Structural Integrity Prediction via Stochastic Local Regression

Seung-Kyum Choi
Systems Realization Laboratory
G. W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Savannah, GA, 31407
email: schoi@me.gatech.edu

Abstract: A primary challenge of stochastic analysis is to discover rigorous ways to forecast the low probability of failure which is critical to reliability constraints. In this paper, a new framework is proposed for the accurate estimation of the low failure probability. Combining the excellent advantages of the polynomial chaos expansion, and local regression method will result in a new simulation-based modeling technique that enables the accuracy of the structural integrity prediction. The proposed procedure can allow for realistic modeling of sophisticated statistical variations and facilitate in order to achieve improved reliability by eliminating unnecessary conservative approximations. An example problem is depicted to illustrate how the method is used to provide a quantitative basis for developing robust designs associated with the low probability of failure.

Keywords: Polynomial Chaos Expansion, Moving Least-Squares, Local Regression, Low Failure Probability

1. Introduction

In recent years, the rapid development and improvement of novel design concepts, especially utilizing novel material systems, is a major request of the aerospace and automobile industry. In addition, new digital and information science technologies are creating the potential for new high-level design fields, such as micro-electro-mechanical systems (MEMS) and multi-scale engineering systems. However, introducing this state-of-the-art technology and new material systems is rapidly increasing the complexity of most engineered systems. There exist significant difficulties in anticipating, understanding, designing, and controlling both normal and abnormal behaviors of the complex systems. In addition, uncertainties in material properties, geometry, manufacturing processes, and operational environments of the complex engineered systems are clearly critical at all scales (nano-, micro-, meso-, and macro-scale). For example, the typical tolerances of geometric accuracy and surface finish are on the order of tenths of microns during the fabrication processes (Maluf, 2004), and the common microfabrication material (i.e.

polycrystalline silicon) has 9~15% variation in its Young's modulus and tensile strength (Sharpe, Turner, and Edwards, 1999).

To compensate the ignorance of uncertainties in input parameters, safety factors have traditionally been incorporated approximately in engineering designs. Generally, the factor of safety is understood to be the ratio of the expected strength to response to the expected load (Choi, Grandhi, and Canfield, 2006). In practice, both the strength and load are variables, the values of which are scattered about their respective mean values. When the scatter in the variables is considered, the factor of safety could potentially be less than unity, and the traditional factor of safety-based design would fail. More likely, the factor of safety is too conservative, leading to an overly expensive design for a given level of safety. Probabilistic methods are convenient tools to describe or model physical phenomena that are too complex to treat with the present level of scientific knowledge. The probabilistic method explicitly incorporates given statistical data into the design algorithms and provides safer designs at given cost, whereas conventional deterministic design with the safety factor discards such data. However, the probabilistic-based approach often requires repeated evaluations of the probability of failure and it induces the computational challenge associated with the large number of computer simulations when the system requires extremely low failure probability, such as 10^{-5}~10^{-7}.

A common approach to the computationally-expensive procedure of the probabilistic methods is to approximate the system response using relatively inexpensive surrogate modeling techniques. In the approximation of the response function, the accuracy depends on the choice of the basis function and the sampling method including the choice of the sampling region and the position of the sampling points. An effective choice of the basis function for the uncertainty analysis is the direct use of stochastic expansions, i.e. Polynomial Chaos Expansion (PCE) (Ghanem and Spanos, 1991), since the stochastic expansions provide analytically appealing convergence properties based on the concept of a random process. The PCE can reduce computational effort of uncertainty quantification in engineering design applications where the system response is computed implicitly. Choi et al. (2006) recently developed an uncertainty analysis framework which can account for nonlinear fluctuations of large-scale system responses by integrating the PCE, the Karhunen-Loeve (KL) transform, and Latin Hypercube Sampling (LHS). This research utilized the stochastic expansion and the dimension reduction procedure to generate the random field and showed the applicability of the method to the complex engineered systems.

The objective of the current study is to provide the accurate estimation of the low failure probability of complex engineered systems by utilizing efficient probabilistic methods which can realistically model complicated statistical variations. To achieve a high quality surrogate model, a local regression method, namely Moving Least-Squares (MLS) method (Lancaster and Salkauskas, 1981), is integrated to a previously developed probabilistic decision support

framework (Choi, Canfield, and Grandhi, 2006). The main advantage of the MLS method is that the regression coefficients are not constant, but rather parameter dependent. This quality allows the data analysis to not be constrained to a specific global function in order to fit a model to the data. Instead, the fitting segments spawn a local-global approximation allowing the data to acclimate to the function over a wide range of parameters. The stochastic modeling process repeats and recalibrates the PCE model with the local regression scheme until sufficient model adequacies are achieved. This will allow for an accurate estimation of the low probability of failure with limited sampling points. The following sections provide a brief description and main ideas behind the local regression method and then focus on the technical details integrating the stochastic approximation procedure to provide the accuracy of the structural integrity prediction of complex engineered systems.

2. Mathematical Basis for Solution Concept

2.1. LOCAL REGRESSION

The efficacy of local regression schemes such as MLS method, lazy learning method, and locally weighted regression method have been successfully shown in recent engineering applications (Lancaster and Salkauskas, 1981; Stone, 1977; Cleveland, 1979; Katkovnik, 1979; Toropov, Scharamm, Sahai, Jones, and Zeguer, 2005). The basic idea of the local regression is to fit curves and surfaces to localized subsets of the data by a multivariate smoothing procedure with moving processes. The detailed steps of the MLS approximation are described in Figure 1. First, we define a local domain based on the domain influence factor or bandwidth, r. In the second step, we construct an approximation at a calculation point, x_i. These procedures can be repeated to each different calculation point by moving the local domain. Therefore, the regression coefficients of the MLS are not constant but a function of the calculation position or location. The “moving” process is analogous to a weighted moving average method, which is a common method in a time series analysis. In fact, applying zero degree polynomials in the local regression yields a weighted moving average. The advantage of the local approximation compared to the classical global fitting methods is that the method does not require a global function of any form to fit a given model and can generate accurate and smooth fitting of nonlinear responses without significant distortions.

Consider the linear regression model

$$y(x) = \beta_0 + \beta_1 p_1(x) + \ldots + \beta_m p_m(x) + \varepsilon \tag{1}$$

where $p_j(x)$, j = 0,1,2,…,m, are the basis polynomial of order m, β_j are the regression coefficients, and ε , the error of the model equation, is assumed to be normally distributed with mean zero and variance σ_e^2 .

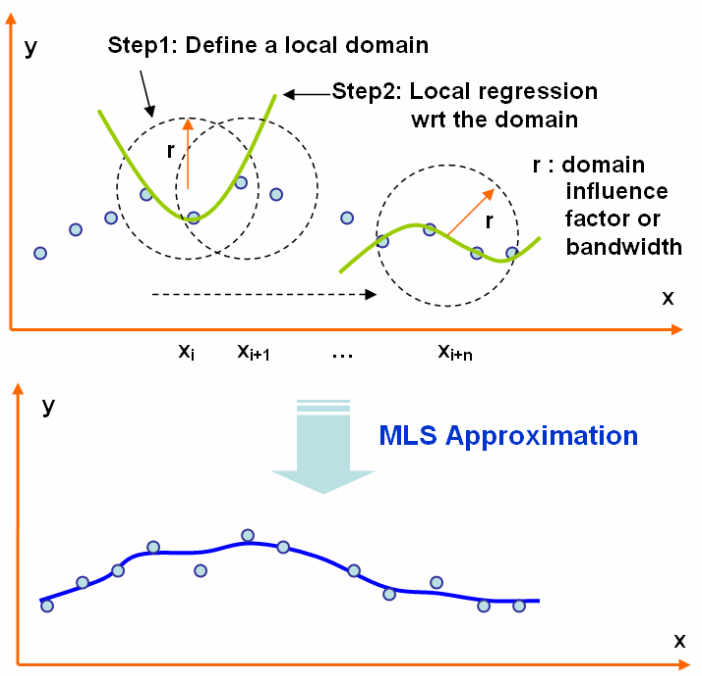


Figure 1. Moving Least-Squares Approximation

Equation (1) can be written in matrix notation for n sample values of x and y as

$$Y = X\hat{\beta} + e \tag{2}$$

where

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & p_1(x_1) & p_2(x_1) \ldots & p_k(x_1) \\ 1 & p_1(x_2) & p_2(x_2) \ldots & p_k(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & p_1(x_n) & p_2(x_n) \ldots & p_k(x_n) \end{bmatrix} \quad \hat{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \textit{and} \quad e = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Here, the simplest polynomial model is the monomials of x^m, i.e., $p^T(x) = [1, x, x^2, ..., x^m]$ and in 2D space, $p^T(x,z) = [1, x, z, x^2, xz, z^2, ..., x^m, z^m]$.

The least-squares procedure results in obtaining the regression coefficients:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \tag{3}$$

The fitted model and the residuals are

$$\hat{Y} = X\hat{\beta} \text{ and } e = Y - \hat{Y} \tag{4}$$

In the method of the Moving Least-Squares (MLS) approximation, the regression coefficient vector, $b(x)$, can be calculated as,

$$b(x) = [X^T W(x) X]^{-1} X^T W(x) Y \tag{5}$$

where X is a n x p matrix of the levels of the regressor variables, Y is a n x 1 vector of the responses, and $W(x)$ is a weight matrix and it is a none zero diagonal matrix:

$$W(x) = \begin{bmatrix} w_1(x - x_1) & 0 & \dots & 0 \\ 0 & w_2(x - x_2) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & w_n(x - x_n) \end{bmatrix} \tag{6}$$

Consequently, the model Y in Eq. (2) can be approximated by MLS approximants $u^h(x)$ as follows

$$u^h(x) = \sum_{j=0}^{m} p_j(x) b_j(x) = p^T(x) b(x) \tag{7}$$

The weight matrix, Eq.(6), is a function of the location or position of x and there are several types of weighting functions:

(a) Exponential weight function

$$w_i(x-x_i) = w(d_i) = \begin{cases} \exp(-(d_i/r_i)^2), & if \quad d_i/r_i \le 1 \\ 0, & if \quad d_i/r_i > 1 \end{cases} \tag{8a}$$

(b) Conical weight function

$$w(d_i) = \begin{cases} 1-(d_i/r_i)^2, & if \quad d_i/r_i \le 1 \\ 0, & if \quad d_i/r_i > 1 \end{cases} \tag{8b}$$

(c) Spline weight function

$$w(d_i) = \begin{cases} 1-6(d_i/r_i)^2+8(d_i/r_i)^3-3(d_i/r_i)^4, & if \quad d_i/r_i \le 1 \\ 0, & if \quad d_i/r_i > 1 \end{cases} \tag{8c}$$

where $d_i = \|x - x_i\|$ is the distance from the sample point x_i to x, and the domain influence factor, r_i, is directly related to the smoothing length; namely, the size of the support for the weight function. It is also called the bandwidth.

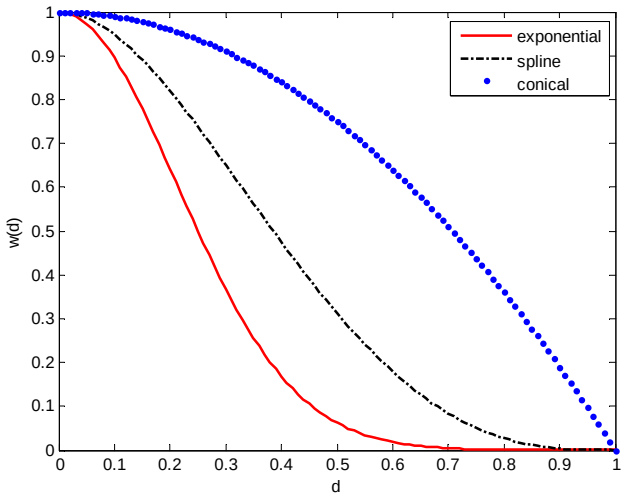


Figure 2. Weight Functions

Figure 2 depicts the three types of the weight functions discussed in this section. It is important to note that the shape of the fitted curve is not critically sensitive to the precise selection of the

weight function. However, the careful adjustment of the domain influence factor of the weight function is critical so that the interval should contain enough data points to obtain the regression coefficients. Otherwise, the regression procedure will envisage a singular matrix. The additional discussion on the effects of several weighting functions and the resulting local approximation can be found in Ref. (Dolbow and Belytschko, 1998).

2.2. STOCHASTIC APPROXIMATION

The Polynomial Chaos Expansion (PCE) stemmed from both Wiener and Ito's work on mathematical descriptions of irregularities (Wiener, 1938). A simple definition of the PCE for a Gaussian random response, $u(\theta)$, as a convergent series is as follows:

$$u(\theta) = a_0\Gamma_0 + \sum_{i=1}^{\infty} a_{i1}\Gamma_1(\xi_{i1}(\theta)) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1} a_{i_1 i_2}\Gamma_2(\xi_{i1}(\theta),\xi_{i2}(\theta)) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3}\Gamma_3(\xi_{i1}(\theta),\xi_{i2}(\theta),\xi_{i3}(\theta)) + \ldots \quad (9)$$

where $\{\xi_i(\theta)\}_{i=1}^{\infty}$ is a set of Gaussian random variables, $\Gamma_p(\xi_{i1},...,\xi_{ip})$ is the generic element of a set of multidimensional Hermite polynomials, usually called homogeneous chaos of order *p*, $a_{i1},...,a_{ip}$ are deterministic constants, and θ represents an outcome in the space of possible outcomes of a random event.

Equation (9) can be written more simply as

$$u(\theta) = \sum_{i=0}^{P} b_i \Psi_i(\vec{\xi}(\theta)) \quad (10)$$

where b_i and $\Psi_i(\vec{\xi}(\theta))$ are one-to-one correspondences between the coefficients $a_{i1},.....,a_{ip}$ and the functions $\Gamma_p(\xi_{i1},.....,\xi_{ip})$, respectively. If *u* is a function of a normally distributed random variable *x*, which has the known mean μ_x and variance σ_x^2, ξ is a normalized variable: $\xi = (x-\mu_x)/\sigma_x$. For example, the two-dimensional case of Eq. (9) can be expanded as:

$$u(\theta) = a_0\Gamma_0 + a_1\Gamma_1(\xi_1) + a_2\Gamma_1(\xi_2) + a_{11}\Gamma_2(\xi_1,\xi_1) + a_{12}\Gamma_2(\xi_2,\xi_1) + a_{22}\Gamma_2(\xi_2,\xi_2)$$

$$+ a_{111}\Gamma_3(\xi_1,\xi_1,\xi_1) + a_{211}\Gamma_3(\xi_2,\xi_1,\xi_1) + a_{221}\Gamma_3(\xi_2,\xi_2,\xi_1) + a_{222}\Gamma_3(\xi_2,\xi_2,\xi_2)\ldots \tag{11}$$

Equation (11) can be recast in terms of $\Psi_i[.]$ and b_i as follows:

$$u(\theta) = b_0\Psi_0 + b_1\Psi_1 + b_2\Psi_2 + b_3\Psi_3 + b_4\Psi_4 + b_5\Psi_5 + \ldots \tag{12}$$

Thus, the term $a_{11}\Gamma_2(\xi_1,\xi_1)$ becomes $b_3\Psi_3$ for this two-dimensional case.

The general expression to obtain the multidimensional Hermite polynomials is given by (Ghanem, and Spanos, 1991)

$$\Gamma_p(\xi_{i1},.....,\xi_{ip}) = (-1)^n \frac{\partial^n e^{-\frac{1}{2}\vec{\xi}^T\vec{\xi}}}{\partial\xi_{i1},.....,\partial\xi_{ip}} e^{\frac{1}{2}\vec{\xi}^T\vec{\xi}} \tag{13}$$

where the vector $\vec{\xi}$ consists of *n* Gaussian random variables $(\xi_{i1},.....,\xi_{in})$. Generally, the one-dimensional Hermite polynomials are defined by

$$\Psi_n(\xi) = (-1)^n \frac{\varphi^{(n)}(\xi)}{\varphi(\xi)} \tag{14}$$

where $\varphi^{(n)}(\xi)$ is the n^{th} derivative of the normal density function, $\varphi(\xi) = 1/\sqrt{2\pi}\ e^{-\xi^2/2}$. This is simply the single-variable version of Eq. (13). From Eq. (14), we can readily find

$$\{\Psi_i\} = \{1,\ \xi,\ \xi^2-1,\ \xi^3-3\xi,\ \xi^4-6\xi^2+3,\ \xi^5-10\xi^3+15\xi,\ldots\} \tag{15}$$

Thus, a second order, 2-D PCE is given by

$$u(\theta) = b_0 + b_1\xi_1(\theta) + b_2\xi_2(\theta) + b_3(\xi_1^2(\theta)-1) + b_4\xi_1(\theta)\xi_2(\theta) + b_5(\xi_2^2(\theta)-1) \tag{16}$$

where $\xi_1(\theta)$ and $\xi_2(\theta)$ are two independent random variables.

PCE can be used to represent the response of an uncertain system in the non-intrusive formulation (Pettit, Canfield, and Ghanem, 2002; Choi, Grandhi, Canfield, and Pettit, 2004). The basic idea of this approach is to project the response and stochastic system operator onto the stochastic space spanned by PCE with the projection coefficients, b_i, being evaluated through an efficient

sampling scheme. We first define the vector x at a particular point $(\xi_1, \xi_2, ..., \xi_m)$ of random variables

$$x^T = [1 \ \Psi_1(\xi_1) \ \Psi_2(\xi_1) ... \Psi_p(\xi_1) \ \ \Psi_1(\xi_2) \ \Psi_2(\xi_2) ... \Psi_p(\xi_2) \ \ \Psi_1(\xi_m) \ \Psi_2(\xi_m) ... \Psi_p(\xi_m)] \tag{17}$$

where p is the order of polynomial and $\Psi_j(\xi_i)$ are PCE. The estimated response at this point is

$$y(x) = x^T \hat{\beta} \tag{18}$$

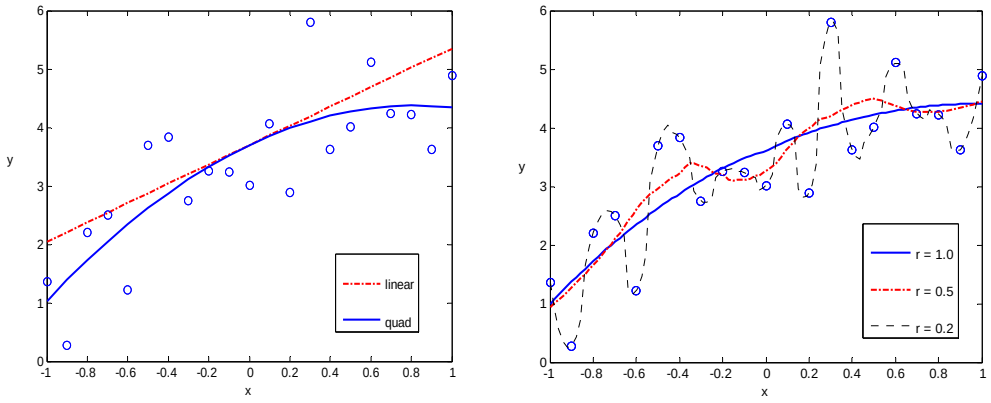
where $\hat{\beta}$ is a set of undetermined coefficients of PCE and it can be obtained from Equation (5).

2.3. SOLUTION STRATEGIES

For the utilization of the local regression method in practice, the selection of its basic components, such as the basis function, the weighting function, and the domain influence factor, *r*, is critical to provide the reliable model adequacy of the approximation. For instance, the domain influence factor has a significant effect on the fitted shape. Depending upon the size of the domain influence factor or the bandwidth, the user can adjust the closeness of fit, and this flexibility can also enable the user to achieve the same result of the interpolation and the global regression as shown in Figure 3b. Further discussions on the fixed bandwidth and nearest-neighbor bandwidth selection in the local regression are available in Refs. (Cleveland, 1979; Katkovnik, 1979). Figure 3a shows the fitted model for the same data by using the global regression method. In the case of the global regression, the data analysis is constrained to a specific global function to fit a model data. It is clear that the local regression method provides sufficient flexibility to achieve good model adequacy. However, when the size of the domain influence factor is small, the obtained response approximation can be unstable against the effects of random fluctuations, or noise phenomena. Therefore, it is important to develop criterions for the selection of the basic components of the local regression method.

In this study, the PCE is employed as a basis function. More satisfactory solutions can be expected because of the orthogonal property of the PCE. For the common polynomial regression model of the monomials of x^m, the columns of X in Eq. (2) can sometimes be nearly collinear, which causes an ill-conditioned problem, because negative values of x produce negative values for all odd powers, and positive values of x produce large positive values for all of the function. Hence, small changes in the basis function lead to relatively large changes in the regression coefficients. Another important issue with the polynomial regression is in determining an appropriate order of polynomials. By using a linear basis function (first-order polynomial) in the local regression often induces rapid changes in the slope. In the local regression method, increasing the degree of polynomials can typically enlarge the bandwidth without introducing intolerable bias; it eventually produces smoother fitting shapes compared to the linear basis

(Lancaster and Salkauskas, 1981; Stone, 1977; Cleveland, 1979). In order to determine the appropriate degree of the polynomials and the size of the domain influence factor, several possible criterions, which involve R^2, C_p statistics (Montgomery, 1997), and the graphical diagnostics, can be considered. The graphical diagnostics, such as the plot of residuals ε versus $\hat{y}$, or y versus $\hat{y}$, can provide a visual assessment of model effectiveness. The visual inspections of residuals are preferable to understand certain characteristics of the regression results because analysts can easily construct the plots and reveal useful information from the unorganized data. However, the visual inspection is a labor intensive process and it is difficult to automate. An advantage of the R^2 and C_p statistics is that the procedure can be automated. It does not require labor intensive processes. Since the automated procedure can underestimate a peak in a surface and sometimes produces a poor solution, an ideal criterion can be a cross-validation by using both the graphical diagnostics and R^2 or C_p statistics.



(a) Global Regression (b) Local Regression with r = 0.2, r = 5.0, and r = 1.0

Figure 3. Effect of Local Regression and Domain Influence Factor

Figure 4 shows the flowchart of the solution strategies for determining the appropriate parameters of the basic components in the MLS approximation. In this procedure, the utilization of the stratified sampling technique known as LHS is expected to decrease the number of simulations needed. To determine the parameters for the MLS approximation, the R^2 value has been checked along with the graphical diagnostics of the regression result as shown in Figure 4. The formula for R is defined by

$$R_{X_1,X_2} = \frac{Cov(X_1, X_2)}{\sigma_{X_1}\sigma_{X_2}} \tag{19}$$

where Cov(·) is the measure of correlation of the fluctuations of the two different quantities; namely, covariance and σ_{X_1} represents the standard deviations for X_1.

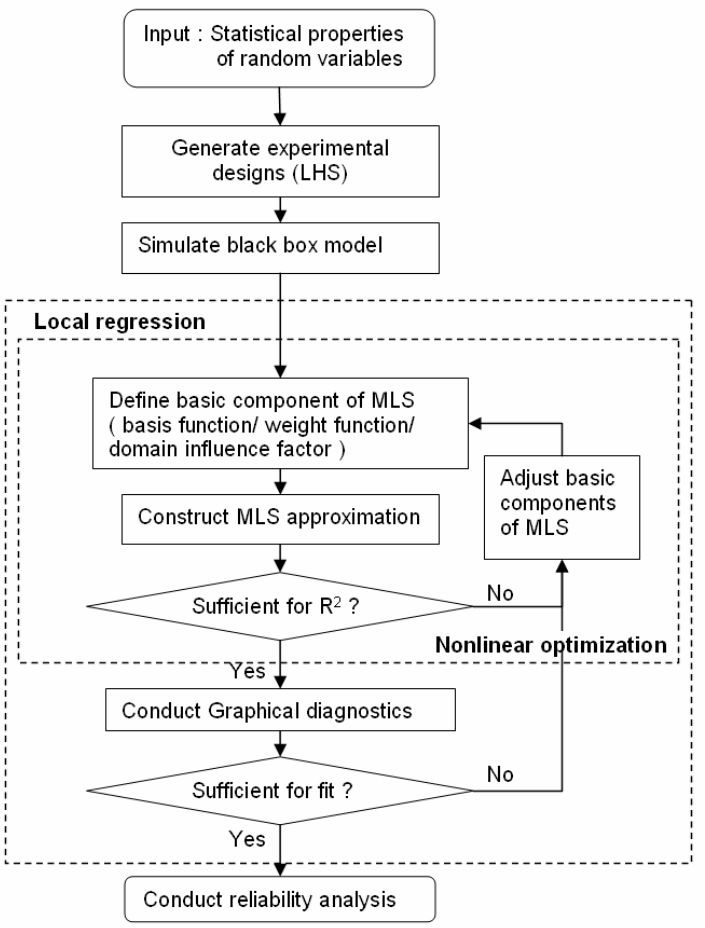


Figure 4. Solution Strategies for Local Regression

R^2 can vary from 0.0 to 1.0, where a R^2 value of 1.0 indicates the regression perfectly fits the data. R^2 is a good measure to automate the determining procedure of the basic components for the MLS approximation using nonlinear optimization. However, when R^2 is misused, the user can produce

an undesirable interpolation with very high order polynomial models. The introduction of the graphical diagnostics step, such as the residual analysis, can detect this undesirable and uncontrollable result with little additional effort. For instance, the abnormality of the residual plots indicates that the selected model is inadequate or that an error exists in the analysis. There are no significant computational costs to obtain statistical properties of the responses after constructing the PCE representation of stochastic responses.

3. Structural Integrity Prediction

3.1. THREE-BAR TRUSS EXAMPLE

Reliability analysis evaluates various statistical properties and the probability of system failure by determining whether the limit-state functions are exceeded. Generally, the limit state indicates the margin of safety between the resistance and the load of structures. The limit state function, $g(\cdot)$, and probability of failure, P_f, can be defined as

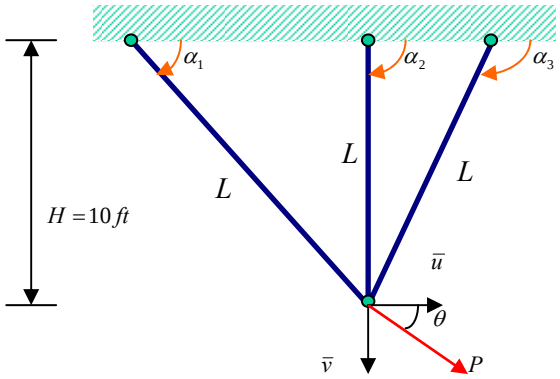
$$g(X) = R(X) - S(X)$$
$$P_f = P\ [g(\cdot) \le 0] \tag{20}$$

where R is the resistance and S is the loading of the system. Both $R(.)$ and $S(.)$ are functions of the random variable X. The notation $g(.) < 0$ denotes the failure region. Likewise, $g(.) = 0$ and $g(.) > 0$ indicate the failure surface and safe region, respectively.

In this section, the estimation of the low failure probability will be discussed by comparisons of a sampling method and the proposed method. An indeterminate, asymmetric system of a three pin-connected truss structure is illustrated in Figure 5. The unloaded length, L_m, and orientation, α_m, of each member are deterministic. Young's modulus, E_m, of each member is also assumed to be deterministic. The load has a random magnitude, P, and direction, θ. The cross-sectional area A for all members is also random. The random quantities are initially considered normally distributed and uncorrelated:

A~N(1 in^2, 0.1 in^2)
P~N(1000 *lb*, 250 *lb*)
θ ~N(45^o, 7.5^o)

where the symbol x ~ N(μ_x, σ_x) denotes that the random variable x is treated as a normal distribution and has the mean of μ_x and standard deviation of σ_x.



$$\alpha_1 = 45^\circ, \alpha_2 = 90^\circ, \alpha_3 = 110^\circ, E = 2.9\times10^6\ psi$$

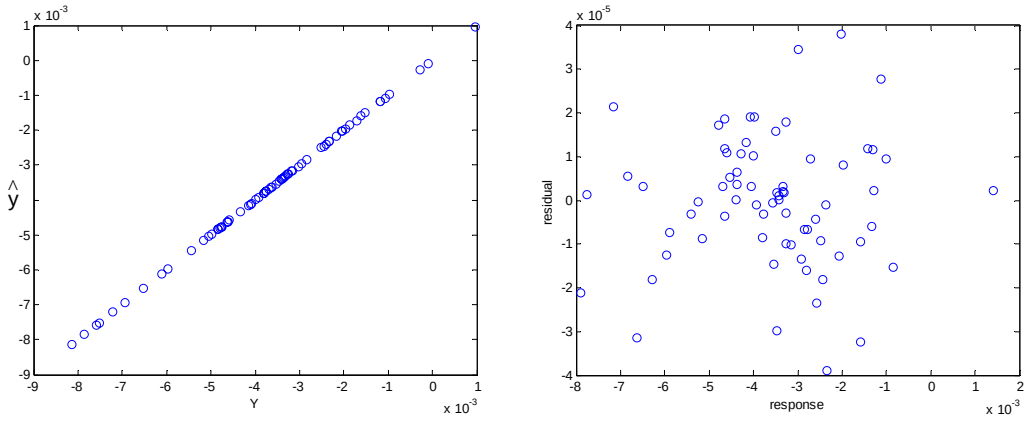
Figure 5. Pin-connected Three-bar Truss

The principle of virtual work is used to calculate the displacement vector $[\bar{u}, \bar{v}]^T$ of the joint at which the load is applied and is given by the solution of the following system of equations:

$$P\cos\theta = \sum_{m=1}^{3}(\bar{u}\cos^2\alpha_m + \bar{v}\cos\alpha_m\sin\alpha_m)\frac{E_m A_m}{L_m}$$
$$P\sin\theta = \sum_{m=1}^{3}(\bar{v}\sin^2\alpha_m + \bar{u}\cos\alpha_m\sin\alpha_m)\frac{E_m A_m}{L_m} \qquad (21)$$

The horizontal deflection of the structure should be $\bar{u}$ < 0.0015 in. This restriction is considered as a limit state. To obtain the probability of failure, P_f, one million simulations were conducted to obtain a converged result in MCS. 200 samples of LHS were used to obtain the surrogate model of the limit state by using the third-order PCE model with the exponential weight function of Eq. (8a). The plots of y versus $\hat{y}$ (Figure 6a) or residuals versus $\hat{y}$ (Figure 6b) provide a visual assessment of model effectiveness in regression analysis. Since the residual plot of Figure 6b exhibits white noise behavior which means there is no abnormality and the residual plot in Figure

6a yields points around the 45^o line, the estimated regression function shows accurate predictions of the values that are actually observed. Therefore, the selected PCE and the weight function model of MLS are sufficient for fitting the given data. After conducting the local regression, P_f is calculated using one million MCS simulations with the obtained PCE model.



(a) Residual Plot: y versus $\hat{y}$ (b) Residual Plot: $\hat{y}$ versus Residual

Figure 6. Residual Analysis

Table 1. Comparison of Methods for Reliability Analysis

	P_f	95% Confidence Interval
MCS	4.70×10^{-6}	$[3.64\times10^{-6}, 5.76\times10^{-6}]$
PCE+MLS	4.34×10^{-6}	$[4.01\times10^{-6}, 4.67\times10^{-6}]$

The corresponding results of the current example are summarized in Table 1. The PCE result converged to $P_f = 4.34\times10^{-6}$ and 95% confidence interval is also obtained. The confidence interval indicates a range of values that likely contains the analysis results. For this case, the user can be 95% confident that the true mean of P_f will be between 4.01×10^{-6} and 4.67×10^{-6}. The confidence interval of MCS is larger than the result of PCE, but it has an overlapping region with the PCE's. The interval can be reduces as the sampling size increases in the case of MCS. The

obtained result exhibits that the use of PCE along with MLS is applicable to the estimation of the low failure probability.

4. Summary

A new framework is proposed for the accurate estimation of the low failure probability of common engineering problems by utilizing efficient probabilistic methods which can realistically model complicated statistical variations. A local regression method, MLS, is integrated to a previously developed probabilistic decision support framework which combines the PCE and LHS. The stochastic modeling process repeats and recalibrates the PCE model with the local regression scheme until sufficient model adequacies are achieved. This allows for an accurate estimation of the low probability of failure with limited sampling sets. This increased capability has the potential to provide significant robust designs with a minimal amount of computational cost.

References

Choi, S., Canfield, R.A, and Grandhi, R.V., "Estimation of Structural Reliability for Gaussian Random Fields," *Structure & Infrastructure Engineering*, Vol. 2, No. 3-4., pp. 161-173, 2006.

Choi, S., Grandhi, R.V., and Canfield, R.A., "Robust Design of Mechanical Systems via Stochastic Expansion," International Journal of Materials and Product Technology, Vol. 25, pp. 127-143, 2006.

Choi, S., Grandhi, R.V., and Canfield, R.A., "Structural Reliability under non-Gaussian Stochastic Behavior," Computers and Structures, Vol. 82, pp.1113-1121, 2004.

Choi, S., Grandhi, R.V., and Canfield, R.A., *Reliability-Based Structural Design*, Springer, London, 2006.

Choi, S., Grandhi, R.V., Canfield, R.A. and Pettit, C.L., "Polynomial Chaos Expansion with Latin Hypercube Sampling for Estimating Response Variability," *AIAA Journal*, Vol. 42, No. 6, pp. 1191-1198, 2004.

Cleveland, W., "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, Vol. 74, pp. 829-836, 1979.

Cureton, E., and D'Agostinao, R., *Factor Analysis: An Applied Approach*, Lawrence Erlbaum Associates, NJ, 1983.

Dolbow J., and Belytschko T., "An Introduction to Programming the Meshless Element Free Galerkin Method," *Archives of Computational Methods in Engineering*, Vol. 5, No. 3, pp. 207-241, 1998.

Ghanem, R., and Spanos, P.D., *Stochastic Finite Elements: A Spectral Approach,* Springer-Verlag, NY, 1991.

Katkovnik,V., "Linear and Nonlinear Methods of Nonparametric Regression Analysis," *Soviet Automatic Control*, Vol. 5, pp. 25-34, 1979.

Lancaster, P. and Salkauskas, K., "Surfaces Generated by Moving Least-Squares Methods," *Math. Comput.*, Vol. 37, pp. 141–158, 1981.

Maluf, N., "An Introduction to Microelectromechanical Systems Engineering," *Artech House*, Boston, 2004.

Montgomery, D.C., *Design and Analysis of Experiments*, Wiley, New York, 1997.

Pettit, C.L., Canfield, R.A., and Ghanem, R., "Stochastic Analysis of an Aeroelastic System," *15th ASCE Engineering Mechanics Conference*, Columbia University, New York, 2002.

Sharpe, W. N., Turner, K.T., and Edwards. R.L., "Tensile Testing of Polysilicon," *Experimental Mechanics*, Vol. 39, No. 3, pp. 162-170, 1999.

Stone, C.J., "Consistent Nonparametric Regression," *The Annals of Statistics*, Vol. 5, pp. 595-645, 1977.

Toropov, V., Scharamm, U., Sahai, A., Jones, R., and Zeguer, T., "Design Optimization and Stochastic Analysis based on the Moving Least Squares Method," *6th World Congresses of Structural and Multidisciplinary Optimization*, Brazil, 2005.

Watanabe, S., "Karhunen-Loève Expansion and Factor Analysis," *In Transactions of the 4th Prague Conference on Information Theory*, Prague, pp. 635-660, 1965.

Wiener, N., "The Homogeneous Chaos," *American Journal of Mathematics*, Vol. 60, pp.897-936, 1938.