20th AIAA Computational Fluid Dynamics Conference
27 - 30 June 2011, Honolulu, Hawaii

AIAA 2011-3069

Robust Optimal Control using Polynomial Chaos and Adjoints for Systems with Uncertain Inputs

Sriram*
General Electric Research Center, Niskayuna, NY, 12309, U.S.A.

Antony Jameson†
Stanford University, Stanford, CA, 94305, U.S.A.

The objective of this note is to show how one can combine Polynomial Chaos Expansions (PCE) and adjoint theory to efficiently obtain sensitivities for robust optimal control. A non-intrusive PCE method is used to analyze the constraint equations for the state (which depends on uncertain inputs), namely the governing equations of the dynamical system. Adjoint solutions are constructed for each of the polynomial basis functions used in the approximate expansion. The combination of the gradient for each basis-adjoint pair is used to form the overall gradient. The resulting gradient can be used to improve an initial guess in an iterative optimization procedure. The repeated use of the non-intrusive PCE method, the adjoint solver and the gradient estimate can be used to determine optimal control laws for the governing system in the presence of uncertainties. The formulation of the optimal control problem is presented in the context of the flow equations where the expected value of a functional is to be minimized. The boundary shape is the control. The associated cost of this approach in an optimization setting is equal to the cost of a PCE analysis ($\approx Q$ deterministic simulations) plus Q (number of unknowns in the PCE expansion) adjoint solves for each iteration of a steepest-descent algorithm. However, this cost can be further reduced for certain objective functions using an intrusive formulation for the adjoint equations.[11]

I. Introduction

Engineering designs are typically guaranteed a certain degree of robustness due to the multi-point nature of the overall design process. This ensures performance at a few key operating points. These operating points are typically further apart in the design space and have very different characteristics. Take-off and landing, cruise and a coordinated turn are some key operating points in the flight envelope of an airplane. At each of these operating points, there is typically some uncertainty (both flow and geometry). One can envision another multi-point design process to ensure robustness to these uncertainties.

Advances in modeling of systems governed by Partial Differential Equations(PDEs) with uncertain inputs and the recent explosion of Uncertainty Quantification (UQ) methods offers a new approach to quantify robustness. UQ methods based on Polynomial Chaos (PC) theory are particularly efficient as they identify the optimal choice of basis functions to represent the characteristics of the system for a given form of input uncertainties. While PC was originally developed for Gaussian uncertain inputs for linear systems,[1] the Wiener-Askey family of orthogonal polynomials provides one-to-one correspondence for most forms of input uncertainties. The use of the PC method reduces the problem of determining the probabilistic quantities of the system to one of determining the coefficients of the polynomial basis functions, and is essentially a spectral approach in probability space. The dependence of the number of basis functions on the required order of expansion and the number of input uncertainties leads to rapid rise in computational cost which will pose a hurdle for immediate acceptance in industrial Engineering environments.

*Mechanical Engineer, Fluid Mechanics Lab
†Thomas V. Jones Professor of Engineering, Department of Aeronautics and Astronautics

The determination of optimal control laws for systems with uncertain inputs requires analysis techniques that propagate the effect of the input uncertainty and this is typically achieved through any UQ analysis. One approach to determining optimal control laws for such systems is through Pontryagin's maximum principle. This control theory approach when used for systems governed by PDEs[2–4] has dramatic computational cost advantages over the finite-difference method of calculating gradients. With the control theory approach the necessary gradients are obtained through the solution of an adjoint system of equations of the governing equations of interest. The adjoint method is extremely efficient since the computational expense incurred in the calculation of the complete gradient is effectively independent of the number of design variables. In this study, a continuous adjoint formulation has been used to derive the adjoint system of equations. Hence, the adjoint equations are derived directly from the governing equations and then discretized. Hence, this approach has the advantage over the discrete adjoint formulation in that the resulting adjoint equations are independent of the form of discretized flow equations. The adjoint system of equations have a similar form to the governing equations of the flow and hence the numerical methods developed for the flow equations can be reused for the adjoint equations. When used for boundary control, every discrete point that defines the boundary is allowed to move, enabling a large variety of boundary control. Any constraints on the boundary shape is imposed by projecting the gradient into an allowable sub-space.

It is the objective of this article to expound on the combined use of polynomial chaos theory and adjoints to provide an alternative frame-work for robust optimization. In Section II a brief overview of the Polynomial Chaos method is provided and in Section III the adjoint theory for deterministic systems is outlined. Section IV combines Section II and III to highlight the central tenet of this article. Section V provides the outline of an algorithm that can be used to estimate robust optimal control laws. Section VII establishes the equivalence between the use of a non-intrusive and intrusive UQ method within the algorithm outlined in Section V. Finally, Section VIII presents some results for an academic aerodynamic problem where the Mach number is treated as an uncertain input to a system that solves the steady-state Euler equations for flow around an airfoil. Optimal shapes obtained with the approach outlined in this study are also shown.

II. Polynomial Chaos Expansions (PCEs)

Under the broad umbrella of uncertainty quantification, a large number of recent investigations have centered around the use of polynomial chaos methods (and their variants). The use these methods can be traced to the seminal work of Wiener[1] and more recently to Spanos and Ghanem[6,7] and Xiu and Karniadakis[8,9] (Please refer to Xiu[5] for a comprehensive survey article). Generalized Polynomial Chaos Expansion (GPCE) methods use hyper-trigonometric polynomials from the Wiener-Askey family to approximate the output of systems due to random inputs. Typically, an output, w, is written as an polynomial expansion

$$w(\omega) = \sum_{i=1}^{P} \hat{w}_i \Phi_i(\xi(\omega))$$

where $\omega \in \Omega$ is an element in the event space, Φ_i is an element of an orthogonal Wiener-Askey family of polynomials and ξ is a random vector that defines ω. Estimates of $\hat{w}_i$ provide surrogate function definitions for w, that can then be used either for analysis or optimization strategies. The process of determining the $\hat{w}_i$ is the focus of most GPCE methods. We will assume that all random variables that are considered here (both input uncertainties and output performance measures) belong to the triplet $(\mathbb{R}, \mathcal{B}, \mathcal{P})$ where $\mathcal{B}$ is the σ-field generated by Borel sets in $\mathbb{R}$ and $\mathcal{P}$ is the set of probability measures on $(\mathbb{R}, \mathcal{B})$. One of the central tenets of polynomial chaos methods is the use of orthogonality relations between the polynomial basis and measures from the probability space $\mathcal{P}$.

The common methods to compute the coefficients are classified under the labels of intrusive and non-intrusive methods. The former uses GPCE for the unknowns in the governing equations and uses the method of weighted residuals and the orthogonality property to form a set of coupled non-linear ODEs for the unknown coefficients while the non-intrusive method use "sampling" techniques to determine the coefficients. Evaluation of intrusive methods is beyond the scope of this work, so we concentrate on non-intrusive methods. The curse-of-dimensionality is a serious issue with GPCE methods that requires efficient sampling techniques (e.g. tensor product for low dimensional problems, sparse grids and cubature rules for higher dimensional problems). While we do not compare these approaches here, please refer to Xiu[5] for a detailed review.

The pseudo-spectral (collocation) approach uses a set of quadrature/collocation points to obtain estimates of the output of the system. The collocation points are the Gauss-Hermite quadrature points and weights and the orthogonality of the polynomials can be used to determine the coefficients. It should be noted that the quadrature points are evaluated for the weighting function $e^{-x^2/2}$.

The pseudo-spectral discrete expansion, w_N^P, uses the orthogonal polynomials $\Phi(\xi(\omega))$ from the Wiener-Askey basis and finds the coefficients of the expansion, $\hat{w}_m$.[5]

$$\mathcal{I}_N^P u \simeq w_N^P(\omega) = \sum_{i=1}^{M} \hat{w}_i \Phi_i(\xi(\omega)); \quad M = \frac{(N+P)!}{N!P!} \tag{1}$$

$$\hat{w}_i = \sum_{j=1}^{Q} u(\omega^j)\Phi_m(\xi(\omega)^j)\alpha^j \tag{2}$$

$Q \in \mathbb{I}$ is the set of sampling points indexed by j used for the PCE reconstruction. For pseudo-spectral (collocation) schemes, $Q = P$. The choice of the nodes should be such that

$$\sum_{j=1}^{Q} f(y^j)\alpha^j \equiv \int f(y)\rho(y)dy = \mathbb{E}[f(y)]$$

where $\rho(y)$ is the measure w.r.t. which the basis, Φ is normalized such that $h_m = 1, \ \forall m$ and h_m is given by

$$\int \rho(y_i)\Phi_m(y_i)\Phi_n(y_i)dy_i = h_m^2 \delta_{mn} \tag{3}$$

where δ_{mn} is the Kroneckar delta function.

Once the coefficients are evaluated statistical information can be easily extracted through the following identities. The expected value or the first moment can be written as

$$\mathbb{E}[w] \approx \int (\sum_{i=1}^{P} \hat{w}_i \Phi_i(y))\rho(y)dy = \hat{w}_1,$$

due to the fact that $\Phi_i(y)$ is orthogonal to $\Phi_1(y) = 1$ (Equation 3) for all $i \geq 2$. Using the same property and the expression for the variance in terms of the expected value, the variance can be written in terms of the PCE coefficients as

$$Var(w) = \mathbb{E}w^2 - \mathbb{E}^2 w \approx \sum_{m=2}^{P} \hat{w}_m^2$$

A similar approach can be used to compute the second-moment and global sensitivity coefficients.[5]

A. Some PCE techniques

We focus our attention on three different methods. They are outlined below for completeness.

1. Non-Intrusive Sampling Method (NIS)

In this method, (which we will sometimes refer to as the Linear Regression method) we use Hermite Polynomials to approximate the outputs of the system. The coefficients of the GPCE are obtained using an regression approach. Hence, we evaluate the output at a set of predetermined points and assemble a linear (over-determined) system of equations

$$\begin{array}{l}
\Phi_0(\xi_1)\hat{w}_1 + \Phi_1(\xi_1)\hat{w}_2 + \cdots + \Phi_P(\xi_1)\hat{w}_P = u(\xi_1) \\
\Phi_0(\xi_2)\hat{w}_1 + \Phi_1(\xi_2)\hat{w}_2 + \cdots + \Phi_P(\xi_2)\hat{w}_P = u(\xi_2) \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\Phi_0(\xi_Q)\hat{w}_1 + \Phi_1(\xi_Q)\hat{w}_2 + \cdots + \Phi_P(\xi_Q)\hat{w}_P = u(\xi_Q)
\end{array}$$

which is solved in a least-squares sense for $\hat{w}_i, \ \ i = 1...Q$ (using MATLAB's \ operator). Here, Q is the number of sampling points and P is the order of the GPCE.

2. *Pseudo-Spectral Method with Hermite Polynomials (PSH)*

In this method, we use a pseudo-spectral (collocation) approach and obtain estimates of the output of the system at a set of quadrature points. The collocation points are the Gauss-Hermite quadrature points and weights and the orthogonality of the polynomials can be used to determine the coefficients. It should be noted that the quadrature points are evaluated for the weighting function $e^{-x^2/2}$.

3. *Pseudo-Spectral Method with Legendre Polynomials (PSL)*

This method is very similar to the previous method expect that Legendre polynomials are used instead of Hermite polynomials. The reason for evaluating this method is two-fold. The support for the Gauss-Hermite quadrature points lies in $[-\infty, \infty]$ and hence higher order evaluations requires deterministic runs at quadrature points where it may be difficult to "converge" numerical solvers. Additionally, we would also like to evaluate the degradation in the pseudo-spectral method when we violate the choice of the Wiener-Askey polynomial basis. This will be useful for problems when the input uncertainties are mixed (e.g. Gaussian and Uniform) and we use Hermite (or Legendre) polynomials to approximate the output.

III. Deterministic Adjoint Theory

We provide a brief overview of the adjoint process[4] for deterministic systems and describe it in the context of Euler (or Navier-Stokes) equations that govern the evolution of fluid flow.

The cost functions are functions of the state variables, u, and the control variables, which may be represented by the function, $\mathcal{F}$, say. Then

$$I = I(u, \mathcal{F}),$$

and a change in $\mathcal{F}$ results in a change

$$\delta I = \frac{\partial I^T}{\partial u}\delta u + \frac{\partial I^T}{\partial \mathcal{F}}\delta \mathcal{F}, \tag{4}$$

in the cost function. Using control theory, the governing equations for the state variables are introduced as a constraint in such a way that the final expression for the gradient does not require re-evaluation of the state. In order to achieve this, δw must be eliminated from equation 4. Suppose that the governing equation R which expresses the dependence of w and $\mathcal{F}$ within the domain D can be written as

$$R(u, \mathcal{F}) = 0 \tag{5}$$

Then δu is determined from the equation

$$\delta R = \left[\frac{\partial R}{\partial u}\right]\delta u + \left[\frac{\partial R}{\partial \mathcal{F}}\right]\delta \mathcal{F} = 0 \tag{6}$$

Next, introducing a Lagrange Multiplier ψ, we have

$$\delta I = \frac{\partial I^T}{\partial u}\delta u + \frac{\partial I^T}{\partial \mathcal{F}}\delta \mathcal{F} - \psi^T\left(\left[\frac{\partial R}{\partial u}\right]\delta u + \left[\frac{\partial R}{\partial \mathcal{F}}\right]\delta \mathcal{F}\right)$$

$$\delta I = \left(\frac{\partial I^T}{\partial u} - \psi^T\left[\frac{\partial R}{\partial u}\right]\right)\delta u + \left(\frac{\partial I^T}{\partial \mathcal{F}} - \psi^T\left[\frac{\partial R}{\partial \mathcal{F}}\right]\right)\delta \mathcal{F}$$

Choosing ψ to satisfy the adjoint equation

$$\left[\frac{\partial R}{\partial u}\right]^T\psi = \frac{\partial I}{\partial u} \tag{7}$$

the first term is eliminated and we find that

$$\delta I = \mathcal{G}\delta\mathcal{F} \tag{8}$$

where

$$\mathcal{G} = \frac{\partial I^T}{\partial \mathcal{F}} - \psi^T \left[\frac{\partial R}{\partial \mathcal{F}}\right] \tag{9}$$

This process allows for elimination of the terms that depend on the flow solution with the result that the gradient with respect with an arbitrary number of design variables can be determined without the need for additional evaluations of the state.

After taking a step in the negative gradient direction, the gradient is recalculated and the process repeated to follow the path of steepest descent until a minimum is reached. In order to avoid violating constraints, the gradient can be projected into an allowable subspace within which the constraints are satisfied. In this way one can devise procedures which must necessarily converge at least to a local minimum and which can be accelerated by the use of more sophisticated descent methods such as conjugate gradient or quasi-Newton algorithms. There is a possibility of more than one local minimum, but in any case this method will lead to an improvement over the original design. The above process solves the following optimization problem:

$$\begin{aligned} &\inf_{\mathcal{F}\in\mathbf{F}} \quad I(u,\mathcal{F}) \\ \text{s.t.} \quad & R(u,\mathcal{F}) = 0 \\ & C(u,\mathcal{F}) \leq 0 \end{aligned} \tag{10}$$

IV. Extension to Systems with Uncertain Inputs : A Partially Intrusive Algorithm

In this section, we discus a partially intrusive approach that uses the PCE to formulate the adjoint system. Expanding the objective, residual and flow solutions in terms of PC expansions, allows for derivation of an adjoint system for each "mode" of the PCE. These modes can then be combined to determine an expression for the stochastic gradient.

The adjoint formulation in Section III can be re-written in the following form for systems with uncertain inputs. The general stochastic optimization problem can be written as

$$\begin{aligned} &\inf_{\mathcal{F}\in\mathbf{F}} \quad \mathcal{J} \\ \text{s.t.} \quad & \mathcal{P}(\hat{R}(\hat{u}(\xi),\mathcal{F})) = \mathcal{P}(r) \\ & \mathcal{P}(\hat{C}(\hat{u}(\xi),\mathcal{F})) = \mathcal{P}(c) \end{aligned} \tag{11}$$

$\mathcal{J}$ can be mean or s.t.d (or other moments of interest) of a primitive function, $\mathcal{I}$. Using a simplified notation for the probabilistic constraints we can write the optimization problem where the objective is to reduce the mean of $\mathcal{I}$ as follows:

$$\begin{aligned} &\inf_{\mathcal{F}\in\mathbf{F}} \quad \mathbb{E}(\hat{I}(\hat{u}(\xi),\mathcal{F})) \\ \text{s.t.} \quad & \hat{R}(\hat{u}(\xi),\mathcal{F}) = 0 \quad a.s. \\ & \hat{C}(\hat{u}(\xi),\mathcal{F}) < 0 \quad a.s. \end{aligned} \tag{12}$$

where the ˆ symbol is used to emphasize that the quantities are probabilistic in nature. Note that the control variables are still deterministic and the stochastic nature of the problem arises from the dependence of $\hat{u}$ on the random inputs ξ. For simplicity, the first moment is used as the objective that is minimized by the

optimal $\mathcal{F}$, but the optimization problem can also be higher moments or other probabilistic quantities. Due to the nature of the basis used in PCE, the first moment, $\mathbb{E}$ offers some simplifications that is highlighted in this Section. A schematic sketch of the control problem is shown in Figure 1
In many common engineering cases the properties of $\mathcal{H}$ can be used to obtain expressions for the variation in $\mathcal{J}$ in terms of variations of I. For example, if $\mathcal{H}$ is the $\mathbb{E}$, then the linearity properties can be used to write $\mathcal{J} = \mathbb{E}\delta\mathcal{I}$. In general,

$$
\begin{aligned}
\mathcal{J} &= \mathcal{H}(I(\hat{w}, \mathcal{F}) \\
\delta\mathcal{J} &\approx \mathcal{J}^+ - \mathcal{J}^0 \\
&= \mathcal{H}(I^+) - \mathcal{H}(I^0) \\
&= \mathcal{H}(I) + \mathcal{H}(\frac{\partial I}{\partial \hat{w}}\delta w + \frac{\partial I}{\partial \mathcal{F}}\delta \mathcal{F}) - \mathcal{H}(I) \\
&= \mathcal{H}(\frac{\partial I}{\partial \hat{w}}\delta w + \frac{\partial I}{\partial \mathcal{F}}\delta \mathcal{F}) \\
&= \mathcal{H}(\delta I)
\end{aligned}
\tag{13}
$$

The key insight that enables the efficient re-use of the adjoint method is the observation that the PCE approach enables re-construction of any system output, w as a linear expansion of the form:

$$w(\omega) = \sum_{i=1}^{Q} \hat{w}_i \Phi_i(\xi(\omega))$$

Hence, $\hat{I}, \hat{R}, \hat{u}$ etc. can be written as a linear combination of basis functions. The choice of H_i is obtained from the Wiener-Askey basis and depends on the nature of the uncertain input. Substituting this expansion into the derivation outlined in Section III formulates the adjoint system and the gradient. This is outlined here for completeness. w_i^x is used to denote the i^{th} co-efficient associated with the random variable x in the PCE expansion.

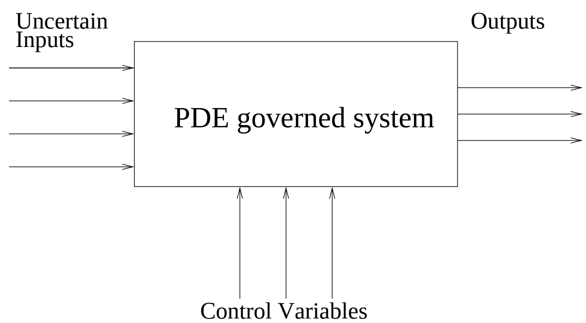


Figure 1. A schematic of the control problem with uncertain inputs. The objective is to find the optimal control variables.

$$
\begin{aligned}
\hat{I}(\omega) &= \sum_{i=1}^{Q} \hat{w}_i^I \Phi_i(\xi(\omega)) = \sum_{i=1}^{Q} \hat{I}_i \\
\hat{R}(\omega) &= \sum_{i=1}^{Q} \hat{w}_i^R \Phi_i(\xi(\omega)) = \sum_{i=1}^{Q} \hat{R}_i \\
\hat{u}(\omega) &= \sum_{i=1}^{Q} \hat{w}_i^u \Phi_i(\xi(\omega)) = \sum_{i=1}^{Q} \hat{u}_i
\end{aligned}
\tag{14}
$$

Recalling that the expected value of a random variable under the PCE approximation is just the first coefficient, w_1 and Φ_1 is typically 1, offers considerable simplifications to the adjoint process. To encompass the general case of any objective function, we retain the sub-script i notation in the following derivation to denote the possible influence of coefficients > 1.

Now following the deterministic approach, variations in cost function, $\hat{I}$, and the constraint, $\hat{R}$, can be written as a sum over the variations with respect to each of the coefficients of the PC basis functions, $\hat{w}_i$. We drop the $\mathbb{E}$ notation for I and use the subscript notation to denote the contribution of the i^{th} term in the PCE for I. Please note that we do not formulate the problem using $\mathbb{E}(R)$ instead of R. With these notational changes Equation 4 can be written as

$$
\begin{aligned}
\delta \hat{I} &= \sum_{i=1}^{Q}\left(\frac{\partial \hat{I}_i^T}{\partial \hat{u}}\delta \hat{u} + \frac{\partial \hat{I}_i^T}{\partial \mathcal{F}}\delta \mathcal{F}\right) \quad \text{but} \quad \hat{u} = g(\hat{w}_i) \\
\delta \hat{I} &= \sum_{i=1}^{Q}\left(\frac{\partial \hat{I}_i^T}{\partial \hat{w}_i}\delta \hat{w}_i + \frac{\partial \hat{I}_i^T}{\partial \mathcal{F}}\delta \mathcal{F}\right) \qquad (15)
\end{aligned}
$$

where $\hat{I}_i$ is the i^{th} contribution from the PCE. Similarly, Equation 6 can be written as

$$
\delta \hat{R} = \sum_{i=1}^{Q}\left(\left[\frac{\partial \hat{R}_i}{\partial \hat{w}_i}\right]\delta \hat{w}_i + \left[\frac{\partial \hat{R}_i}{\partial \mathcal{F}}\right]\delta \mathcal{F}\right) = 0 \quad a.s. \qquad (16)
$$

Multiplying Equation 16 by a Lagrange multiplier ψ are before but in indexed form (to correspond to the PC basis function $\hat{w}_i$) and combining Equations 15 and 16 and grouping terms leads to

$$
\delta \hat{I} = \sum_{i=1}^{Q}\left[\left(\frac{\partial \hat{I}_i^T}{\partial \hat{w}_i} - \psi_i^T\left[\frac{\partial \hat{R}_i}{\partial \hat{w}_i}\right]\right)\delta \hat{w}_i + \left(\frac{\partial \hat{I}_i^T}{\partial \mathcal{F}} - \psi_i^T\left[\frac{\partial \hat{R}_i}{\partial \mathcal{F}}\right]\right)\delta \mathcal{F}\right] \qquad (17)
$$

Now the ψ_i can be chosen to eliminate the dependence of $\hat{I}$ on $\hat{w}_i$ and these become the adjoint system of equations to solve. The form of the adjoint equations (after multiplying through with Φ_j, integration with respect to an appropriate weighting function and using the orthogonal property in Equation 3) is,

$$
\left[\frac{\partial \hat{R}_i}{\partial \hat{w}_i}\right]^T \psi_i = \frac{\partial \hat{I}_i}{\partial \hat{w}_i} \quad i = 1, 2, ..., Q \qquad (18)
$$

and the gradient can be written as

$$
\delta \hat{I} = \hat{\mathcal{G}}\delta \mathcal{F} \qquad (19)
$$

where

$$
\hat{\mathcal{G}} = \sum_{i=1}^{Q}\left[\frac{\partial \hat{I}_i^T}{\partial \mathcal{F}} - \psi_i^T\left[\frac{\partial \hat{R}_i}{\partial \mathcal{F}}\Phi_i(\xi)\right]\right] \qquad (20)
$$

Hence, to evaluate the gradient, one needs to be able to reconstruct the i^{th} approximation to the partials required in the adjoint system (Equation 18) and the gradient equation (Equation 20). As the non-intrusive method using PCEs for UQ scales weakly with the number of outputs, this is not major stumbling block. However, the number of adjoint solves scales with the number of unknowns in the PCE expansion which is known to have rapid growth with order and number of input uncertainties. For engineering estimates, order 2 has been found to be sufficient.

After taking a step in the negative gradient direction, the gradient is recalculated and the process repeated to follow the path of steepest descent until a minimum is reached. As in the deterministic case, in order to avoid violating constraints, the gradient can be projected into an allowable subspace within which the constraints are satisfied. In this way one can devise procedures which must necessarily converge at least to a local minimum where it is robust the uncertain inputs to the system.

V. Overview of the Optimization Process

The process outlined in the previous section can be summarized as an algorithm:
Note : This algorithm is partially intrusive as it requires reformulation of the adjoint system developed for deterministic calculations. In Appendix A some worked out examples are presented. While this approach requires reworking of the adjoint solver, if the objective function is simple like the mean, it only requires the solution of adjoint system for *all* the uncertain inputs, significantly reducing the cost of the adjoint component.

Data: Define uncertain inputs to system (μ, σ etc.), Choose PCE from Wiener-Askey basis
Initialize $\hat{G} = 0$;
while $\|\hat{\mathcal{G}}\| \neq 0$ **do**
  $\mathcal{F} = \mathcal{F} - \lambda\hat{G}$, λ is a constant;
  Use non-intrusive method to estimate PCE;
  Use PCE to estimate quantities required for adjoint solve;
  for $i \in 1, 2, ...Q$ **do**
    Solve adjoint system in Equation 18;
  end
  Compute $\hat{G}$ using Equation 20;
  Project $\hat{G}$ into allowable sub-space to satisfy constraints;
end

Algorithm 1: A Partially-Intrusive, Steepest-Descent version of the Optimization Algorithm

VI. Extension to Systems with Uncertain Inputs : A Non-Intrusive Approach

In this section, we outline a completely non-intrusive approach. Here, we solve for the flow and the adjoint equations at each of the sampling points. The gradient is then constructed using a combination of the flow and adjoint solutions and gradient at each sampling point. This gradient is then represented as a PCE and the appropriate coefficients (depending on the objective function) are used to determine the change to the geometry that results in improvement of the performance metric.

The algorithm can be outlined as follows:

Data: Define uncertain inputs to system (μ, σ etc.), Choose PCE from Wiener-Askey basis
Initialize $\hat{G} = 0$;
while $\|\hat{\mathcal{G}}\| \neq 0$ **do**
  $\mathcal{F} = \mathcal{F} - \lambda\hat{G}$, λ is a constant;
  Use non-intrusive method to estimate PCE of flow;
  Use non-intrusive method to estimate PCE of adjoint;
  At each sampling point, compute $\hat{G}_j$ using Equation 9;
  Using $\hat{G}_j$, reconstruct a PCE of the gradient $\hat{G}$;
  Project $\hat{G}_j$ into allowable sub-space to satisfy constraints;
end

Algorithm 2: A Non-Intrusive, Steepest-Descent version of the Optimization Algorithm

VII. (Lack of) Equivalence between the two approaches

In this section, we show that the method outlined in Section VI that is based on a non-intrusive UQ method is an approximation to one with a partially intrusive method (Section IV). The non-intrusive approach is easier to implement and can readily re-use deterministic codes used for aerodynamic design but is computationally expensive. The partially intrusive approach requires additional coding (and derivation) but could significantly reduce the cost of the adjoint simulation. This is due to the fact that as shown in Figure 2 the number of sampling counts grows rapidly with increase in number of input random variables. On the other hand, the partially intrusive approach only requires solution of a number of adjoint systems equivalent to the number of unknown coefficients in the PCE expansion. In addition, for common aerodynamic objective functions like the expected value of drag, one only needs the solution of the adjoint system for the first term in the PCE expansion. The proof of this equivalence is obtained by equating the expression for the coefficients of the PCE expansion for the gradient through the non-intrusive approach and that constructed through the PCE expansion for the adjoints in the partially intrusive approach. The two forms are slightly different and the difference contains higher-order terms in the PCE expansion and terms that couple the adjoint solution and the perturbation terms at different sampling points.

The key connection between the two different methods is through the explicit expression for the coef-

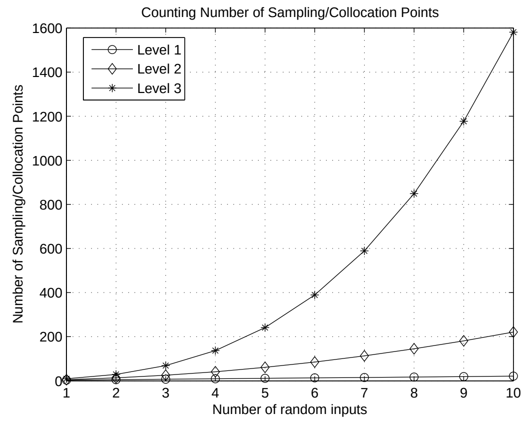


Figure 2. Number of sampling points using a Smolyak Sparse grid and Clenshaw-Curtis abscissas

ficients in terms of the solution at the different sampling points. We will use the subscripts j to denote the sampling points and i to denote the coefficients of the PCE. Recalling that the expression for the m^{th} coefficient is of the form:

$$\hat{w}_m = \sum_{j=1}^{Q} u(\omega^j)\Phi_m(\xi(\omega)^j)\alpha^j$$

where Q is the number of sampling points. Hence, the gradient expression in Section VI can be written as

$$\hat{G}_m = \sum_{j=1}^{Q} \hat{G}(\omega^j)\Phi_m(\xi(\omega)^j)\alpha^j$$

which for $m = 0$ reduces to a linear combination of the gradient at various sampling points. The gradient at the different sampling points can be written using the deterministic formulation (Equation 9) as

$$\hat{G}_j = \frac{\partial \hat{I}_j^T}{\partial \mathcal{F}} - \hat{\psi}_j^T \left[\frac{\partial \hat{R}_j}{\partial \mathcal{F}}\right] \tag{21}$$

and the coefficients of the gradient for the non-intrusive approach can be written as

$$\hat{G}_i = \sum_{j=1}^{Q} \hat{G}_j \Phi_i(\xi_j)\alpha^j \tag{22}$$

Comparing this coefficient to that obtained from the partially intrusive approach in Equation 20

$$\begin{aligned} \hat{G} &= \sum_i \hat{G}_i \Phi_i(\xi) \qquad (23) \\ \hat{G} &= \sum_{i=1}^{Q} \left[\frac{\partial \hat{I}_i^T}{\partial \mathcal{F}} - \psi_i^T \left[\frac{\partial \hat{R}_i}{\partial \mathcal{F}}\right]\right] \Phi_i(\xi) \\ \hat{G}_i &= \left[\frac{\partial \hat{I}_i^T}{\partial \mathcal{F}} - \psi_i^T \left[\frac{\partial \hat{R}_i}{\partial \mathcal{F}}\right]\right] \end{aligned}$$

and expanding $\hat{I}$ and $\hat{\psi}$ in terms of the values at the different sampling points, we can write the coefficients $\hat{G}_i$ as

$$\hat{G}_i = \sum_{j=1}^{Q} \Phi_i(\xi_j)\alpha^j \left[\frac{\partial \hat{I}_j^T}{\partial \mathcal{F}}\right] - \sum_{j=1}^{Q} \Phi_i(\xi_j)\alpha^j \left[\psi_j^T\right] \sum_{j=1}^{Q} \Phi_i(\xi_j)\alpha^j \left[\frac{\partial \hat{R}_j}{\partial \mathcal{F}}\right]$$

Comparing this with the coefficients in Equation 22, it is clear that the double summation term is only partly accounted for in the non-intrusive approach. Hence, the contribution of the product term involving the adjoint solutions and the variation in the residual due to boundary surface movement across various sampling points is not accounted for in the expressions for the gradient. Our numerical studies suggest that lack of completeness in the gradient expression is relevant for coefficients $m >\geq 1$ and hence could affect computations where higher-order moments of the objective function are being optimized. Alternately, one can easily modify the gradient expressions in the non-intrusive approach to include these terms

VIII. Results

A. Assessment of PCE for an Analytical Problem

We start with a simple (but representative) analytical problem. Consider a problem with one random variable, $\xi \in [-1, 1]$ and an output, $u = 2 + \xi^2$ if $\xi > 0.0$ and $u = 2 + 0.1\ \xi^2$ if $\xi <= 0.0$. We assume that ξ is normally distributed with $\mu = 0$ and $\sigma = 1.0$. The form of u is assumed to be **representative** of the drag-divergence characteristics of transonic airfoils. We will use the results here to decide on the choice of the non-intrusive method for the optimal control problem we are interested in.

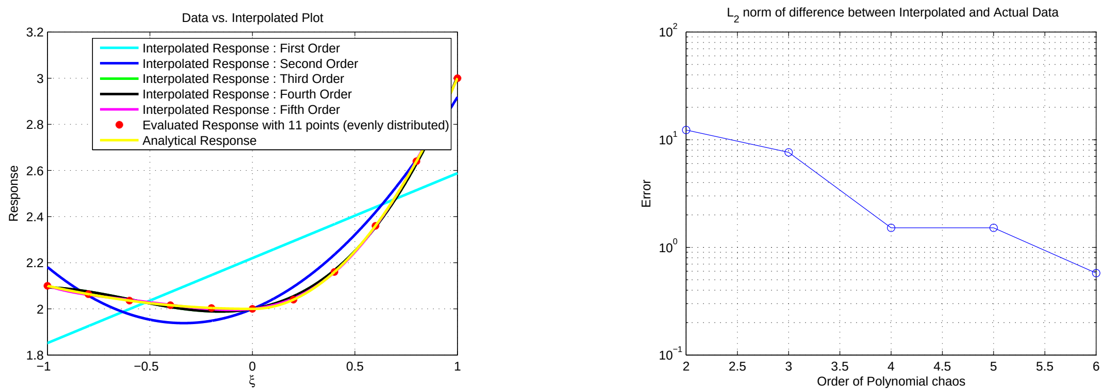
For the non-intrusive sampling approach, an equally distributed set of points in $[-1, 1]$ was sampled and used to reconstruct varying orders of polynomial chaos. A linear regression approach was used to determine the unknown coefficients of the expansion. ../figures 3(a) compares the interpolated curve using various expansion orders. This plot would suggest that a fourth order fit is sufficient for most engineering calculations. However, Fig. 3(b) suggests that the L_2 norm of the difference between the fitted and analytical data shows poor convergence rate with increasing p. In addition, the regression approach was found to be sensitive to the choice of sampling. Both Gauss-Hermite quadrature points and random sampling (using the MATLAB rand function) resulted in poor error convergence with increasing P. Figure 4 shows the error in the mean and the variance for increasing polynomial order. The lack of reduction in the variance suggests that alternate variance reduction sampling need to be explored if accurate variance is a requirement.

For the pseudo-spectral approach, the moments of the output are the immediate outputs. Figure 6 studies the effect of the number of quadrature points (Gauss-Hermite) on the convergence of mean and variance for a fixed polynomial order (P=2). This suggests that about 8 quadrature/sampling points are required to determine the mean and variance within an error of 10^{-3}. The interpolated fits using different orders of polynomial chaos are shown in Fig. 5(a). In comparison to Fig. 3(a), the interpolation polynomial generated by the pseudo-spectral approach looks less accurate. This is also evident from the L_2 errors shown in Fig. 5(b). However, one needs to be cautious while using higher-order expansions as it will require obtaining converged computational solutions at quadrature points that belong to the set $[-\infty, \infty]$. For example, if the standard deviation in incidence angle is around 3 degrees, then we might require solutions for incidence angles of 9 degrees for the eighth order expansion (quadrature points at $\approx \pm 3$).

Alternately, the Gauss-Legendre points have compact support (in $[-1, 1]$) and hence could be an alternative candidate for the pseudo-Galerkin approach. Figure 8 studies the effect of the number of quadrature points (Gauss-Legendre) on the convergence of mean and variance for a fixed polynomial order (P=2). The degradation in convergence is stark. The interpolated fits using different orders of polynomial chaos are shown in Fig. 7(a). Again, the error is larger than the error in Fig. 3(a).

Figures 9(b), 10(a) and 10(b) compare the Probability Density Function (pdf) of the response for different orders of polynomial chaos obtained using the linear regression and the stochastic collocation methods. When compared to Fig. 9(a) all three methods capture the general trend of the Monte-Carlo results. The non-intrusive sampling approach provides convergent pdfs for increasing polynomial chaos order. A fifth order expansion seems to be sufficient to capture the trend of the pdf. It is clear that the stochastic collocation methods converges to the pdf obtained using the analytical expression.

The conclusions for this analytical problem are summarized in Table 1. While the pseudo-spectral approach with Gauss-Hermite quadrature points results in a good surrogate function, the error in the mean and variance is much smaller at a lower polynomial order for the NIS and PSL methods. In addition, both these methods require fewer function evaluations making them attractive candidates for probabilistic



(a) Comparison of Interpolated and Analytical Data (NIS)

(b) L_2 error between fitted and analytical data (NIS)

Figure 3. Surrogate fits and Error for NIS

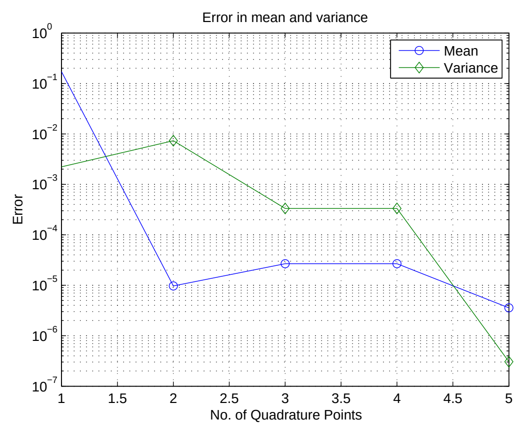
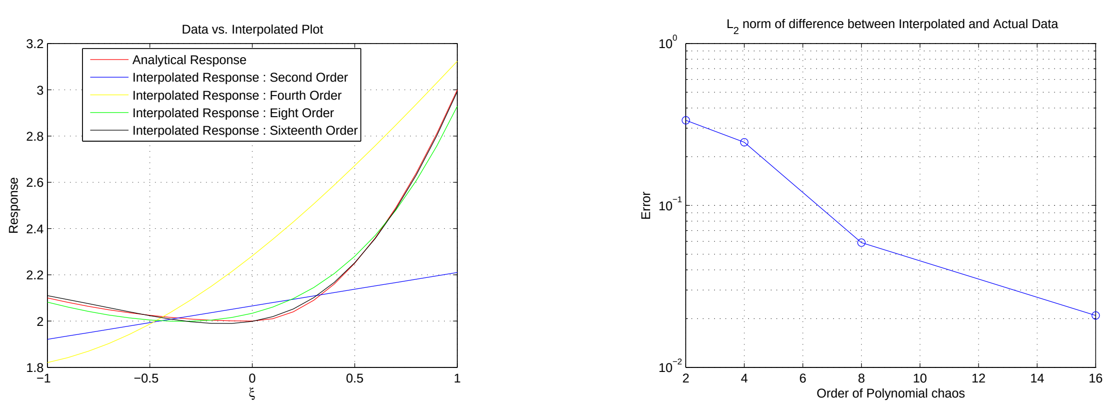


Figure 4. Error in the mean and variance (NIS)



(a) Comparison of Interpolated and Analytical Data (PSH)

(b) L_2 error between fitted and analytical data (PSH)

Figure 5. Surrogate fits and Error for PSH

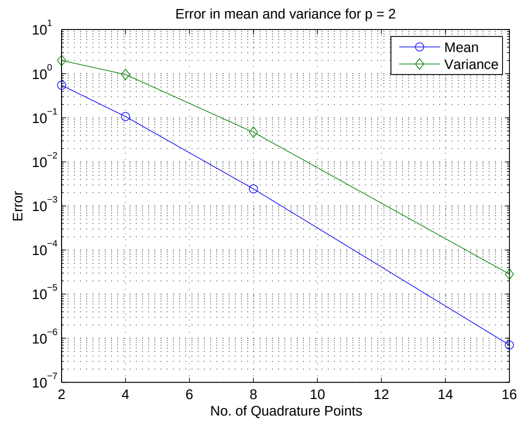
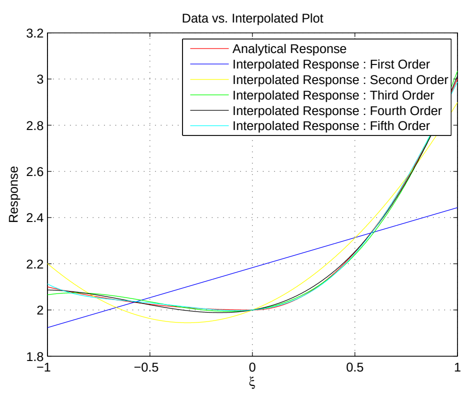
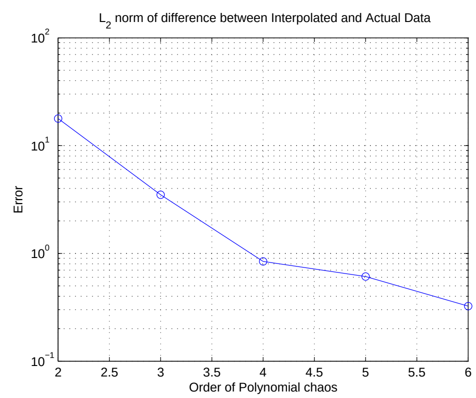


Figure 6. Convergence of mean and variance (PSH)



(a) Comparison of Interpolated and Analytical Data (PSL)



(b) L_2 error between fitted and analytical data (PSL)

Figure 7. Surrogate fits and Error for PSL

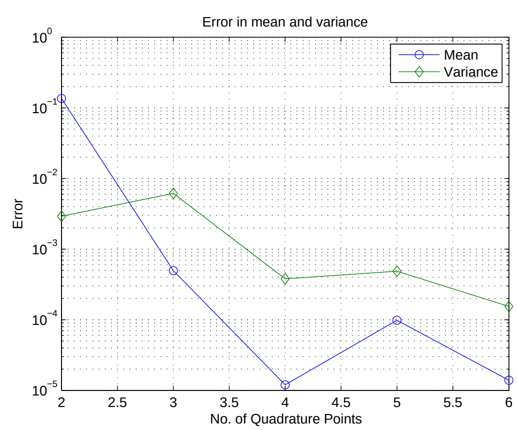
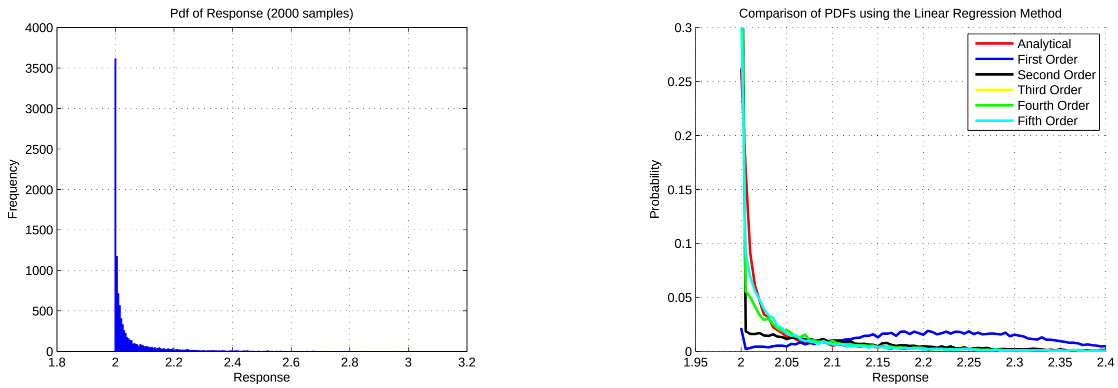


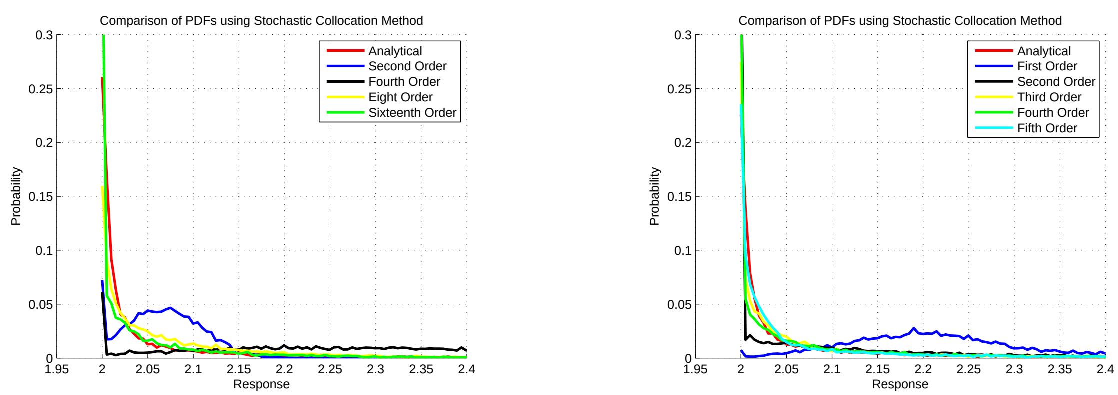
Figure 8. Error in the mean and variance (PSL)



(a) PDF of Monte-Carlo Simulation

(b) PDF (NIS)

Figure 9. PDF : Monte-Carlo and NIS



(a) PDF (PSH)

(b) PDF (PSL)

Figure 10. PDF : PSH and PSL

analysis in engineering environments. However, design procedures that rely on surrogates may consider the PSH method as it provides better functional approximations.

Method	Function Evaluations	Order	L_2 Err.	Mean Err.	Var Err.
NIS	3	1	12.315	0.1712	0.0022
NIS	4	2	7.6345	9.7e-6	0.0073
NIS	6	3	1.5175	2.6e-5	3.3e-4
NIS	9	4	1.5175	2.6e-5	3.3e-4
NIS	11	5	0.5771	3.5e-5	3.0e-7
PSH	2	2	0.0335	0.4845	0.8114
PSH	4	4	0.0246	0.0779	0.4797
PSH	8	8	0.0059	0.0015	0.0273
PSH	16	16	0.0022	4.2e-7	1.2e-4
PSL	2	1	16.699	0.1356	0.0029
PSL	3	2	6.6708	4.9e-4	0.0062
PSL	4	3	1.4703	1.2e-5	3.8e-4
PSL	5	4	1.1021	9.8e-5	4.8e-4
PSL	6	5	0.4808	1.4e-5	1.5e-4

Table 1. Comparison of different methods for the Analytical Problem

B. Aerodynamic Analysis for Normal Random Free-Stream Mach Number

We now assume an input random variable, the free-stream Mach number which is normally distributed with mean, μ 0.75 and standard deviation, $\sigma = 0.0025$ (11) and use PCE expansions to determine the pdf of the aerodynamic drag of a RAE2822 airfoil in inviscid flow.

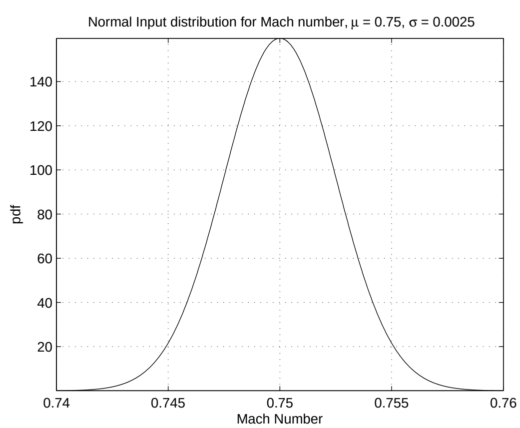
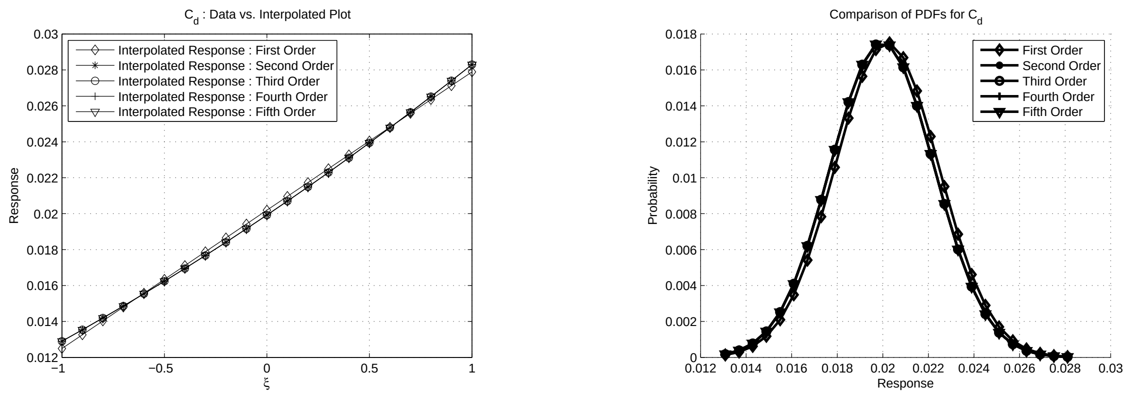


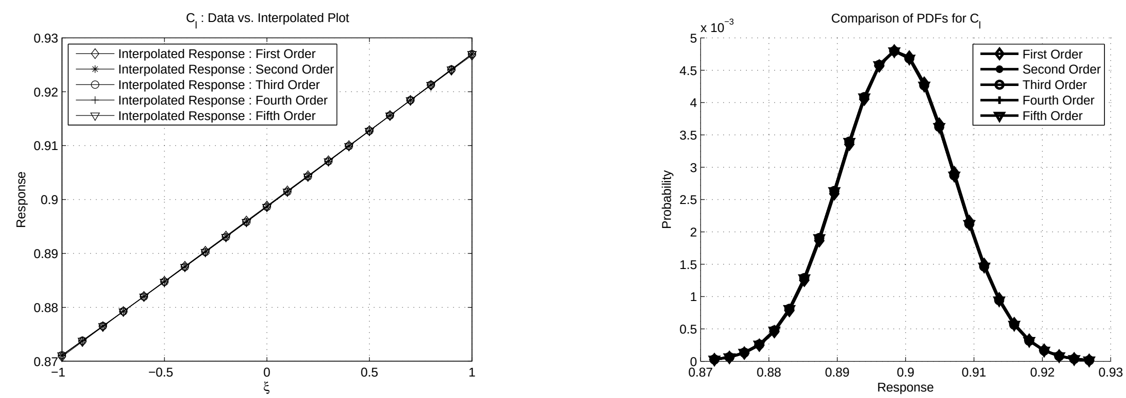
Figure 11. The pdf of the free-stream Mach number

We evaluate the drag using FLO82 at a predetermined set of sampling points for the PSH method. Figures 13 and 12 shows the evaluation of C_d and C_l at a fixed angle-of-attack of 2 degrees. As expected the drag and lift response is roughly linear. This now resembles a linear system and this is further evident in the pdf of C_l and C_d which are roughly Gaussian. Furthermore, PCE expansions above order 2 are sufficient for C_d and even the first order expansion is sufficient for C_l.



(a) Interpolated Fits with PSH for C_d (b) Comparison of PDF with PSH for C_d

Figure 12. Estimates for C_d



(a) Interpolated Fits with PSH for C_l (b) Comparison of PDF with PSH for C_l

Figure 13. Estimates for C_l

C. Multi-Point Design : Unconstrained Lift

We now use the traditional multi-design approach to obtain a baseline optimum solution. We will compare the robust optimization approaches that are the focus of this study against the solution obtained with the multi-point approach. We choose 5 points equally spaced at intervals of 0.005 in the Mach number range of $[0.74, 0.76]$. The optimization algorithm was driven with adjoints and a steepest-descent algorithm and 20 iterations performed in this multi-point design with the intent of reducing C_d. The only constraints that are imposed is that the thickness of the redesigned airfoils do not fall below that of the initial geometry. This is done by projecting the combined gradient into the allowable sub-space. All points in the multi-point design were weighted equally. Table 2 shows the resulting reduction in C_d. Figure 15 shows the pdf of C_d before and after redesign. The pdfs were computed using the NIS method as the sampling points did not follow any quadrature rules. While it is clear that the mean has reduced from 0.0204 to 0.0100, the reduction in σ is less significant (from 0.00772 to 0.00616). Regardless, this reduction of approximately 50% in the mean and around 20% in the standard deviation, illustrates the ability of multi-point optimization to recover robust designs. In general, we would expect multi-point design exercises to reduce the mean of the performance measure while having little control over the standard deviation (or higher moments) unless one tailors the weights of the design points. A sample study using the weights of the Gauss-Hermite quadrature rule for a fourth order method (hence four sampling points) shows a much larger decrease in the mean and the standard deviation. The normalized weights are 0.1267, 0.3685, 0.3685 and 0.1267. The sampling points and the resulting C_d are tabulated in Table 2 and the pdf is shown in Figure 16. The results show that even naive tailoring of the multi-point optimization exercise can help target statistical quantities that add robustness to the design. It is however not clear that the simplicity of this problem (normal distribution of input and approximately linear response of the system)

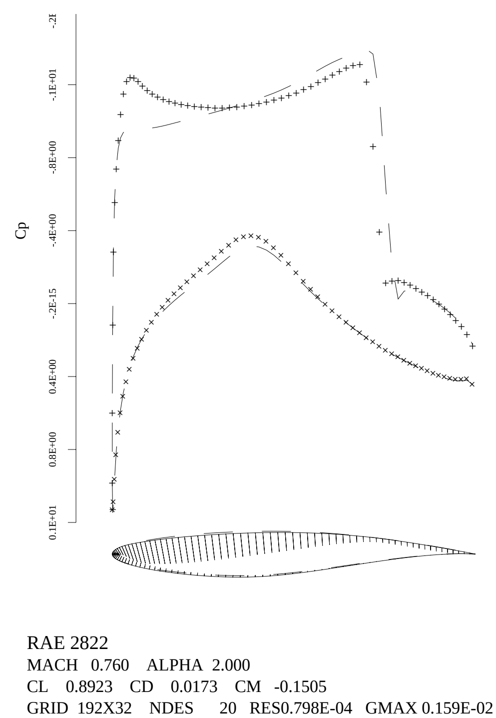


Figure 14. Multi-Point Design : C_p distribution at Mach 0.76 before (—) and after (*,+) the design. The lines on the airfoil suggest the direction of movement by the optimization.

D. PCE-based Design : Unconstrained Lift

In this section, we present results from using the PCE-based approach for robust optimization. For comparison we use the same problem as in the previous sub-section (C) and try to reduce the statistical properties of

Mach	Baseline Geometry	After Optimization with equal weights	After Optimization Gauss-Hermite Weights
0.740	0.0129	0.0050	0.0039
0.745	0.0162	0.0072	0.0042
0.750	0.0199	0.0100	
0.755	0.0240	0.0134	0.0092
0.760	0.028	0.0173	0.0127

Table 2. Multi-Point Design : Drag before and after optimization

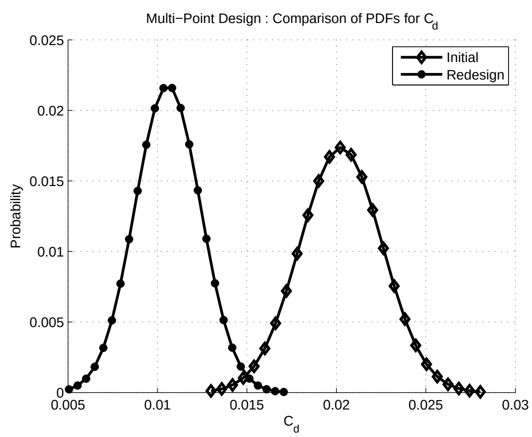


Figure 15. Comparison of pdf (using NIS and third order PCE approximation) before and after the multi-point redesign.

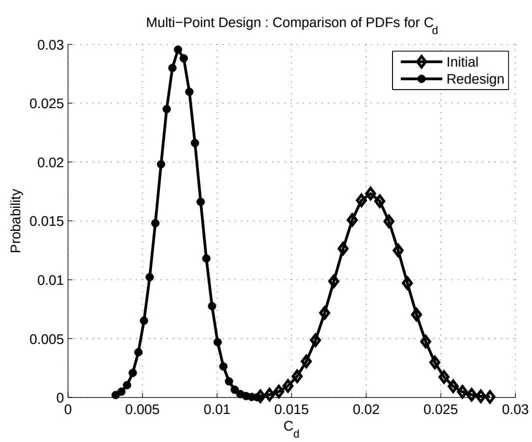
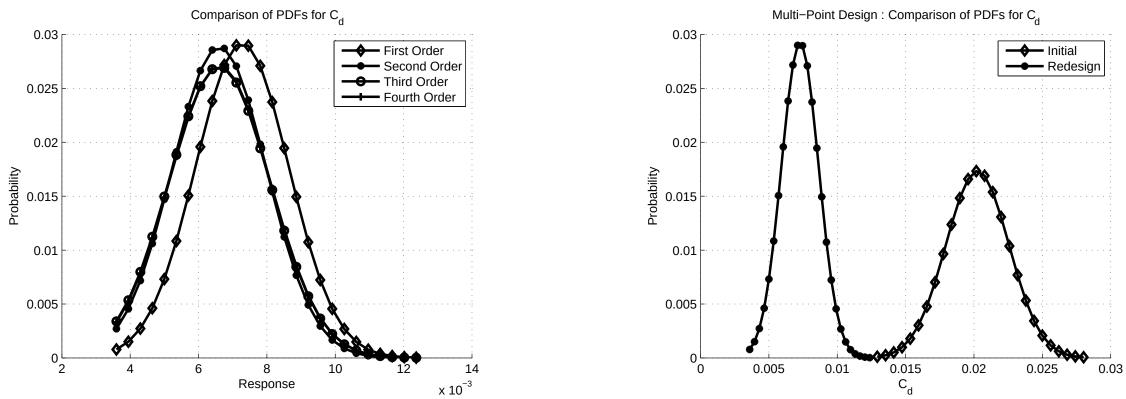


Figure 16. Comparison of pdf (using weights for the Gauss-Hermite Quadrature Rule) before and after the multi-point redesign.

C_d due to the uncertain operating condition of free-stream Mach number (B). To eliminate the dependence of the results on the sampling points, the same points from the multi-point optimization exercise in Section C are used again. As these may not be the quadrature points for the PSH or PSL method, we revert to the NIS approach to construct the PCE of the required quantities. The algorithm for the optimization follows that outlined in Section VI.

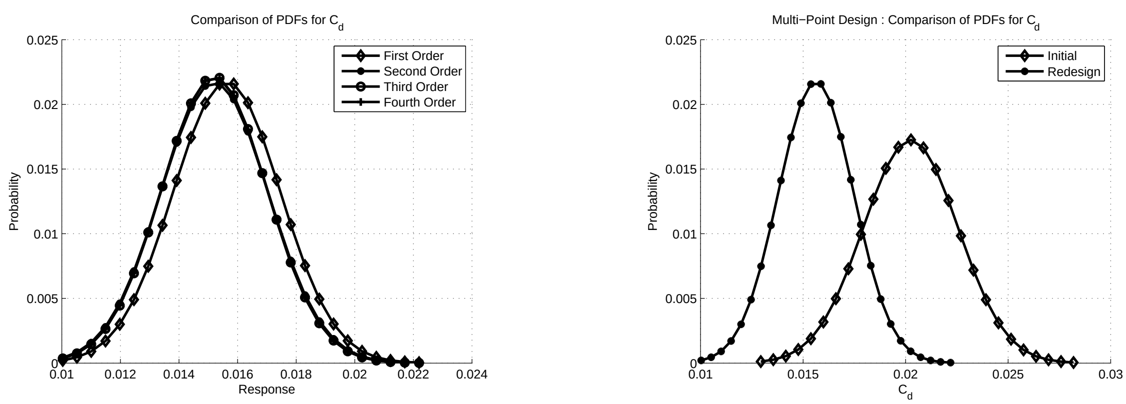
Figure 17 shows the pdf of C_d when the objective was to reduce its mean. For comparison, Figure 18 shows the pdf of C_d when the objective was to reduce its variance. In both calculations 5 points were used for the PCE re-construction. The reduction in the mean and variance when only the mean of C_d is the objective function is comparable to that obtained using the Gauss-Hermite weights for the multi-point optimization. Surprisingly, the reduction in variance obtained using the variance of C_d as the objective function is lower than that obtained with the mean as the objective function. the results at each sampling point are tabulated in Table 3. Figures 19(a),19(b),20(a),20(b) and 21(a),21(b),22(a),22(b) show the pressure distribution before and after the re-design using the mean and variance of C_d as objective function respectively.



(a) Convergence of PDF after optimization for μ of C_d

(b) PDF before and after optimization for μ of C_d

Figure 17. Estimates for C_d



(a) Convergence of PDF after optimization for σ of C_d

(b) PDF before and after optimization for σ of C_d

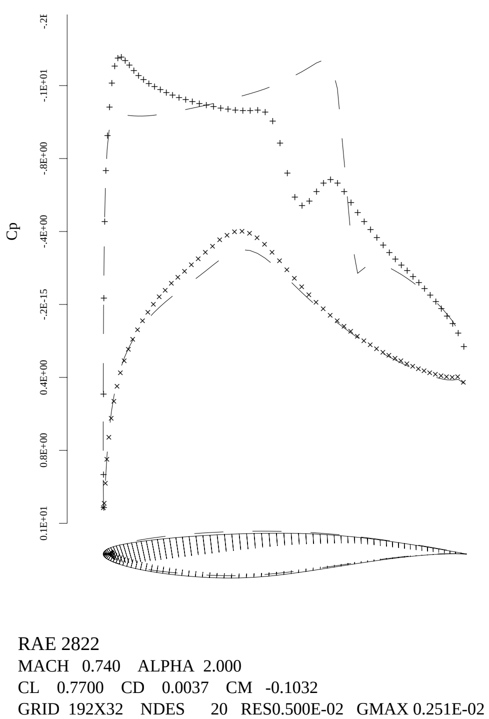
Figure 18. Estimates for C_d

IX. Conclusions

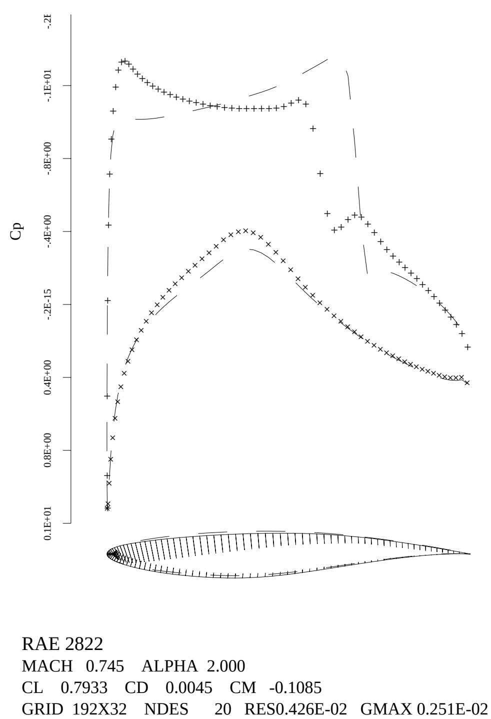
This note outlines the process of efficiently combining non-intrusive PCE based methods UQ methods and adjoint techniques to obtain robust optimal controllers for dynamical systems. The associated cost of the non-intrusive approach in an optimization setting is combination of a UQ analysis and Q adjoint solves for each iteration of a steepest-descent algorithm, where Q is the number of PC coefficients. In the partially intrusive approach, it is possible (with code rewrites) to reduce the number of adjoint solves to as low as 1

Mach	Baseline Geometry	After Optimization with mean	After Optimization variance
0.740	0.0129	0.0036	0.0100
0.745	0.0162	0.0044	0.0124
0.750	0.0199	0.0065	0.0151
0.755	0.0240	0.0093	0.0184
0.760	0.028	0.0126	0.0223

Table 3. PCE-based Design : Drag before and after optimization

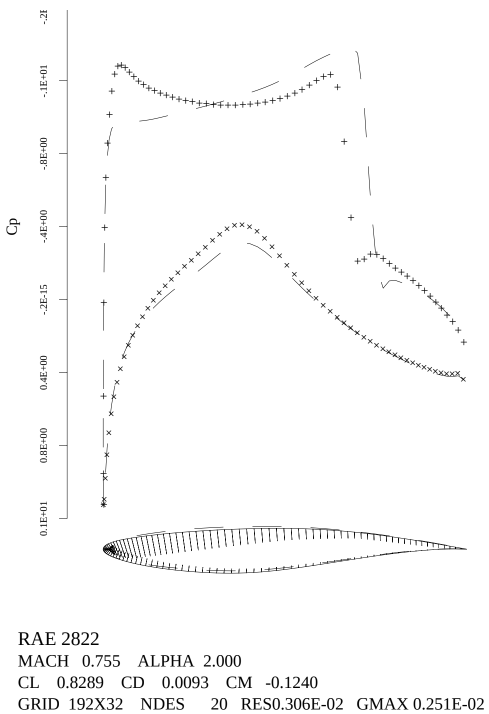


(a) Pressure distribution for $M = 0.74$

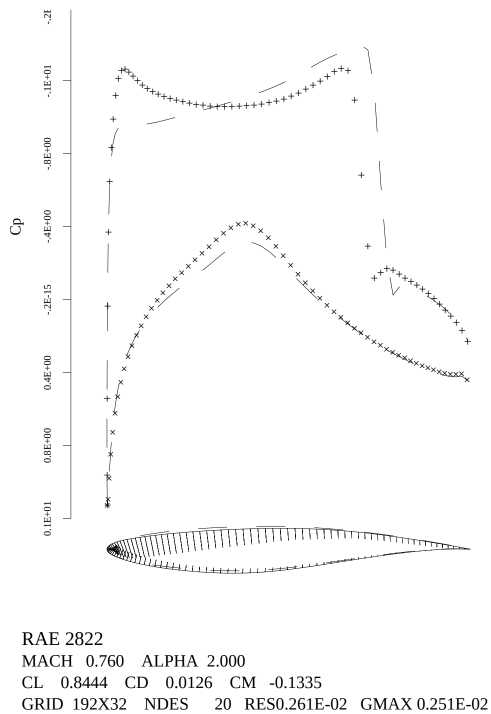


(b) Pressure distribution for $M = 0.745$

Figure 19. Pressure distrubution before (—) and after (*,+) optimization. The lines on the airfoil suggest the direction of movement by the optimization.

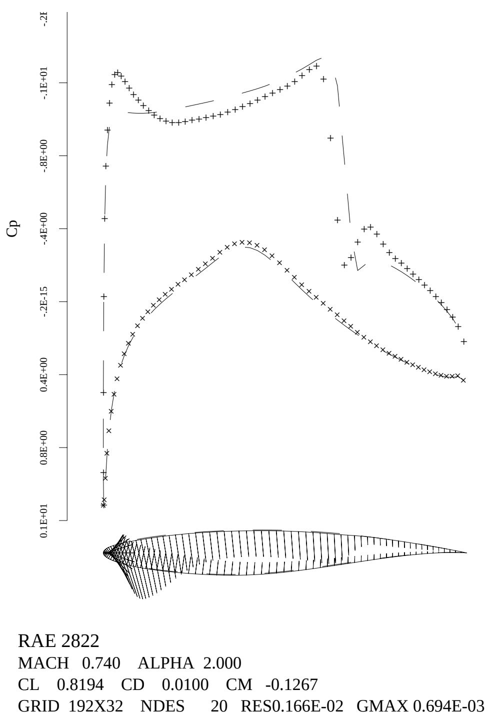


(a) Pressure distribution for $M = 0.755$

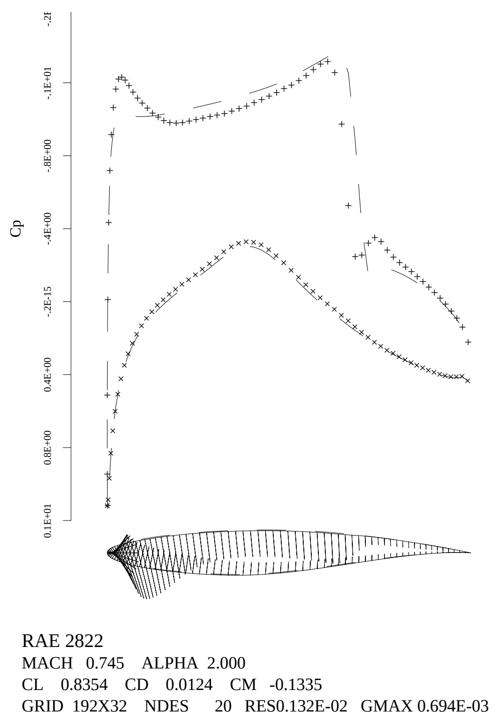


(b) Pressure distribution for $M = 0.760$

Figure 20. Pressure distribution before (—) and after (*,+) optimization. The lines on the airfoil suggest the direction of movement by the optimization.

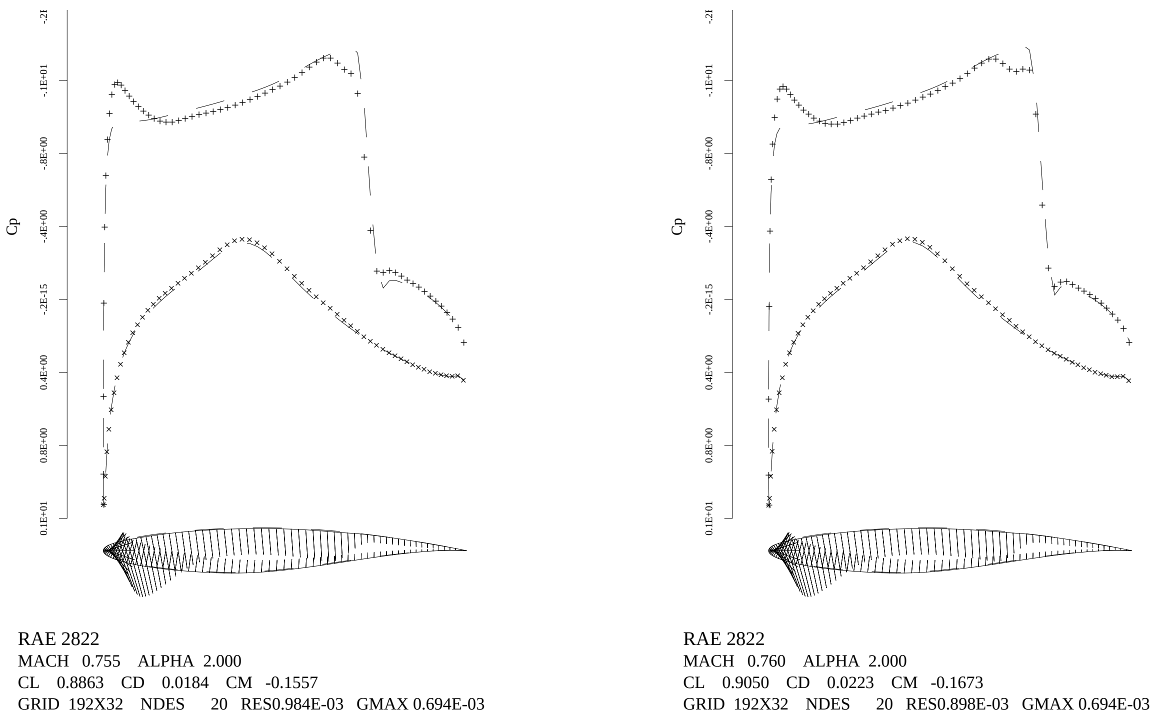


(a) Pressure distribution for $M = 0.74$



(b) Pressure distribution for $M = 0.745$

Figure 21. Pressure distribution before (—) and after (*,+) optimization. The lines on the airfoil suggest the direction of movement by the optimization.



(a) Pressure distribution for $M = 0.755$

(b) Pressure distribution for $M = 0.760$

Figure 22. Pressure distribution before (—) and after (*,+) optimization.The lines on the airfoil suggest the direction of movement by the optimization.

(when the objective is to reduce the mean) and as high as Q when the objective function is higher moments. While still expensive to use routinely in an industrial setting, the ability to assess robustness of designs during the evolution of the design process is invaluable and we hope that some method of this nature will find use in the near future.

X. Acknowledgements

The author would like to thank the GE company for permission to publish this document. The first author would like to thank the GE company for permission to publish this document. The second author would like to thank Dr. Kajal Gupta for support under STTR Phase1 and Phase 2 grants administered by NASA Dryden.

A. Some Common Aerodynamic Objective Functions

A. Expectation of Drag Coefficient

The inviscid drag coefficient for aerodynamic problems can be written as

$$\hat{C}_d = \int_b p c_j n_j dS \quad c_j$$

is the j^{th}direction cosine. Note that this is a linear expression in pressure which greatly simplifies the math. Using the PCE expansion for pressure, p, one can rewrite the above equation as

$$\hat{C}_d = \int_b \sum_{i=1}^{Q} w_i^p \Phi_i(\xi) c_j n_j dS \hat{C}_d = \sum_{i=1}^{Q} \int_b w_i^p \Phi_i(\xi) c_j n_j dS$$

$$\frac{\partial \hat{C}_d}{\partial w_i^p} = \int_b \Phi_i(\xi) c_j n_j dS \tag{24}$$

Hence, if $\hat{I}$ is $\mathbb{E}C_d$ then Equation 19 can be written as

$$\delta \mathbb{E}\hat{C}_d = \mathbb{E}\hat{\mathcal{G}}\delta\mathcal{F}$$

which allows for each of computation of the gradient as just the expectation of Equation 20 which turns out to be just the term corresponding to $i = 1$.

B. Variance of Drag Coefficient

Under the assumption that I and δI are independent, the variation in the variance of I can be simplified as follows:

$$\begin{aligned}
Var(I) &= \mathbb{E}I^2 - \mathbb{E}^2 I & (25)\\
\delta Var(I) &= Var(I + \delta I) - Var(I) & (26)\\
&= \mathbb{E}(I + \delta I)^2 - \mathbb{E}^2(I + \delta I) - Var(I) & (27)\\
&= \mathbb{E}I + 2\mathbb{E}(I\delta I) + (\delta I)^2 - \mathbb{E}^2(X) - 2\mathbb{E}I\mathbb{E}\delta I - \mathbb{E}^2(\delta I) - Var(I) & (28)\\
&= \mathbb{E}(\delta I)^2 - \mathbb{E}^2(\delta I) & (29)\\
&= Var(\delta I) & (30)
\end{aligned}$$

Hence, if $\hat{I}$ is $Var(C_d)$ then Equation 19 can be written as

$$\delta Var(\hat{C}_d) = Var(\delta(\hat{C}_d) = Var(\hat{\mathcal{G}})\delta\mathcal{F}$$

C. Expectation for Inverse Problems

A typical inverse problem is the requirement to meet a specified target loading distribution. The cost function can be written as

$$\begin{aligned}
\hat{I} &= \int_b (\hat{p} - \hat{p_d})^2 dS & (31)\\
&= \int_b \sum_{i=1}^{Q} (w_i^p \Phi_i(\xi) - w_i^{p_d}\Phi_i(\xi))^2 dS & (32)\\
&= \int_b \sum_{i=1}^{Q} \Phi_i(\xi) \left[(w_i^p)^2 - (w_i^{p_d})^2 \Phi_i(\xi) + 2\sum_{i=1}^{Q}\sum_{j=i}^{Q} w_i^p w_j^{p_d} \Phi_j(\xi) \right] dS & (33)\\
&= \sum_{i=1}^{Q} \int_b \Phi_i(\xi) \left[(w_i^p)^2 - (w_i^{p_d})^2 \Phi_i(\xi) + 2\sum_{i=1}^{Q}\sum_{j=i}^{Q} w_i^p w_j^{p_d} \Phi_j(\xi) \right] dS & (34)
\end{aligned}$$

For brevity, lets assume that the target loading is prescribed as the mean. This allows us to drop the coefficients $w_i^{p_d}$ where $i > 1$. (Note that the process of orthogonalization that results in the adjoint system requires multiplication with $\Phi_i(\xi)$ and integration w.r.t. weight functions. As the polynomials are orthogonal few terms in the double summation will remain due to the $\Phi_j(\xi)$ term.) Now the R.H.S. of the adjoint system of equations in Equation 18 can be written as

$$\frac{\partial \hat{I}}{\partial w_i^p} = \int_b 2\Phi_i(\xi)(w_i^p - w_i^{p_d}) dS$$

Now, for each i, we have to multiply the term on the right with $\Phi_i(\xi)$ and integrate with respect to an appropriate weight function. As due to Equation 3, $\int \Phi_i^2(\xi)\rho(\xi)d\xi = 1$, the R.H.S. for the adjoint system for each i is $\int_b 2(w_i^p - w_i^{p_d})dS$. As before, Equation 19 can be written as

$$\delta(\mathbb{E}\hat{I}) = \mathbb{E}\hat{\mathcal{G}}\delta\mathcal{F}$$

which allows for the computation of the gradient as just the expectation of Equation 20 which turns out to be just the term corresponding to $i = 1$.

D. Variance for Inverse Problems

The additional complexity in this case as opposed to the previous sub-section is the need to pre-calculate higher-order integrals of the the polynomials, $\Phi_i(\xi)$. To identify the form of these higher-order terms we write the variance as

$$Var(\hat{I}) = \mathbb{E}(\hat{I})^2 - \mathbb{E}^2(\hat{I}) \tag{35}$$

and repeat the derivation as in the previous subsection.

References

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

[2]Lions, J. L., '*Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, New York, 1971. Translated by S.K. Mitter.

[3]Pironneau, O., *Optimal Shape Design for Elliptic Systems*, Springer-Verlag, New York, 1984.

[4]Jameson, A., *Aerodynamic design via control theory*, Journal of Scientific Computing, 3:233–260, 1988.

[5]Xiu, D, *Fast Numerical Methods for Stochastic Computations: A Review*, Communications in Computational Physics, Vol 5, No 2-4, pp 242-272, 2009.

[6]Spanos, P., and Ghanem, R., *Stochastic finite element expansion for random media*, Journal of Engineering Mechanics, ASCE, Vol. 115, No. 5, pp. 1035-1053, May 1989.

[7]Ghanem, R., and Spanos, P., *Polynomial chaos in stochastic finite element*, Journal of Applied Mechanics, ASME, Vol. 57, No. 1, pp. 197-202, March 1990.

[8]Xiu, D. and Karniadakis, G. E., *The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations*, SIAM J. Sci. Comput., 24(2), 619-644, 2002.

[9]Xiu, D. and Karniadakis, G. E., *Modeling Uncertainty in Flow Simulations via Generalized Polynomial Chaos*, J. Comput. Phys. 187, 137-167, 2003.

[10]Wang, D. X. and He, L., *Coupled aerodynamic and aeromechanic design optimization by using the continuous adjoint method*, Proc., 7th Asian Computational Fluid Dynamics Conference, Bangalore, Nov, 2007.

[11]Sriram, *Robust Optimal Control of Systems with Uncertain Inputs*, GE Global Research Internal Report, March 2010 (please contact the first author for a copy).