Goal-Oriented Uncertainty Propagation Using Stochastic Adjoints

Karthikeyan Duraisamy*, Juan Alonso†

Stanford University, Stanford CA 94305, USA

Praveen Chandrashekar‡

Tata Institute of Fundamental Research, Bangalore 560065, India

We propose a framework based on the use of adjoint equations to formulate an adaptive sampling strategy for uncertainty quantification for problems governed by algebraic or differential equations involving random parameters. The approach is non-intrusive and makes use of discrete sampling based on a collocation on simplex elements in stochastic space. Adjoint or dual equations are introduced to estimate errors resulting from possible inexact reconstruction of the solution within the simplex elements. The approach is demonstrated to be accurate in estimating error in statistical moments of interest and shown to exhibit super-convergence, in accordance with the underlying theoretical rates. Goal-oriented error indicators are then built using the adjoint solution and used to identify regions for adaptive sampling. The error-estimation and adaptive refinement strategy is applied to a range of problems including those governed by algebraic equations as well as scalar and systems of ordinary and partial differentiatial equations. The strategy holds promise as a reliable method to set and achieve error tolerances for efficient uncertainty quantification in complex problems.

I. Introduction

With the advances in computational technology in recent years, mathematical models are being increasingly called upon to predict a variety of phenomena relevant to science and engineering. When applied to complex systems, predictive models are subject to a range of uncertainties, arising either from natural variabilities present in the system, or from an improper knowledge of the system and the conditions influencing it. The process of quantifying such uncertainties and characterizing their manifestations on the output quantities of interest is being recognized to be very important in all areas of computational science and has emerged as an important discipline termed *uncertainty quantification* (UQ)[1,2]. In the present work, the aspect of uncertainty that is due to the variability in the system will be considered. Otherwise referred to as *aleatoric uncertainty*, this is concerned with the propagation of uncertainty in system parameters to system outputs, with the variability in the parameters assumed to be representable by known probability distributions.

Uncertainty quantification approaches can be classified as *intrusive*[3–6] or *non-intrusive*[7–10]. The former involves a reformulation of the governing equations by projecting them in stochastic (probability) space and typically results in a one-time solution that can be post-processed to recover the statistics of interest. The latter typically involves deterministic or random sampling of the governing equations in stochastic space. While there are advantages and disadvantages[2] to either approach, in this work we consider a non-intrusive method. Intrusive methods like stochastic finite element methods[3], are inherently difficult to implement as a reformulation of the governing equations leading to a large system of coupled PDE, development of new numerical schemes and also a complete rewrite of complicated codes are required. Typically, the goal of the UQ analysis is to estimate the probability density of some random output functional which also involves estimating certain statistical moments like the mean and variance. In this work, recent ideas on adjoint based techniques for numerical solution of partial differential equations using goal-oriented grid adaptation[11,12] are extended to stochastic space and a new framework is introduced to estimate and control the numerical error in calculating some statistical moments, which we will sometimes refer to as the objective function. In other words, an approach to *manage* the error introduced due to an imperfect characterization of the output in stochastic space will be presented. This approach requires development of adjoint solvers, which might be considered as an intrusive approach; however the use of automatic differentiation (AD)[13] tools greatly simplifies the development of adjoint solvers for complex simulation codes. Moreover, with increasing use of adjoint-based optimization in design, it is usually the case that an adjoint solver is already available.

The main contribution of the paper is the presentation of a new method for accurate estimation of statistical moments of a random functional of interest using adjoint solutions. The use of adjoint solutions is shown to double the accuracy of the moments leading to superconvergence. The adjoint approach also gives error estimates for the

* Consulting Assistant Professor, Department of Aeronautics & Astronautics, dkarthik@stanford.edu

† Associate Professor, Department of Aeronautics & Astronautics, jjalonso@stanford.edu

‡ Fellow, Center for Applicable Mathematics, praveen@math.tifrbng.res.in

functional that are used to adaptively sample the solution in stochastic space leading to a goal-oriented adaptive approach to uncertainty quantification. The proposed approach can be used either in a continuous or a discrete adjoint framework; in the present work, we use the discrete approach due to the ease with which discrete adjoint solvers can be developed for complex simulation codes, e.g., using AD.

The rest of the paper is organized as follows. In section II and III, we give the formulation of the problem of estimating statistical moments of a random functional of interest and introduce the adjoint equation. Section IV discusses the use of adjoint solutions to reduce the error in the objective functional while section V provides an adaptive sampling procedure. In section VI, we apply these ideas to simple model problems and the two-dimensional Euler equations involving one and two random variables. Section VII discusses some extensions and further generalization of this apprach.

II. Stochastic Expansions and Sampling

Consider a system of governing equations written in a compact form as:

$$R(\vec{u}(\vec{x}, t, \vec{\xi})) = 0, \tag{1}$$

where $\vec{u}$ represents the state vector of the unknowns in space $\vec{x}$ and time t, and the quantity $\vec{\xi}$ represents the aleatoric or uncertain variables. We assume that $\vec{\xi} \in \Omega$ is an N_ξ-dimensional vector with a precisely specified probability structure. For purposes of clarity, symbols signifying spatial and temporal dependence are omitted (as is the vector notation) without any loss of generality of the formulation. Let $I(u(\xi))$ be a functional of interest; the goal of UQ is typically to estimate the statistical moments of this random output functional. In the rest of the paper, we will take the objective function to be the mean of the random functional defined as

$$\mathcal{J}(u) = \int_\Omega I(u(\xi))\mathrm{d}\xi \tag{2}$$

The random variables are assumed to be uniformly distributed. Extensions to other random variables is easy and is described in Section (VII.); conceptually, this amounts to taking $\mathrm{d}\xi$ to be the appropriate probability measure of the random variables involved. In order to compute the above integral, we will assume that the random space Ω is divided into N_E simplex elements consisting of N_S vertices. The stochastic problem is converted to a deterministic problem corresponding to the N_S realizations of the random variables and the objective function is computed by performing a quadrature on the simplex elements

$$\mathcal{J}(u) = \sum_{i=1}^{N_E} \int_{E_i} I(u(\xi))\mathrm{d}\xi, \tag{3}$$

where, $\{E_i\}$ represents a division of the random space Ω into N_E simplex elements. However, the random solution $u(\xi)$ is available only at the vertices of the simplex elements. In order to perform the quadrature, we construct an approximation $\tilde{u}(\xi)$ using the solution at the vertices of the simplex elements using finite element techniques. In essence, based on N_S samples, an unstructured mesh composed of N_E elements in N_ξ dimensions is built, similar to the concept of Stochastic Finite Elements[3,14]. The accuracy in computing the objective function depends on how closely the reconstructed solution $\tilde{u}(\xi)$ within these elements is able to represent the true distribution $u(\xi)$.

III. Formulation of adjoint problem in stochastic space

Each sample in stochastic space $u(\xi_j)$, $j = 1, \ldots, N_S$ is obtained by solving eqn.1, and hence the residual $R(u(\xi_j)) = 0$ at the vertices of the elements. If the reconstructed solution $\tilde{u}(\xi)$ is not exact, then its residual will not be identically zero in Ω. In recognition of this, a global adjoint problem is defined in Ω. Using the theory of adjoint-based error estimation[11,12], it would then be possible to explicitly derive the relationship between locally non-zero residuals and the global error in computing the objective function.

Given a non-linear *forward* equation $R(u(\xi)) = 0$ with $\xi \in \Omega$ and a *scalar* objective function $\mathcal{J}(u) = \int_\Omega I(u(\xi))\mathrm{d}\xi$, the governing equations for the adjoint variable v can be constructed[15] as

$$\mathcal{L}_u^* v = -g(u), \tag{4}$$

where $\mathcal{L}_u$ is the Frechet derivative of R. Assuming $\mathcal{M}_u$ to be the Frechet derivative of $\mathcal{J}$, the adjoint operator $\mathcal{L}_u^*$ and the source term $g(u)$ are given by

$$\begin{aligned} (\mathcal{L}_u \hat{u}, v)_\Omega &= (\hat{u}, \mathcal{L}_u^* v)_\Omega & (5)\\ \mathcal{M}_u \hat{u} &= (g(u), \hat{u})_\Omega, & (6) \end{aligned}$$

where, $(\cdot,\cdot)$ is an inner product depending on the function spaces to which u, v belong. In the discrete approach that we follow here, the inner product is the usual Euclidean inner product[17]. Since there is no cross-coupling between different locations in Ω, the solution of the global adjoint problem reduces to the individual solutions of the adjoint equations at each sampling location.

IV. Error estimation procedure

The error estimation procedure used in the present work is based on a discrete version of the conceptual adjoint problem introduced in the previous section. Thus R in equation (1) represents the system of discrete equations obtained by applying a numerical scheme to the differential equations governing the system. In computing the statistical average (2), the integral is approximated using a quadrature on simplex elements which is exact for polynomials of degree s,

$$\mathcal{J} \approx \sum_{i=1}^{N_E} \sum_{j=1}^{N_q} w_{ij} I(u(\xi_{ij})) + O(\Delta\xi^{s+1}),$$

where ξ_{ij} is the j^{th} Gaussian quadrature point in the element i. For purposes of clarity, the subscripts $(.)_{ij}$ will be used to denote the value of the quantity $(.)$ at the quadrature point ξ_{ij}. Also, summations over i and j imply looping through the elements and quadrature points within the elements, respectively. The quantities I and R can depend explicitly on the random variables without changing the formulation, but we do not indicate this dependence for purpose of clarity.

Note that $u(\xi)$ is known only at the vertices of the simplex, $\xi_j, j = 1, \ldots, N_S$. Using these vertex values of u we reconstruct the solution $\tilde{u}(\xi)$ in random variables, which would typically be a linear or quadratic reconstruction. Assuming I is a smooth function of u,

$$\begin{aligned} I(u_{ij}) &= I(\tilde{u}_{ij}) + [I(u_{ij}) - I(\tilde{u}_{ij})] \\ &= I(\tilde{u}_{ij}) + \frac{\partial I}{\partial u}(\tilde{u}_{ij})(u_{ij} - \tilde{u}_{ij}) + O(\|u - \tilde{u}\|^2) \end{aligned}$$

Assuming smoothness of u with respect to the random variable ξ, if $\tilde{u}$ is exact for polynomials of degree r, then we can expect $\|u - \tilde{u}\| = O(\Delta\xi^{r+1})$ where $\Delta\xi$ is a measure of the stochastic element size, e.g., the length of the largest side. Since

$$R(u_{ij}) = 0$$

we can write[§]

$$\begin{aligned} I(u_{ij}) &= I(\tilde{u}_{ij}) + \left\{ \frac{\partial I}{\partial u}(\tilde{u}_{ij})(u_{ij} - \tilde{u}_{ij}) + v_{ij}^T [R(u_{ij}) - R(\tilde{u}_{ij}) + R(\tilde{u}_{ij})] \right\} \\ &\quad + O(\|u - \tilde{u}\|^2) \\ &= I(\tilde{u}_{ij}) + \left[\frac{\partial I}{\partial u}(\tilde{u}_{ij}) + v_{ij}^T \frac{\partial R}{\partial u}(\tilde{u}_{ij}) \right] (u_{ij} - \tilde{u}_{ij}) + v_{ij}^T R(\tilde{u}_{ij}) \\ &\quad + O(\|u - \tilde{u}\|^2) \end{aligned}$$

It has to be mentioned that this procedure mirrors the numerical error estimation procedure of Venditti et al.[12], but the reconstruction and expansion is performed in probability space instead of within a spatial mesh element. If we choose v_{ij} to be the solution of the adjoint equation

$$\frac{\partial I}{\partial u}(\tilde{u}_{ij}) + v_{ij}^T \frac{\partial R}{\partial u}(\tilde{u}_{ij}) = 0,$$

then the corrected functional is

$$I(u_{ij}) = I(\tilde{u}_{ij}) + v_{ij}^T R(\tilde{u}_{ij}) + O(\|u - \tilde{u}\|^2)$$

This, however, requires the solution of adjoint equations at the quadrature points which would require possibly unjustifiable extra work. Using a solution of the adjoint equations $v(\xi)$ at the vertices of the simplex elements, we propose to use them to reconstruct a higher order representation in stochastic space, $\tilde{v}(\xi)$. Then the corrected functional is

$$\begin{aligned} I(u_{ij}) &= I(\tilde{u}_{ij}) + \tilde{v}_{ij}^T R(\tilde{u}_{ij}) \\ &\quad + \left[\frac{\partial I}{\partial u}(\tilde{u}_{ij}) + \tilde{v}_{ij}^T \frac{\partial R}{\partial u}(\tilde{u}_{ij}) \right] (u_{ij} - \tilde{u}_{ij}) + (v_{ij} - \tilde{v}_{ij})^T R(\tilde{u}_{ij}) \\ &\quad + (v_{ij} - \tilde{v}_{ij})^T \frac{\partial R}{\partial u}(\tilde{u}_{ij})(u_{ij} - \tilde{u}_{ij}) + O(\|u - \tilde{u}\|^2) \end{aligned}$$

The second term on the right hand side involves known quantities and is termed the computable correction (CC), whereas the third and fourth terms involve the unknown exact solutions u_{ij}, v_{ij} and constitutes the remaining error (RE). Again, assuming smoothness properties, if $\tilde{v}$ is also exact for polynomials of degree r, then the fifth term is of similar order as the neglected term of $O(\|u - \tilde{u}\|^2)$, and hence can be ignored. The mean value of the functional is given by

$$\mathcal{J} = \sum_{i=1}^{N_E} \sum_{j=1}^{N_q} w_{ij} \{ I(\tilde{u}_{ij}) + \tilde{v}_{ij}^T R(\tilde{u}_{ij}) \} + RE + O(\Delta\xi^{2(r+1)}) + O(\Delta\xi^{s+1}) \tag{7}$$

[§]As I is a scalar, we use the convention that $\frac{\partial I}{\partial u}$ is a row vector.

In the present approach, the second term is used to correct the predictions of the objective function and an estimate of the remaining error is used to drive the refinement in stochastic space (adaptive sampling) in an attempt to reduce the reconstruction error. Since $\tilde{u}$, $\tilde{v}$ are approximations to the primal and adjoint solutions, we can expect the remaining error to be small, especially because of the targeted refinement. If RE is sufficiently small, then the corrected objective function with the error estimate based on CC is capable of providing super-convergence.

Using only the computable terms, the corrected mean value of the functional is given by

$$\mathcal{J} \approx \sum_{i=1}^{N_E} \sum_{j=1}^{N_q} w_{ij} \{ I(\tilde{u}_{ij}) + \tilde{v}_{ij}^T R(\tilde{u}_{ij}) \} =: J + CC \tag{8}$$

If the remaining error is small, then from the preceding derivation, the estimate J converges at the rate $(r+1)$ while the adjoint corrected estimate $J + CC$ converges at the rate $2(r+1)$. In later sections, we provide numerical evidence of this super-convergence property of adjoint corrections. Note that the accuracy of the numerical quadrature must be sufficiently high so that quadrature error is smaller than the error of reconstruction, which requires using a quadrature rule so that $s \geq 2r+1$.

V. Adaptation procedure

The remaining error which is not precisely computable can be used as an error indicator to decide where to perform additional sampling. Since we do not know the exact solutions u_{ij}, v_{ij} we need to estimate the terms $(u_{ij} - \tilde{u}_{ij})$ and $(v_{ij} - \tilde{v}_{ij})$. This can be achieved by using two reconstruction operators of different accuracy[12]. Let L and Q denote linear and quadratic reconstruction operators in stochastic space. Then an estimate of remaining error in element i is given by

$$\epsilon_i = \left| \sum_{j=1}^{N_q} w_{ij} \left\{ (Qv_{ij} - Lv_{ij})^T R(\tilde{u}_{ij}) + [R_{ij}^*(\tilde{v}_{ij})]^T (Qu_{ij} - Lu_{ij}) \right\} \right| \tag{9}$$

where

$$R_{ij}^*(v) := \left[\frac{\partial I}{\partial u}(\tilde{u}_{ij}) \right]^T + \left[\frac{\partial R}{\partial u}(\tilde{u}_{ij}) \right]^T v$$

and Qu_{ij} etc. mean that the reconstructed solution is evaluated at the quadrature point ξ_{ij}, e.g., $Qu_{ij} = (Qu)(\xi_{ij})$. The total remaining error is estimated as a sum of element errors ϵ_i

$$RE = \sum_{i=1}^{N_E} \epsilon_i \tag{10}$$

which can be used to set error tolerance for adaptation. The element error indicator ϵ_i is then used as an adaptation indicator. If any element has a large local error ϵ, then it can be divided into two or more elements by adding new sample points.

In the present work, we implement an adaptive sampling technique based on remaining error as follows. In all the examples, we take the case of two independent random variables, and assume that the random variable space $\Omega \subset \mathbb{R}^2$ is a rectangle. The following algorithm is illustrated in figure (1).

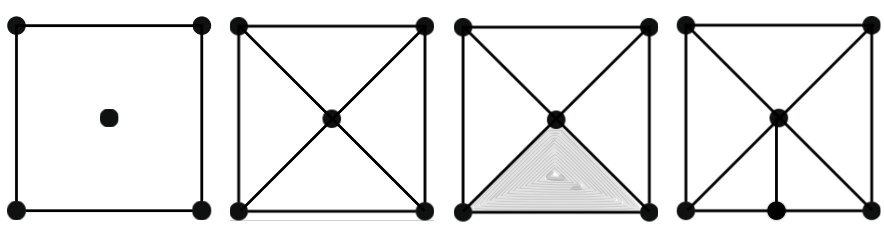

Figure 1. Adaptive refinement of simplex elements: The shaded element is flagged for refinement and is divided into two elements by adding a new sample in the middle of its largest side.

Adaptive sampling algorithm:

1. Sample the primal solution $u(\xi)$ and adjoint solution $v(\xi)$ at the four vertices and center of Ω, i.e., $N_S = 5$
2. Construct a grid of simplex elements using the N_S samples
3. Construct the stochastic approximations $\tilde{u}(\xi)$ and $\tilde{v}(\xi)$ to the primal and adjoint solutions

4. Compute J and CC from equation (8)

5. For each simplex element, compute the remaining error from equation (9)

6. Choose the element with the largest value of the element error and add a new sample at the middle of the largest side

7. Set N_S = N_S + 1 and go to step 2

The above algorithm can be modified to add more than one sample in each iteration depending on the available computational resources. Note that the availability of an error estimate (and an estimate of the total remaining error (equation 9)) presents an opportunity to set an error tolerance and this can serve as a stopping criterion. We could, for instance, choose to stop the refinement when the error estimate or the estimate of the remaining error is reduced below a user-specified percentage of the objective function. Note that this is possible only because the error estimate has the same units as that of the objective function (which is unique to adjoint based methods). For problems in which the randomness is large, a combination of error indicators based on mean and variance of the random functional can also be used. It has to be noted, however, that errors resulting from spatial and/or temporal discretization at each sample point could potentially overwhelm the error in the objective functional, and hence, these have to be accounted for. Although this aspect is not addressed in the present work (all the results presented are grid converged), the availability of the adjoint solution at each sample point may be used to estimate the error in $I(u)$ and ensure that it remains a very small fraction of $I(u)$ using adaptive spatial and/or temporal refinement if necessary.

Remark

Even if the numerical quadrature uses linear reconstructions $\tilde{u}, \tilde{v}$, the estimation of remaining error requires both linear and quadratic reconstructions, which is not an optimal approach. However, we would like to demonstrate that the theoretical convergence rates are achieved with respect to the number of samples for both linear and quadratic reconstructions, and the resulting superconvergence when adjoint corrections are added. In practice, one would use quadratic reconstruction for better accuracy and faster convergence, provided the primal and adjoint solutions have sufficiently smooth dependance on the random variables, see also the remarks in Section (VII.). Alternately, when using linear reconstruction in stochastic space, one can choose a constant in element approximation as the lower order reconstruction L and linear reconstruction as the higher order reconstruction Q, and then estimate the remaining error.

VI. Numerical test cases

To demonstrate the super-convergence properties of the adjoint-based correction and the adaptation procedure, we use four test cases, the first three of which involve simple scalar differential equations while the fourth is based on the solution of the two dimensional Euler equations.. The first two examples involve algebraic governing equations whereas the third is a convection-diffusion problem. The discrete model is assumed to be the *truth* model, i.e., we do not consider the error arising in the numerical discretization, but only aim to reduce the error of stochastic integration. For linear reconstruction, the values at the simplex vertices are used to define a linearly varying function which corresponds to P_1 finite element interpolation. Quadratic reconstruction is performed using the values and gradients at the simplex vertices, the gradients being computed using a least-squares based procedure [16]. Integration within each simplex element is performed using the symmetric Dunavant quadrature rule [19,20]. The quadrature points are located inside the simplex elements; for linear reconstruction it is enough to use four quadrature points while quadratic reconstruction requires seven quadrature points. This choice ensures that the quadrature error is less than the error of interpolation, so that we can observe the theoretical convergence rates with respect to the number of samples.

To verify the convergence rate of the mean of the random functional, a uniform refinement of the simplex elements is first performed. The domain Ω of the random variables in two dimensions is taken to be rectangular and is initially divided into four simplex elements, by sampling at the four corners and the center, see figure (1). In every refinement step, a new sample is added at the middle of each edge. The number of samples is thus 5, 13, 41, 145 and 545 when performing four levels of uniform refinement. Since we are dealing with two random variables, the error is expected to decrease as $N_S^{-p/2}$ where p is the rate of convergence of the objective function. To compute the convergence rate, a straight line is fit to log(Error) as a function of log($\sqrt{N_S}$), the slope of which indicates the value of p. For linear reconstruction, the theoretical rate is p = 2 and p = 4 for the uncorrected and corrected functionals respectively, while for the quadratic reconstruction, it is p = 3 and p = 6 respectively.

A. Algebraic example I

The governing equation is given by

$$R(u, \xi_1, \xi_2) := \frac{1}{2}u^2 - \xi_1 u - \frac{1}{2}\exp(-50(\xi_2 - 1/2)^2) = 0$$

where the random variables $\xi_1 \in [1/2, 1]$, $\xi_2 \in [0, 1]$ are uniformly distributed. The random functional of interest is $I(u) = u^2$ and we would like to estimate

$$\mathcal{J} = \int_0^1 \int_{1/2}^1 I(u) \mathrm{d}\xi_1 \mathrm{d}\xi_2$$

Figure (2) shows the contour plot of I as a function of the random variables, which indicates a sensitive dependence on ξ_2. Figure 3 shows the convergence properties under uniform refinement; for the linear reconstruction, the convergence rates with and without error estimation are 2.7 and 4.8, which are close to the theoretical values of 2 and 4. For the quadratic reconstruction, the values are 3.7 and 6.4 which are again close to the theoretical values of 3 and 6. Figure 4 shows the effectiveness and accuracy of the adaptive refinement procedure; perhaps quite striking is the highly accurate recovery of the error by the estimation procedure, even when the total number of samples is small. The convergence with adaptive refinement is not monotonic until a sufficient number of samples are used (Figure 4). The adjoint corrected functional is seen to converge more monotonically compared to the uncorrected functional. The error recovery is seen to be very good for $N_S > 60$. The simplex grid shows that the error indicator is able to identify the regions of the random space where the functional is more sensitive leading to more refinement in those regions. Figure 5 compares the reduction of error with increasing samples. The benefits of the error correction in both the uniform and adaptive refinement is quite clear. The adjoint-based adaptive refinement/error correction strategy is seen to be advantageous. Also shown for comparison is a Richardson-based error-estimation strategy, which is, in this particular case, not accurate for a low number of samples. It is possible that the Richardson extrapolation will provide improved results in the asymptotic convergence range.

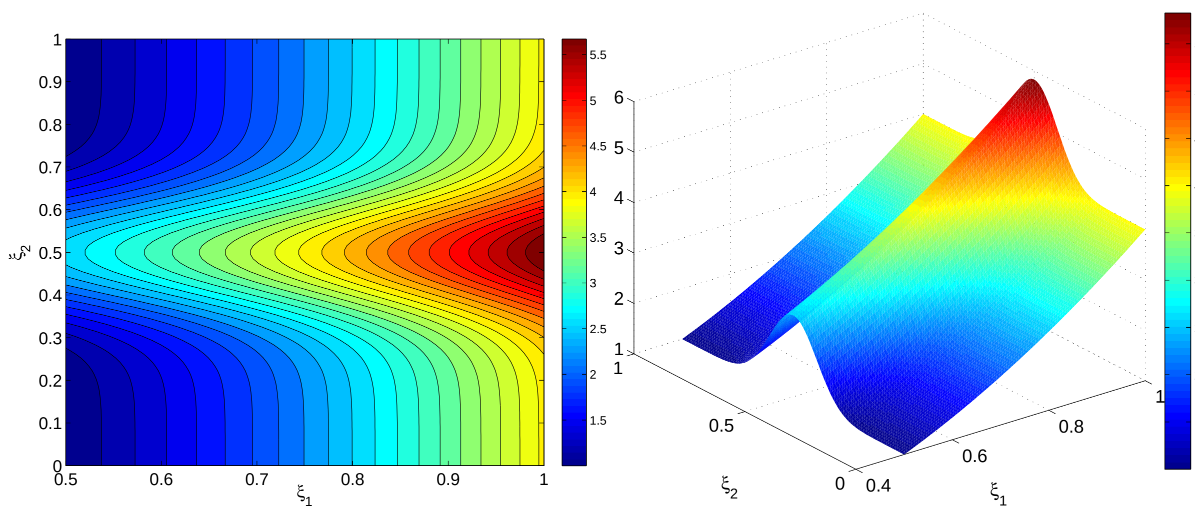

Figure 2. Random functional I for algebraic example I

B. Algebraic problem II

To assess the applicability and performance of the present approach in the case of a discontinuous response surface, the following governing equation is considered:

$$R(u, \xi) := \frac{1}{2}u^2 - \frac{3}{4}u - \frac{a(\xi)}{2} = 0$$

where the random variable $\xi \in [0, 1]$, is uniformly distributed and $a(\xi)$ is discontinuous and assumes the values $\sin \pi\xi$ for $\xi < 0.6$ and $\sin \pi\xi/2$ for $\xi > 0.6$. The objective functional is given by

$$\mathcal{J} = \int_0^1 u^2 \mathrm{d}\xi$$

The primal and adjoint solutions are depicted in Figure 6 and both of them are discontinuous. Figure 7 shows the convergence of the error in the mean functional with the number of samples. The legends P_1 and P_2 refer to linear

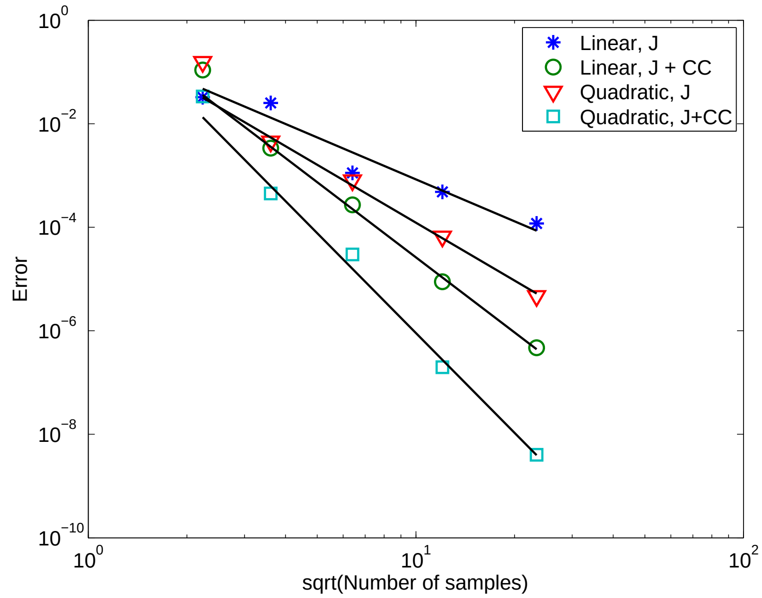

Figure 3. Convergence rates for algebraic example I under uniform refinement: $(*)$ **2.7,** $(\circ)$ **4.8,** $(\triangledown)$ **3.7,** $(\square)$ **6.4**

and quadratic reconstruction of the primal and adjoint solutions. It is again seen that the adjoint-based error correction strategy shows great benefit, especially for the linear reconstruction approach. The adaptive refinement is clearly superior to uniform refinement while Richardson extrapolation is not effective. Similar to the Burgers example, however, the error correction from the quadratic reconstruction does not prove to be more advantageous than the linear reconstruction, thus warranting a non-oscillatory scheme for reconstruction in stochastic space to handle discontinuities.

C. Stochastic convection-diffusion equation

Consider the steady viscous Burgers equation with a random forcing term given by

$$\begin{aligned} uu_x = u_{xx} + s, \qquad & x \in (0,1) \\ u(0) = u(1) = 0 & \end{aligned} \tag{11}$$

The source term $s(x;\xi_1,\xi_2)$ is chosen so that the exact solution is

$$u(x;\xi_1,\xi_2) = 10x(1-x)\sin(\pi(\xi_1 + \xi_2 x))$$

where $\xi_1, \xi_2 \in [0.9, 1.1]$ are uniform random variables. This problem can be considered as a simplified model of Navier-Stokes equations since it has non-linear convection and diffusion effects. The ODE (11) is approximated using a finite volume method on M cells so that the governing equations $R : \mathbb{R}^M \times \Omega \to \mathbb{R}^M$ are given by

$$R(u)_i := \frac{F_{i+1/2} - F_{i-1/2}}{\Delta x} - \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2} - s_i = 0, \quad 1 \le i \le M$$

where $F_{i+1/2} = F(u_i, u_{i+1})$ is a numerical flux function, e.g., the Lax-Friedrich's flux defined as

$$F(u,v) = \frac{1}{2}(f(u) + f(v)) - \frac{1}{2}\lambda(v-u) \tag{12}$$

where we choose $\lambda = 1$ which leads to stable solutions. The boundary conditions are implemented by appropriate definition of the boundary fluxes. The random functional of interest is

$$I = \Delta x \sum_{i=1}^{M} u_i^2 \approx \int_0^1 u^2 \mathrm{d}x$$

and in the computations we use $M = 20$ cells. The functional is plotted in figure (8) as a function of the random variables. We are interested in the mean of the random functional $I(\xi)$ given by

$$\mathcal{J} = \int_{0.9}^{1.1} \int_{0.9}^{1.1} I \mathrm{d}\xi_1 \mathrm{d}\xi_2$$

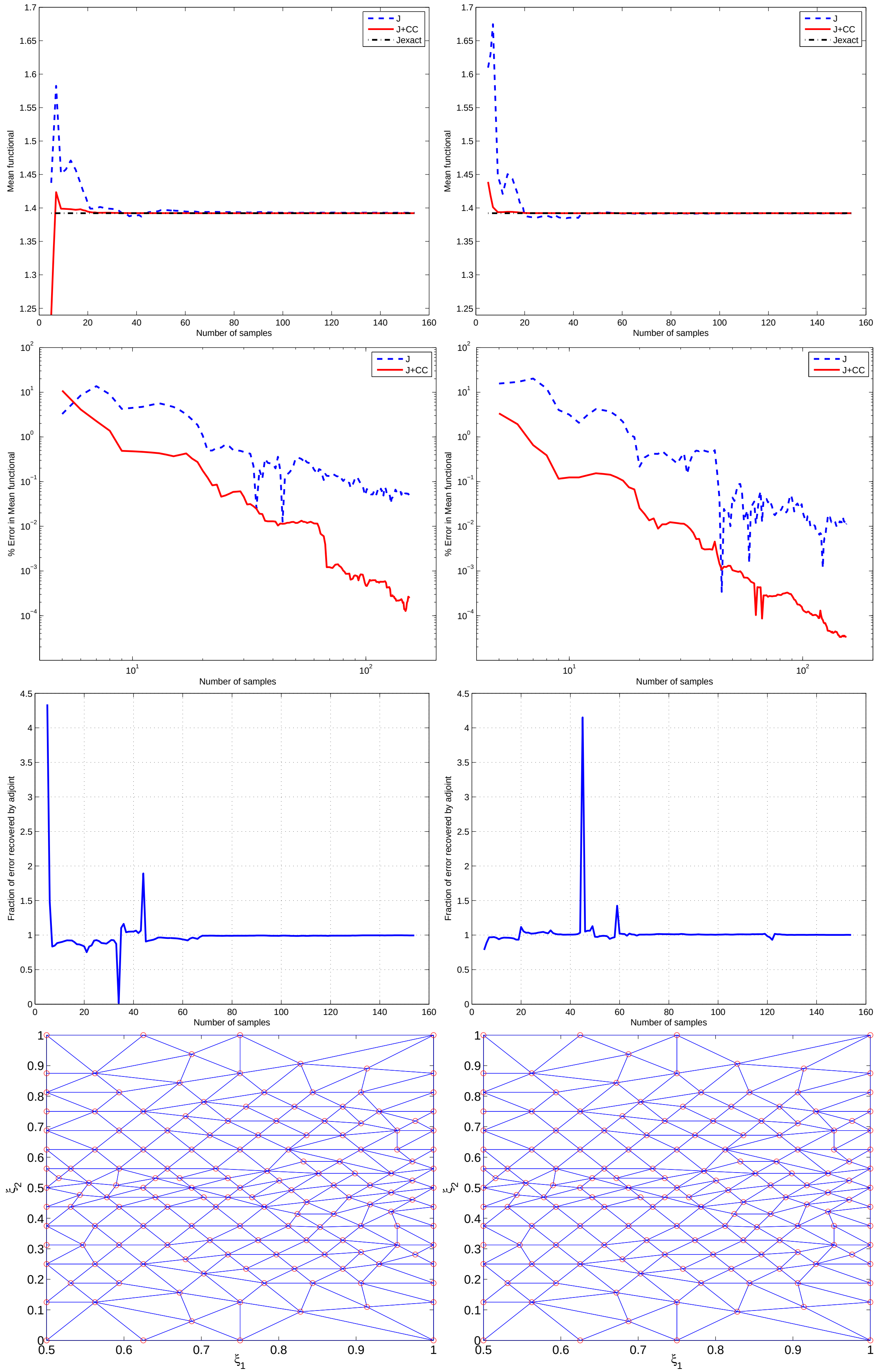

Figure 4. Adaptive sampling for algebraic example I: linear (left) and quadratic reconstruction (right) of state variables in stochastic space

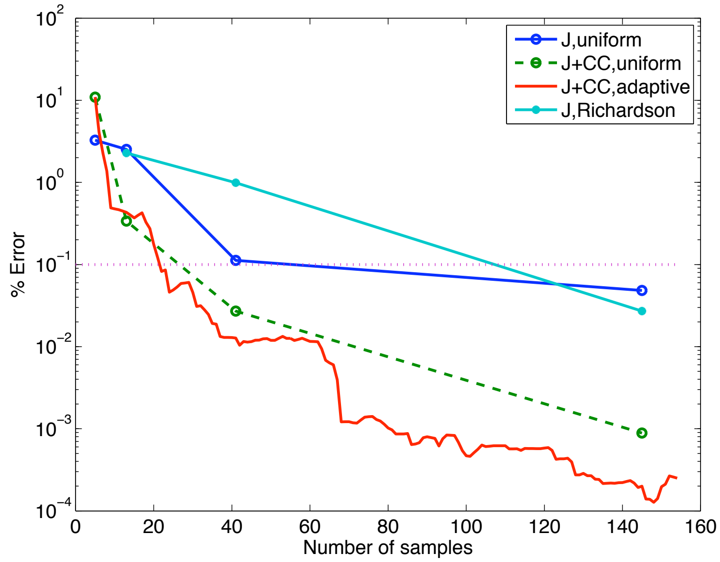

Figure 5. Comparison of error for algebraic example I using linear reconstruction

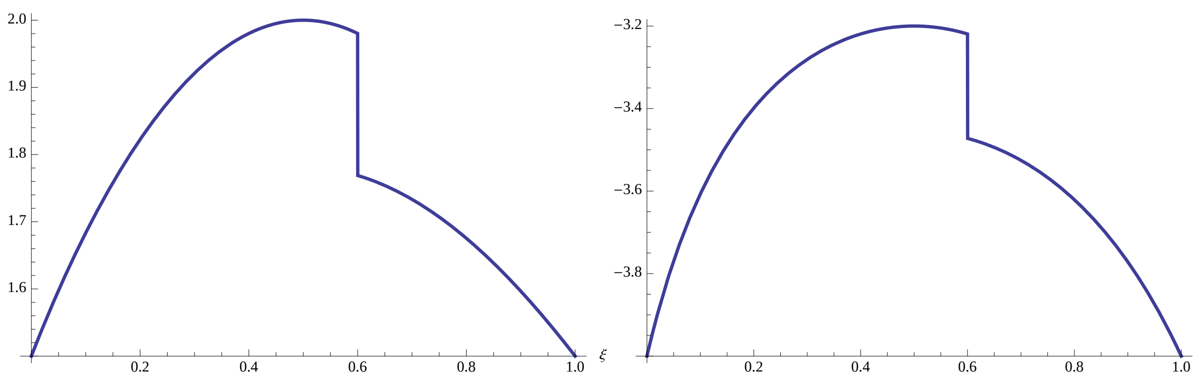

Figure 6. Primal and adjoint response surfaces for Algebraic problem II

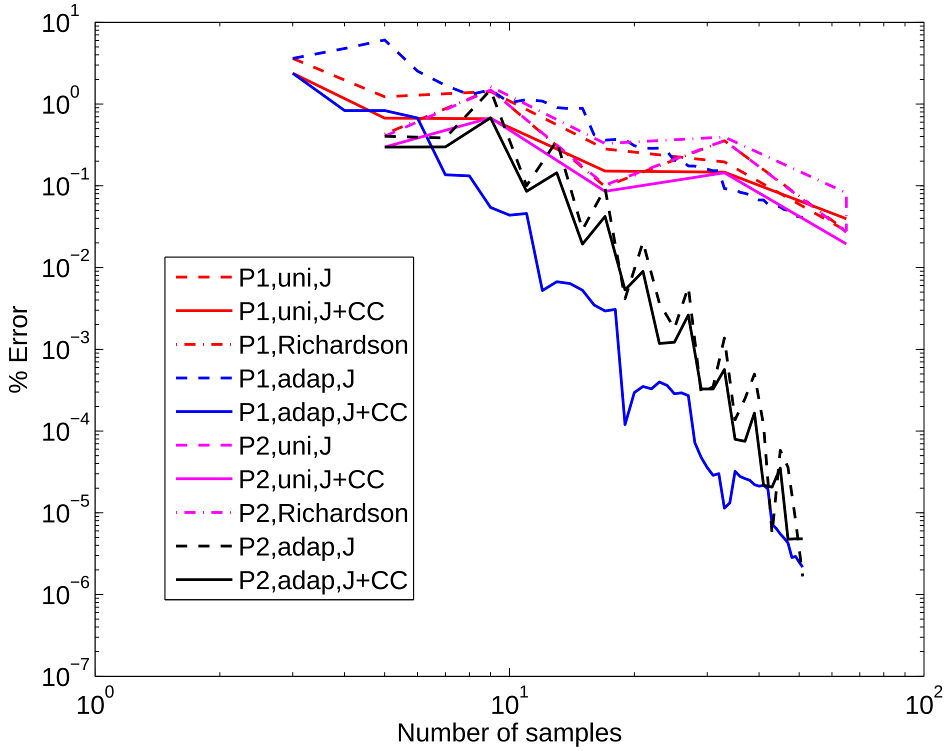

Figure 7. Primal and adjoint response surfaces for Algebraic problem II

As seen from Figures 9 and 10, theoretical convergence and super-convergence rates are observed for uniform refinement and the adaptive refinement is well behaved. Figure 11 shows that the adjoint-based error correction strategy is more accurate than Richardson-based error-estimation. However, the adaptive refinement is not appreciably more efficient than the uniform refinement.

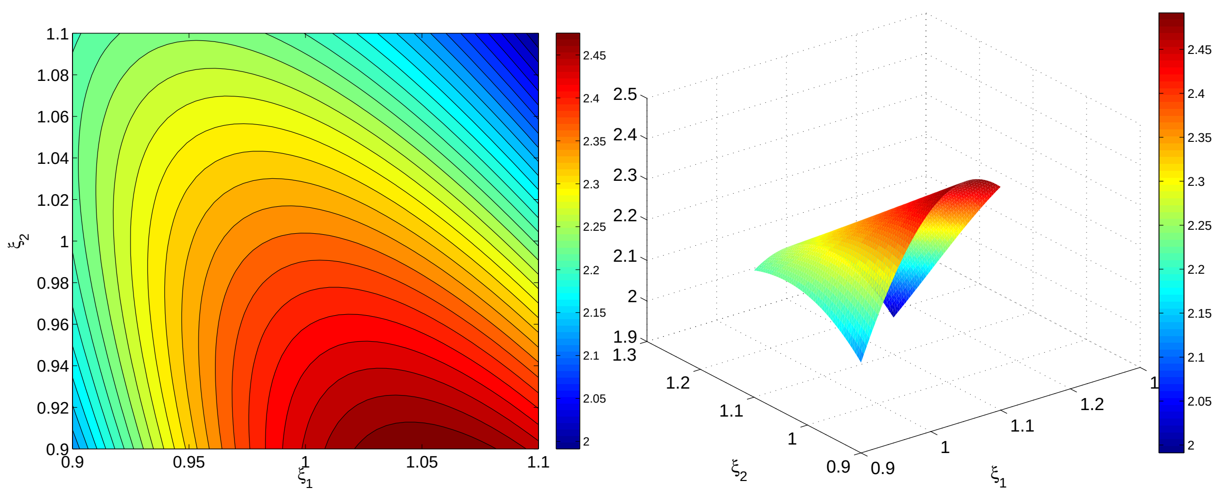

Figure 8. Random functional I for convection-diffusion problem

D. 2D Euler equations

As a final test case, inviscid transonic flow over a NACA 0012 airfoil is considered. The two dimensional Euler equations are solved on a spatial grid consisting of 9200 control volumes (figure 12). A second order accurate upwind spatial discretization is employed and discrete adjoints are computed using the automatic differentiation[21] tool-kit ADOL-C[22]. Two different test problems are considered, the first involving a situation where the freestream mach number varies between 0.8 and 1.0 and a two-parameter problem where, in addition to the afore-mentioned Mach

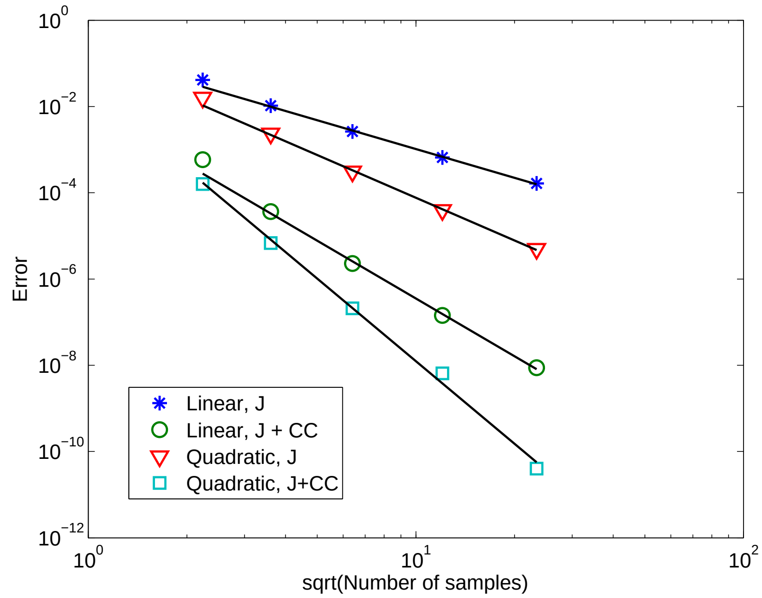

Figure 9. Convergence rates for convection-diffusion problem under uniform refinement: ($*$) 2.2, ($\circ$) 4.4, ($\triangledown$) 3.3, ($\square$) 6.3

number variation, the angle of attack also varies unifomly between $0°$ and $2°$. The objective functional is taken to be the mean drag coefficient over the parameter range.

Figure 13 shows the initial and final "response surfaces" for the linear and quadratic reconstructions. The adaptive refinement is seen to identify the peak in drag coefficient and as observed in figure 14 show rapid functional convergence compared to the uniform refinement strategy. In all these cases, the target error was set to be 1×10^{-6}, and hence the adaptive simulations are terminated when the computable correction reaches this level. The effectiveness of the computatable correction is also evident in the two parameter case (figures 15,16), as is the fact that the quadratic reconstruction shows drastic improvement over the linear reconstruction.

VII. Comments, generalizations and extensions

In the numerical examples, we have demonstrated the performance of the proposed UQ method for two random variables which are assumed to be uniformly distributed. Here we indicate further generalizations of the adjoint-based UQ approach.

1. The results presented in the paper make use of uniformly distributed random variables. However, it is easy to extend the approach to general random variables. If $P(\xi)$ is the probability density function, then it is enough to replace the measure $\mathrm{d}\xi$ with the probability measure $P(\xi)\mathrm{d}\xi$; for the discrete equations, this amounts to replacing the quadrature weights w_{ij} with $w_{ij}P(\xi_{ij})$ in all the formulae. The error estimate in each simplex is not only weighted by the interpolation error but also the probability measure for the simplex.

2. In principle, it is possible to use other types of stochastic elements for numerical integration. But the use of simplex elements has some advantages in terms of requiring smaller number of samples[14]. For example, for a P_2 simplex, all the samples are located at the vertices or on the boundary of the simplex, and hence they are shared by the neighbouring simplices. When a simplex is refined as a consequence of adaptation, the number of new samples required are also small, as compared to d-cubes.

3. The extension to more than two random variables using multi-dimensional simplex elements is straight-forward in principle, but of course the computational cost increases due to the curse of dimensionality. Due to this reason, the present approach may be limited to a small number of random variables. It is primarily aimed at those situations where simulation tools are computationally expensive and are being used to design some critical system for which reliable uncertainty quantification is of paramount importance. The adaptive sampling procedure then allows us to choose the samples in such a way that the error in the moments of interest are reduced in an optimal way. The quadratic reconstruction procedure from[12] can be extended to multi-dimensional case. Alternatively, we can also use P_2 simplex elements as in[14].

4. The convergence rates and superconvergence hold under sufficient smoothness of $I(u)$ and $R(u)$, and smooth dependance of the primal and adjoint solutions on the random variable ξ. As seen in Burgers problem, the primal

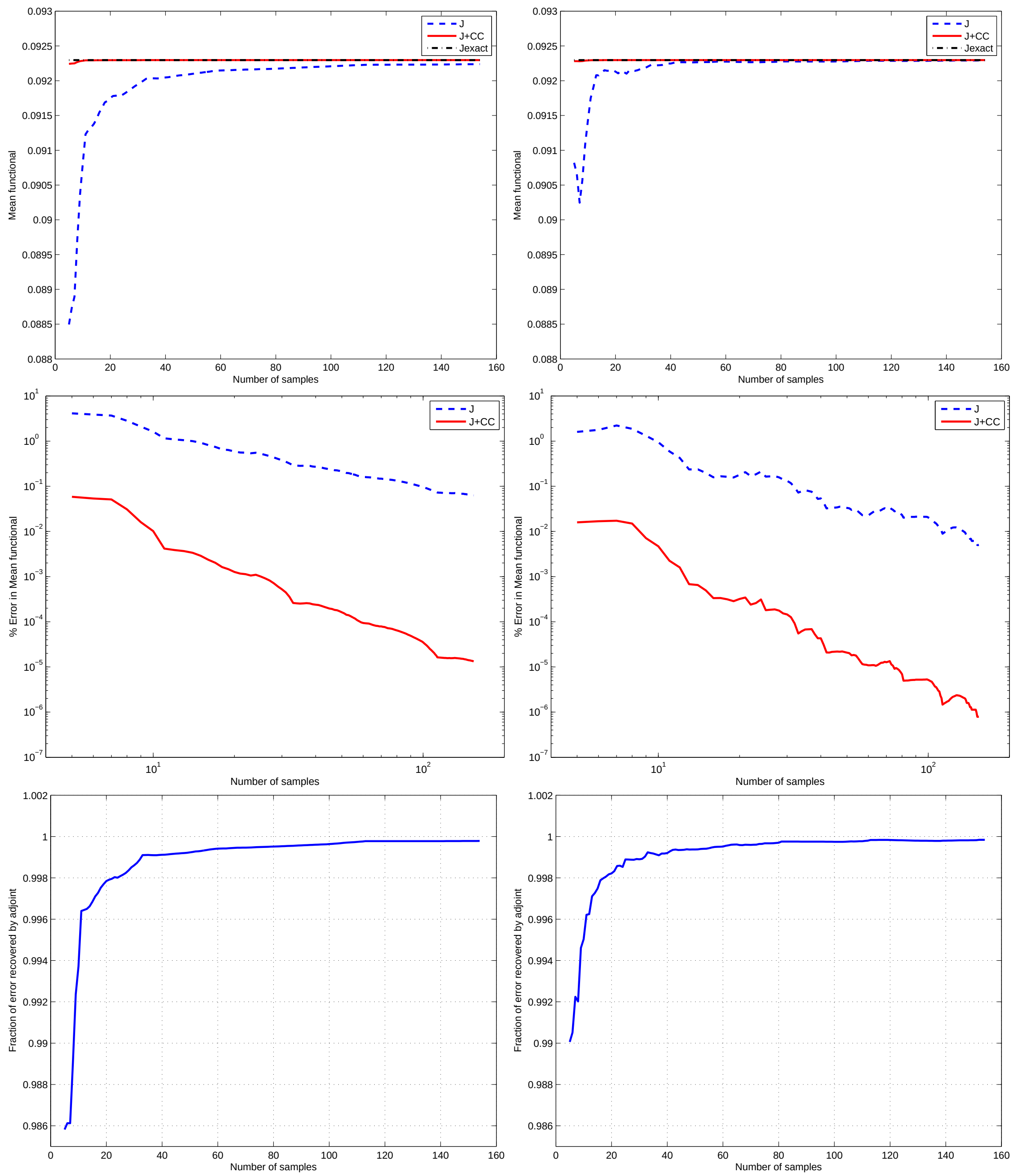

Figure 10. Adaptive sampling for convection-diffusion problem: linear (left) and quadratic reconstruction (right) of state variables in stochastic space

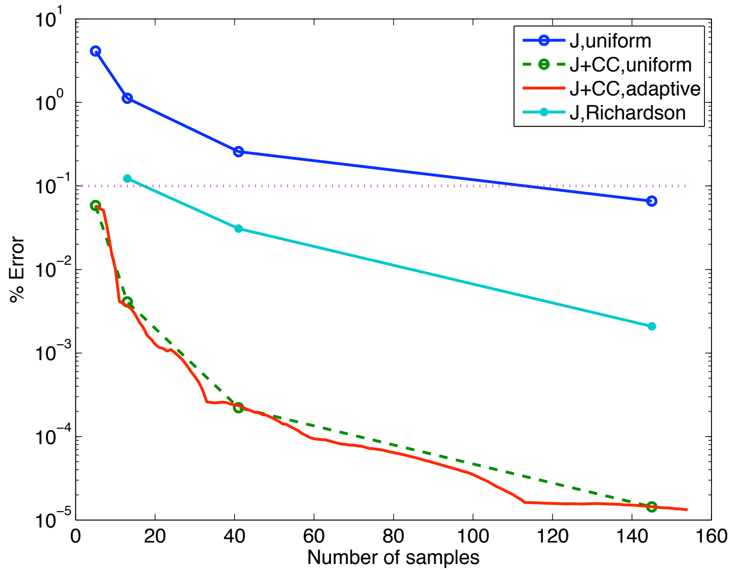

Figure 11. Comparison of error for convection-diffusion example using linear reconstruction

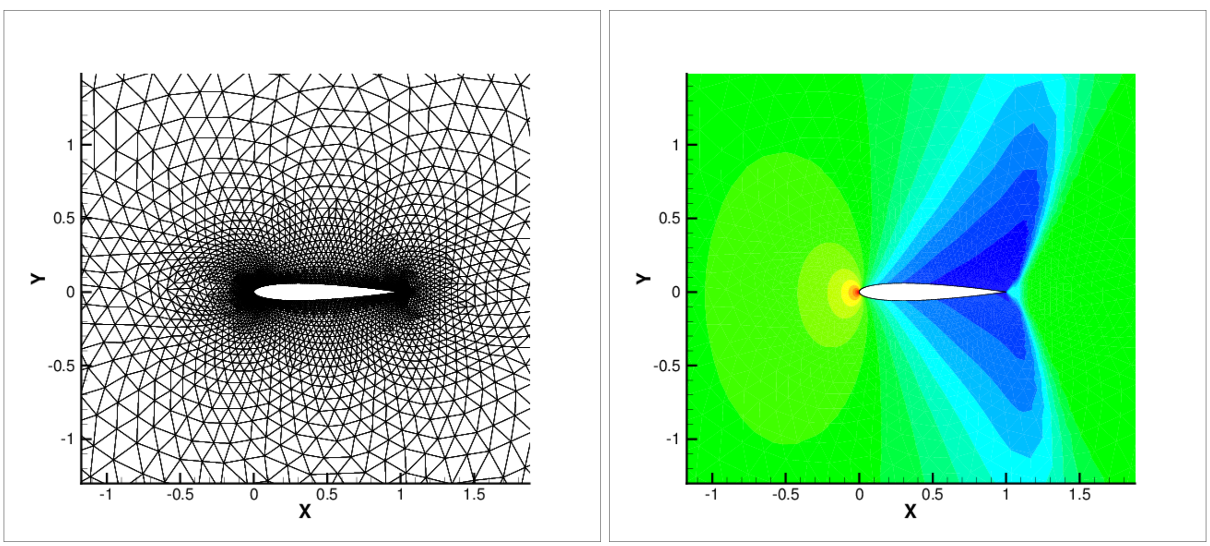

Figure 12. Spatial mesh (left) and sample flow (density) solution (right) for 2D Euler problem

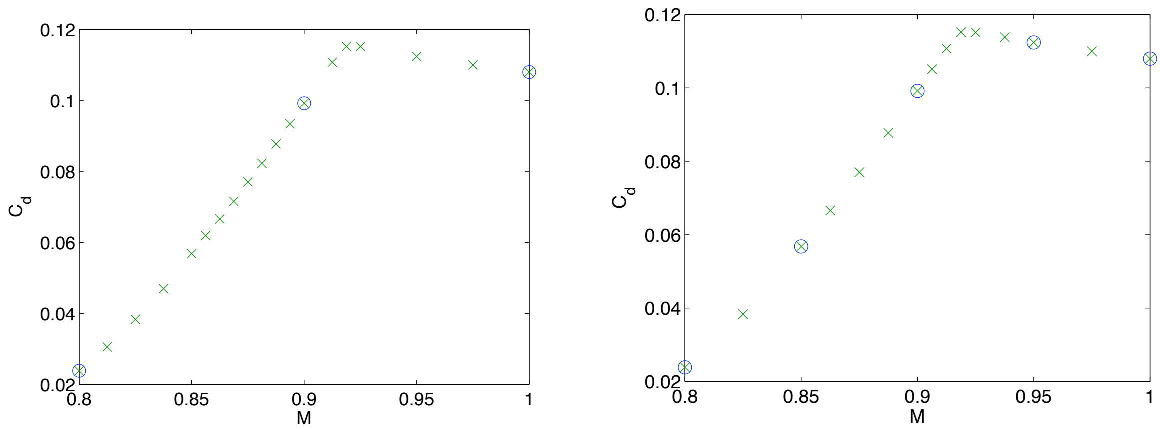

Figure 13. Initial (circles) and final (crosses) samples for one parameter 2D Euler problem. Left: P1, Right: P2 reconstruction

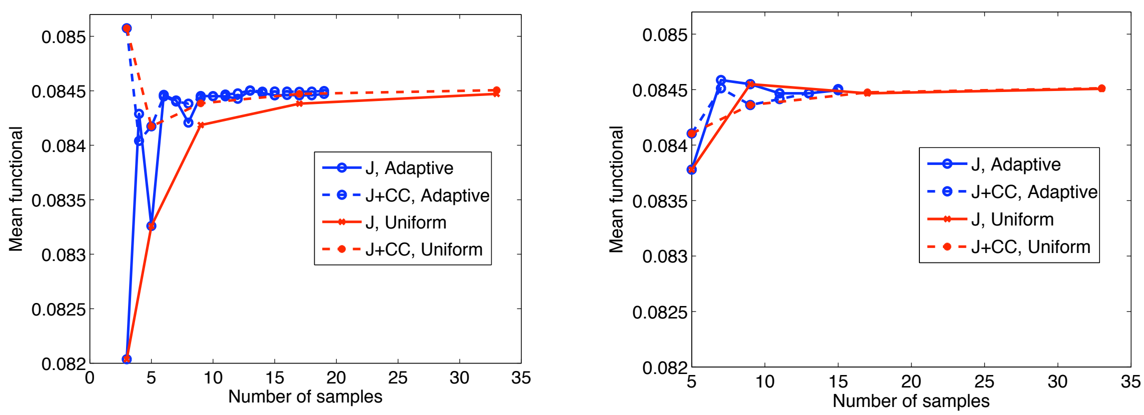

Figure 14. Convergence of functional for one parameter 2D Euler problem. Left: P1, Right: P2 reconstruction

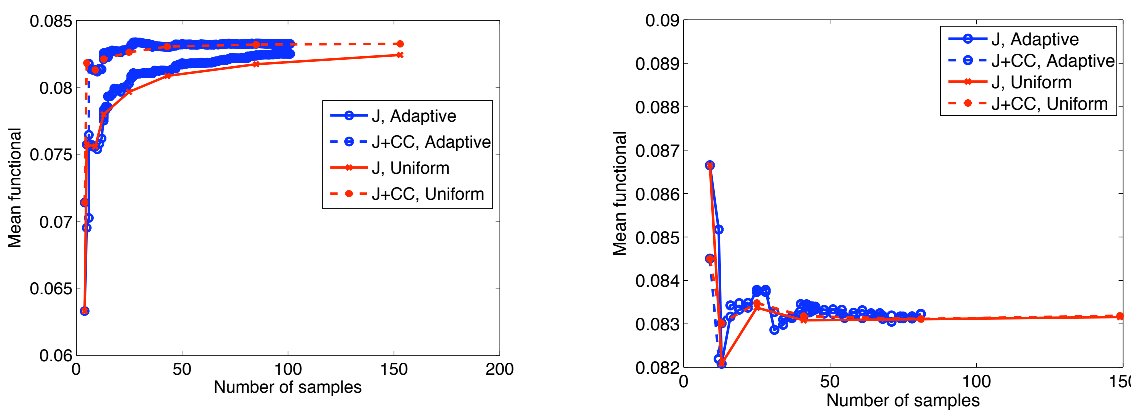

Figure 15. Convergence of functional for two parameter 2D Euler problem. Left: P1, Right: P2 reconstruction

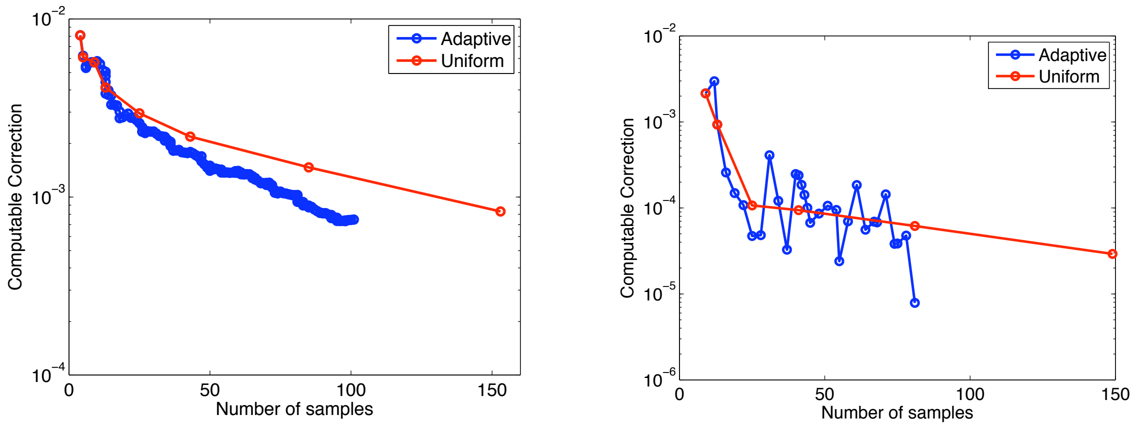

Figure 16. Convergence of error estimate for two parameter 2D Euler problem. Left: P1, Right: P2 reconstruction

solution has a discontinuous dependance on the random parameters since the shock location depends on these parameters. In such situations, quadratic reconstruction becomes oscillatory leading to loss of accuracy of the interpolation in stochastic space. This situation can be improved using a monotonicity or TVD reconstruction procedure in which the reconstruction order is reduced to linear on any stochastic simplex element in which a discontinuity is detected [14]. The reconstruction order can thus vary from one element to another within the same simplex grid, similar to p-refinement in the finite element method.

VIII. Conclusion

A new framework is introduced to estimate and control the numerical error in aleatoric uncertainty quantification of functional outputs using adjoint variables in stochastic space. Super-convergent error estimates and successful adaptive mesh refinement is demonstrated for a number of problems involving algebraic equations and ordinary and partial differential equations. The corrections to statistical moments based on adjoint solutions is seen to be able to recover a large percentage of error, leading to more accurate estimation of the moments. This strategy for error estimation and control coupled with the unstructured and non-intrusive sampling is hence a potentially effective way of handling highly complex response surfaces in multiple dimensions. The current approach also provides a platform to set error tolerances in computing numerical solutions at the sampling points themselves and to effectively budget the various sources of errors associated with uncertainty quantification.

Acknowledgment

This work was partially supported by the Predictive Science Academic Alliance Program (PSAAP) of the Department of Energy at Stanford University. The authors thank G. Iaccarino for early discussions on this work.

References

[1] R. Ghanem and S. F. Wojtkiewicz, *Special Issue on Uncertainty Quantification*, SIAM J. Sci. Comput. 26(2): vii-vii, 2004.

[2] H. N. Najm, *Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics*, Annu. Rev. Fluid Mech. 41:35-52, 2009.

[3] R. Ghanem and P. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Dover Publications, 2003.

[4] B. Debusschere, H. Najm, P. Pebay, O. Knio, R. Ghanem, and O. Le Matre, *Numerical challenges in the use of polynomial chaos representations for stochastic processes*, SIAM J. Sci. Comput. 26:698-719, 2004.

[5] O. Le Matre, H. Najm, P. Pebay, R. Ghanem, and O. Knio, *Multi-resolution-analysis scheme for uncertainty quantification in chemical systems*, SIAM J. Sci. Comput. 29:864-89, 2007.

[6] X. Wan, and G. Karniadakis, *Multi-element generalized polynomial chaos for arbitrary probability measures*, SIAM J. Sci. Comput. 28:901-28, 2006.

[7] R. Ghanem, *Hybrid stochastic finite elements and generalized Monte Carlo simulation*, ASME J. Appl. Mech. 65:1004-9, 1998.

[8] D. Ghiocel and R. Ghanem, *Stochastic finite-element analysis of seismic soil-structure interaction*, J. Eng. Mech. 128:66-77, 2002.

[9] S. Hosder, R. Walters, and R. Perez, *A nonintrusive polynomial chaos method for uncertainty propagation in CFD simulations*, AIAA Paper 2006-0891, 2006.

[10] J. A. S. Witteveen, *Efficient and robust uncertainty quantification methods for flow and fluid-structure interaction simulations*, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 2009.

[11] N. Pierce and M. Giles, *Adjoint and Defect Error Bounding and Correction for Functional Estimates*, J. Comp. Physics, 200:769-794, 2004.

[12] D. A. Venditti, and D. L. Darmofal, *Grid adaptation for functional outputs: Application to 2-D inviscid, compressible flow*, J. Comp. Physics, 176:40-69, 2002.

[13] A. Griewank, *Evaluating derivatives: Principles and techniques of algorithmic differentiaion*, SIAM, 1987.

[14] J. A. S. Witteveen, G. J. A. Loeven, and H. Bijl, *An adaptive stochastic finite elements approach based on Newton-Cotes quadrature in simplex elements*, Computers and Fluids, 38:1270-1288, 2009.

[15] M. Giles and N. Pierce, *An introduction to the adjoint approach to design*, Flow, Turbulence and Combustion, 65(3-4):393-415, 2000.

[16] D. J. Mavriplis, *Revisiting the Least-Squares Procedure for Gradient Reconstruction on Unstructured Meshes*, AIAA paper 2003-3986, 2004.

[17] E. Kreyszig, *Introductory Functional Analysis with Applications*, Wiley, 1989.

[18] Ch. Hirsch, *Numerical Computation of Internal and External Flows, vol 2*, Wiley, 1990.

[19] D. Dunavant, *High Degree Efficient Symmetrical Gaussian Quadrature Rules for the Triangle*, Int. J. Num. Methods Eng., 21:1129-1148, 1985.

[20] J. Burkardt, *DUNAVANT - Quadrature Rules for the Triangle*, MATLAB implementation available at http://people.sc.fsu.edu/~jburkardt/m_src/dunavant/dunavant.html

[21] A. Griewank, *Evaluating Derivatives*, Principles and Techniques of Algorithmic Diferentiation, Society for Industrial and Applied Mathematics, 2000.

[22] D. J. A. Griewank, J. Utke, *Algorithm 755: ADOL-C: A package for the automatic differentiation of algorithms written in c/c++*, ACM Transactions on Mathematical Software 22, 131–167, 1996.