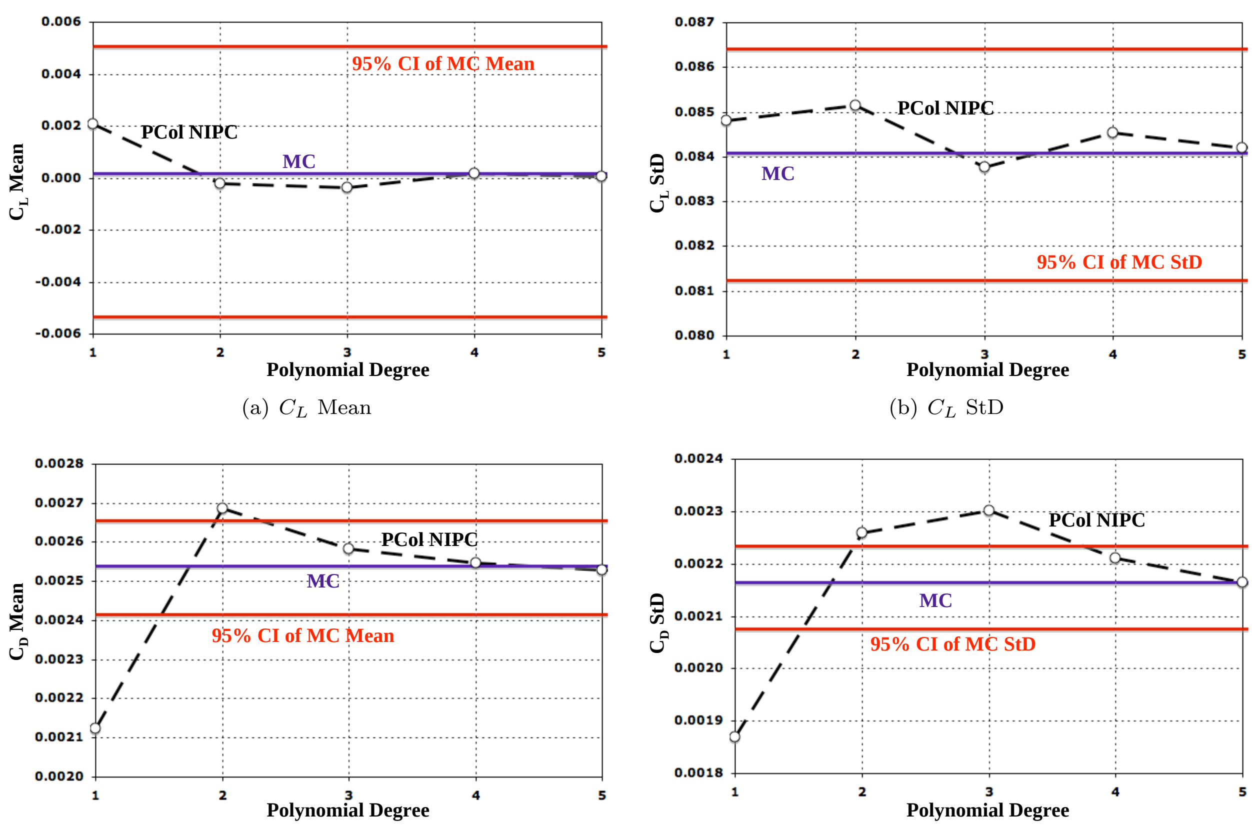


Figure 10. AGARD 445.6 Wing and the surface grid

For the stochastic aerodynamics problem, the free-stream Mach number (M_∞) and the angle of attack (α) are treated as uncertain variables. The Mach number is modeled as a uniform random variable between $M_\infty = 0.8$ and 1.1, and the angle of attack is defined as a uniform random variable between $\alpha = -2.0$ and 2.0 degs. Stochastic solutions to the problem were obtained with two approaches: Latin Hypercube Monte Carlo with 1000 samples and Point-Collocation NIPC. Based on our observations from the stochastic model problems, Hammersley Sampling with $n_p = 2$ has been used for the selection of collocation points for the NIPC approach. The chaos expansions were obtained up to a polynomial degree of 5. Table 4 gives the computational cost associated with each stochastic approach. As can be seen from this table, it took 46.6 hours to finish the Monte Carlo simulations whereas the computational time was approximately 2 hours for the Point Collocation NIPC with a polynomial degree of 5.

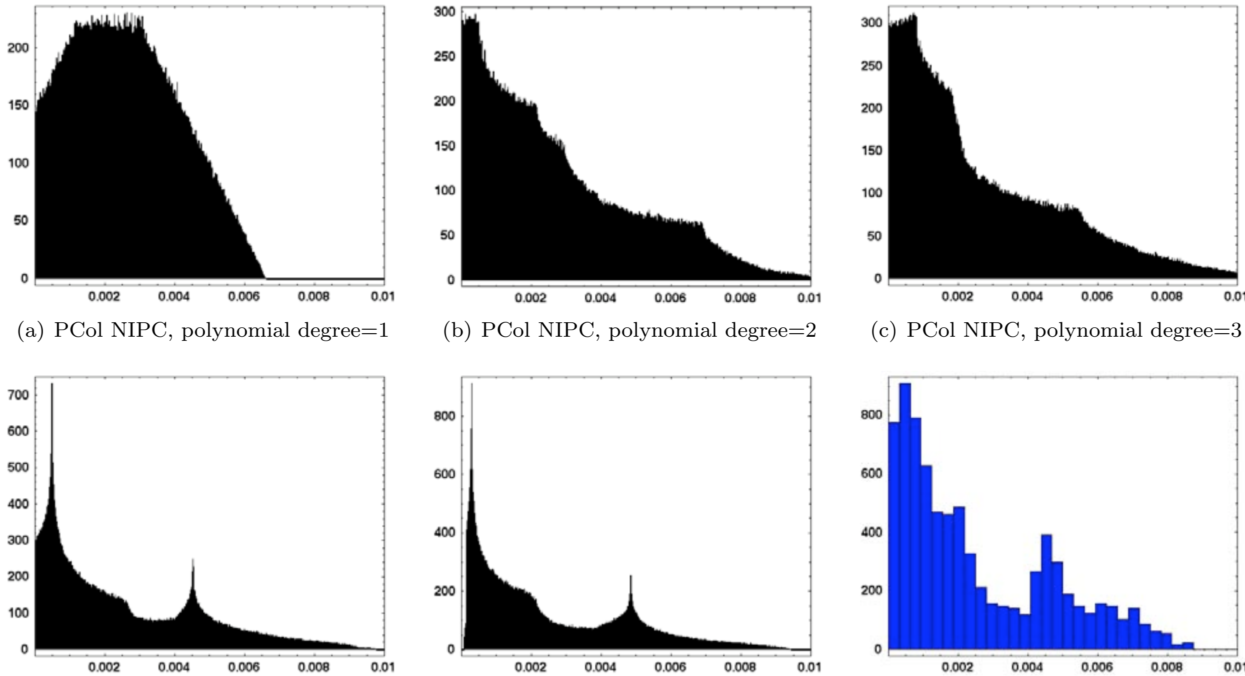
Table 5 gives the statistics and the 95% confidence intervals of the lift (C_L) and drag (C_D) coefficients obtained with the Latin Hypercube Monte Carlo simulations. In figure 11, the convergence of C_L and



(a) C_L Mean (b) C_L StD

(c) C_D Mean (d) C_D StD

Figure 11. Mean and standard deviation of C_L and C_D obtained with Point-Collocation NIPC and Latin Hypercube MC



(a) PCol NIPC, polynomial degree=1 (b) PCol NIPC, polynomial degree=2 (c) PCol NIPC, polynomial degree=3

(d) PCol NIPC, polynomial degree=4 (e) PCol NIPC, polynomial degree=5 (f) Latin Hypercube MC

Figure 12. The histograms of C_D obtained with the Latin Hypercube Monte Carlo (MC) simulations and the Point-Collocation (PCol) NIPC.

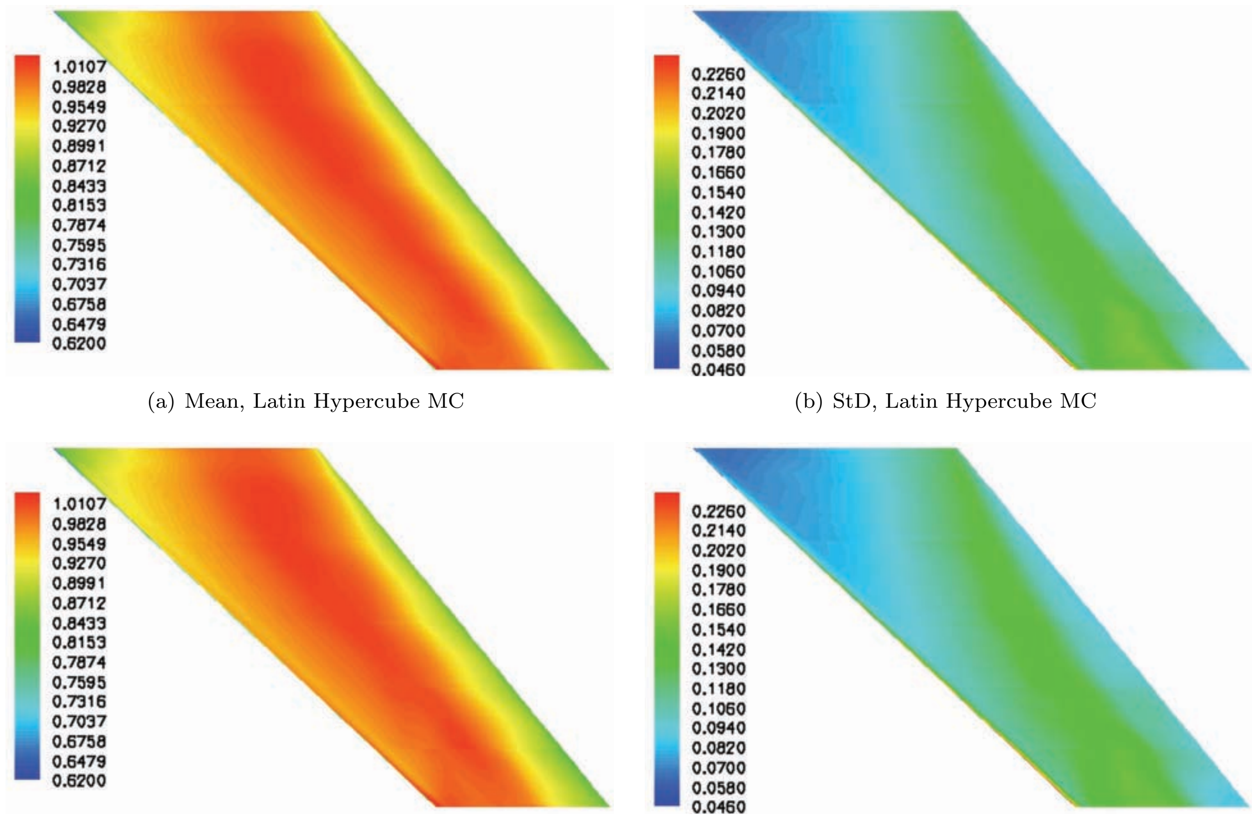
Table 4. The computational cost for evaluation of the Latin Hypercube Monte Carlo simulations and the Point-Collocation NIPC for the stochastic aerodynamics problem. The CFD runs were performed on a SGI Origin 3800 with 64 processors. (p is the degree of the polynomial chaos expansion).

		NIPC				
	Monte Carlo	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
number of CFD runs	1000	6	12	20	30	42
wall clock time (hours)	46.6	0.28	0.56	0.94	1.4	1.96

Table 5. The Latin Hypercube Monte Carlo statistics for C_L and C_D. The 95% confidence intervals for the MC statistics were calculated using the Bootstrap method.

	Mean	95% CI of Mean	StD	95% CI of StD
C_L	0.000169661	[-0.005324476, 0.005058219]	0.084082119	[0.081239769, 0.08641931]
C_D	0.002538528	[0.002414416, 0.002655811]	0.002164797	[0.002075901, 0.002234283]

C_D statistics obtained with the Point-Collocation NIPC is presented. For all polynomial degrees, the NIPC approximations to the mean and the standard deviation of the lift coefficient stays within the 95% confidence interval of the Monte Carlo statistics. The mean of the drag coefficient falls within the confidence interval at a polynomial degree of 3 whereas the standard deviation enters the interval with a polynomial degree of 4. The histogram of the drag coefficient obtained with the NIPC approach gets very similar to the histogram shape of the Monte Carlo simulations at a polynomial degree of 5 (Figure 12). The mean and the standard deviation distributions of the Mach number on the wing upper surface is given in Figure 13. The Monte Carlo and the NIPC statistics are in a good qualitative agreement in most regions of the wing.



(a) Mean, Latin Hypercube MC

(b) StD, Latin Hypercube MC

(c) Mean, PCol NIPC

(d) StD, PCol NIPC

Figure 13. Mean and standard deviation of Mach number on the upper surface of the AGARD 445.6 Wing. The Point-Collocation (PCol) NIPC results are obtained with 5th degree chaos expansions.

Overall, this computational exercise shows that a 5^{th} degree Point-Collocation NIPC expansion obtained with Hammersley sampling and $n_p = 2$ is capable of estimating the statistics at an accuracy level of 1000 Latin Hypercube Monte Carlo simulations with a significantly lower computational cost.

V. Conclusions

In this paper, we have addressed the accuracy and the computational efficiency of a Point-Collocation Non-Intrusive Polynomial Chaos (NIPC) method applied to various stochastic problems with multiple uncertain input variables. Two stochastic model problems were studied to determine the effect of different sampling methods (Random, Latin Hypercube, and Hammersley) for the selection of the collocation points. The first stochastic problem had two uniform random variables and the second problem included four uniform random variables. The effect of the number of collocation points on the accuracy of polynomial chaos expansions were also investigated.

The results of the stochastic model problems show that all three sampling methods exhibit a similar performance in terms of the the accuracy and the computational efficiency of the chaos expansions. However, the convergence of the Point-Collocation NIPC statistics obtained with Hammersley and Latin Hypercube sampling exhibits a much smoother (monotonic) convergence compared to the cases obtained with Random sampling. It has been observed that using a number of collocation points that is twice more than the minimum number required gives a better approximation to the statistics at each polynomial degree. This improvement can be related to the increase of the accuracy of the polynomial coefficients due to the use of more information (collocation points) in their calculation. The results of the stochastic model problems also indicate that for problems with multiple random variables, improving the accuracy of polynomial chaos coefficients in NIPC approaches may reduce the computational expense by achieving the same accuracy level with a lower order polynomial expansion.

To demonstrate the application of Point-Collocation NIPC to an aerospace problem with multiple uncertain input variables, a stochastic computational aerodynamics problem which includes the numerical simulation of steady, inviscid, transonic flow over a three-dimensional wing with an uncertain free-stream Mach number and angle of attack has been studied. In this study, for various output quantities of interest, it has been shown that a 5^{th} degree Point-Collocation NIPC expansion obtained with Hammersley sampling was capable of estimating the statistics at an accuracy level of 1000 Latin Hypercube Monte Carlo simulations with a significantly lower computational cost.

Acknowledgments

This research was supported by the National Institute of Aerospace (NIA) under grant # 415375. The authors would like to thank Dr. Robert Tolson at NIA for his guidance and support. The authors also would like to thank Dr. Walter Silva of NASA LaRC for his contribution to the transonic wing case.

References

[1]Ghanem, R. and Spanos, P. D., "Polynomial Chaos in Stochastic Finite Elements," *Journal of Applied Mechanics*, Vol. 57, March 1990, pp. 197–202.

[2]Ghanem, R., "Stochastic Finite Elements with Multiple Random Non-Gaussian Properties," *Journal of Engineering Mechanics*, January 1999, pp. 26–40.

[3]Ghanem, R. G., "Ingredients for a General Purpose Stochastic Finite Element Formulation," *Computational Methods in Applied Mechanical Engineering*, Vol. 168, 1999, pp. 19–34.

[4]L. Mathelin, M.Y. Hussaini, T. Z. and Bataille, F., "Uncertainty Propagation for Turbulent, Compressible Nozzle Flow Using Stochastic Methods," *AIAA Journal*, Vol. 42, No. 8, August 2004, pp. 1669–1676.

[5]Xiu, D. and Karniadakis, G. E., "Modeling Uncertainty in Flow Simulations via Generalized Polynomial Chaos," *Journal of Computational Physics*, Vol. 187, No. 1, May 2003, pp. 137–167.

[6]Wiener, N., "The Homogeneous Chaos," *American Journal of Mathematics*, Vol. 60, No. 4, 1938, pp. 897–936.

[7]Walters, R., "Towards stochastic fluid mechanics via Polynomial Choas-invited, AIAA-Paper 2003-0413," 41^{st} *AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, January, 2003, CD-ROM.

[8]Debusschere, B. J., Najm, H. N., Pebay, P. P., Knio, O. M., Ghanem, R. G., and Maitre, O. P. L., "Numerical Challenges in the Use of Polynomial Chaos Representations for Stochastic Processes," *SIAM Journal on Scientific Computing*, Vol. 26, No. 2, 2004, pp. 698–719.

[9]Reagan, M., Najm, H. N., Ghanem, R. G., and Knio, O. M., "Uncertainty Quantification in Reacting Flow Simulations through Non-Intrusive Spectral Projection," *Combustion and Flame*, Vol. 132, 2003, pp. 545–555.

[10]Isukapalli, S. S., "Uncertainty Analysis of Transport-Transformation Models, PhD Dissertation," Tech. rep., Rutgers, The State University of New Jersey, New Braunswick, NJ, 1999.

[11]L. Mathelin, M.Y. Hussaini, T. Z., "Stochastic Approaches to Uncertainty Quantification in CFD Simulations," *Numerical Algorithms*, Vol. 38, No. 1, March 2005, pp. 209–236.

[12]Hosder, S., Walters, R., and Perez, R., "A Non-Intrusive Polynomial Chaos Method For Uncertainty Propagation in CFD Simulations, AIAA-Paper 2006-891," 44^{th} *AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, January, 2006, CD-ROM.

[13]Loeven, G. J. A., Witteveen, J. A. S., and Bijl, H., "Probabilistic Collocation: An Efficient Non-Intrusive Approach for Arbitrarily Distributed Parametric Uncertainties, AIAA-Paper 2007-317," 45^{th} *AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, January, 2007, CD-ROM.

[14]Wiener, N. and Wintner, A., "The Discrete Chaos," *American Journal of Mathematics*, Vol. 65, No. 2, 1943, pp. 279–298.

[15]Huyse, L., Bonivtch, A. R., Pleming, J. B., Riha, D. S., Waldhart, C., and Thacker, B. H., "Verification of Stochastic Solutions Using Polynomial Chaos Expansions, AIAA-Paper 2006-1994," 47^{th} *AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Newport, RI, May, 2006, CD-ROM.

[16]Walters, R. W. and Huyse, L., "Uncertainty Analysis for Fluid Mechanics with Applications," Tech. rep., ICASE 2002-1, NASA/CR-2002-211449, NASA Langley Research Center, Hampton, VA, 2002.

[17]Giunta, A. A., Jr., S. F. W., and Eldred, M. S., "Overview of Modern Design of Experiments Methods for Computational Simulations, AIAA-Paper 2003-649," 41^{st} *AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, January, 2003, CD-ROM.

[18]Hammersley, J. M. and Handscomb, D. C., *Monte Carlo Methods*, Methuen's Monographs on Applied Probability and Statistics, Fletcher & Son Ltd., Norwich, 1964.

[19]"AGARD Standard Aeroelastic Configurations for Dynamic Response I - Wing 445.6," Tech. rep., NASA TM-100492, 1987.

[20]Krist, S. L., Biedron, R. T., and Rumsey, C. L., "CFL3D User's Manual (Version 5.0)," Tech. rep., NASA TM-1998-208444, NASA Langley Research Center, Hampton, VA, 1998.