

## Distributional Sensitivity for Uncertainty Quantification

Akil Narayan\* and Dongbin Xiu

*Department of Mathematics, Purdue University, 150 N. University Street,  
West Lafayette, IN 47904, USA.*

Received 16 February 2010; Accepted (in revised version) 30 July 2010

Communicated by George Karniadakis

Available online 30 March 2011

---

**Abstract.** In this work we consider a general notion of *distributional sensitivity*, which measures the variation in solutions of a given physical/mathematical system with respect to the variation of probability distribution of the inputs. This is distinctively different from the classical sensitivity analysis, which studies the changes of solutions with respect to the values of the inputs. The general idea is measurement of sensitivity of outputs with respect to probability distributions, which is a well-studied concept in related disciplines. We adapt these ideas to present a quantitative framework in the context of uncertainty quantification for measuring such a kind of sensitivity and a set of efficient algorithms to approximate the distributional sensitivity numerically. A remarkable feature of the algorithms is that they do not incur additional computational effort in addition to a one-time stochastic solver. Therefore, an accurate stochastic computation with respect to a prior input distribution is needed only once, and the ensuing distributional sensitivity computation for different input distributions is a post-processing step. We prove that an accurate numerical model leads to accurate calculations of this sensitivity, which applies not just to slowly-converging Monte-Carlo estimates, but also to exponentially convergent spectral approximations. We provide computational examples to demonstrate the ease of applicability and verify the convergence claims.

**AMS subject classifications:** 52B10, 65D18, 68U05, 68U07

**Key words:** Uncertainty quantification, epistemic uncertainty, distributional sensitivity, generalized polynomial chaos.

---

## 1 Introduction

Uncertainty quantification (UQ) has become an important tool for modelling in recent years. Many physical systems have uncertainties caused by unknown parameters in the

---

\*Corresponding author. *Email addresses:* acnaraya@purdue.edu (A. Narayan), dxiu@purdue.edu (D. Xiu)

model or by measurement noise plaguing experiments. In such cases, it is critical to understand and predict how the uncertainty affects quantities of interest (QoI) of the systems. This introduces a new paradigm for scientific computing and extends the traditional deterministic simulations to stochastic simulations.

One of the major challenges for stochastic computation and UQ is the simulation cost, as the dimensionality of simulations depends on the total number of random variables that one employs to parameterize the inputs. The larger the dimensionality the higher the cost, (the curse of dimensionality). To circumvent the difficulty, it is crucial to conduct *sensitivity analysis* (SA) prior to simulations. The goal of the sensitivity analysis is to determine which input variables have notable effects on the QoI and eliminate those with negligible effects on the simulation.

In this work we discuss a different kind of sensitivity analysis-distributional sensitivity analysis (DSA), which is intended to quantify the impact on the QoI with respect to changes in the probability distribution of the inputs. This is motivated by the fact that in many cases there is not sufficient data or evidence to fully specify the probability distribution of the inputs. Such kind of uncertainty is often referred to as *epistemic uncertainty*, as opposed to *aleatory uncertainty* where probabilistic information about the inputs is fully specified. For many practical systems, uncertain inputs often present themselves in the form of epistemic uncertainty, and acquiring more information to specify their probability can be a (highly) costly, and sometimes impossible, task. One of the immediate goals of DSA is to provide a guideline to direct the modeling effort. For inputs with large distributional sensitivity (DS), more effort will be required to acquire their probabilistic information; for inputs with small and negligible DS, it is acceptable to specify their distribution with something of computational convenience. By doing so, we can reduce the total number of epistemic variables to a minimum. It is worth remarking on the difference between the DS and the traditional sensitivity. While an input with small sensitivity in the traditional sense naturally implies small DS, there is no direct association on the other hand, i.e., an input with large sensitivity in the traditional sense does not necessarily imply large DS, and vice versa. Therefore, while the traditional SA is a necessary step to reduce the computational burden for (aleatory) stochastic simulations, the DSA is a necessary step to reduce the simulation effort for dealing with epistemic uncertainty.

Indeed, the underlying concern of DSA is the study of how assumptions about probability densities affects outputs. Unsurprisingly, this notion already exists in related fields; one manifestation of this is the "score function" approach [1, 13]. The score function method assumes a parameterization of a family of input distributions and primarily uses a Monte-Carlo estimate to compute sensitivities. The "what-if" problem, extrapolation of the QoI values to unsimulated density locations, is not a consideration of this paper. The spirit of DSA is also captured in the study of local/global sensitivity analysis from Bayesian statistics; this analysis studies the effect that the assumed prior has on the resulting posterior [4, 9, 14]. Our problem is not in the context of Bayesian statistics; in particular we are not concerned with updating our assumed "prior".

Since computational effort is of great concern in stochastic computations, it is desir-

able to keep the simulation efforts for DSA to a minimum. In this paper, in addition to proposing a definition of DS, we also present a set of numerical algorithms that require only a forward stochastic simulation, which is needed for any traditional UQ computations, and reduce the evaluation of DS to a post-processing step. The forward stochastic simulation will be based on two of the most widely used methods, generalized Polynomial Chaos (gPC) based methods, and sampling-based methods.

This paper is organized as follows. Section 2 introduces notation and some requisite definitions for distributional sensitivity. Section 3.1 discusses convergence theory for a Galerkin-like discretization and Section 3.2 considers a sampling-based approximation. Section 4 discusses the implementation aspects of computing distributional sensitivity. Finally, we conclude with a few numerical examples in Section 5 and closing remarks in Section 6.

## 2 Distributional sensitivity

### 2.1 Notations

Let  $\mathcal{Z} \in \mathbb{R}^D$  be a random variable on probability space  $(\Omega, \mathcal{F}, \nu)$ , i.e.,

$$P[\mathcal{Z} \in A] = \int_A d\nu.$$

We assume in this paper that any measure of interest  $\nu$  is absolutely continuous with respect to Lebesgue measure ( $\lambda$ ) so that a density  $\rho$  exists:

$$P[\mathcal{Z} \in A] = \int_A \rho d\lambda.$$

The density  $\rho$  of a measure  $\nu$  will be denoted by its Lebesgue-Radon-Nikodym derivative  $d\nu/d\lambda$ . We use  $E_\nu$  to mean the expectation of a quantity with respect to the measure  $\nu$  and  $L^p(\nu)$  to mean the space of  $L^p$  integrable functions under the measure  $\nu$ ; the norm on this space is denoted  $\|\cdot\|_{L^p(\nu)}$ . When  $\nu = \lambda$ , we use shorthand notation  $\|\cdot\|_p$  for the norm. On an inner product space, we denote  $\langle \cdot, \cdot \rangle_\rho$  to be the inner product with respect to Lebesgue measure weighted by the density  $\rho$ . If a measure  $\eta$  is absolutely continuous with respect to a measure  $\nu$  we write  $\eta \ll \nu$ . If the norm on  $L^p(\eta)$  dominates the norm on  $L^p(\nu)$ , i.e., if there exists a constant  $K$  such that for all  $u \in L^p(\eta)$  we have  $\|u\|_{L^p(\nu)} \leq K\|u\|_{L^p(\eta)}$ , then we write  $\|\cdot\|_{L^p(\nu)} \leq \|\cdot\|_{L^p(\eta)}$ .

Since we assume that all probability measures are dominated by  $\lambda$ , a density function exists that is semi-positive on the interior of the domain; therefore the measure  $\nu$  admits a collection of  $L^2(\nu)$ -orthogonal polynomials. For example, if  $\nu \sim \mathcal{N}(\mu, \sigma^2)$ , the normal distribution, then  $d\nu/d\lambda$  is proportional to  $e^{-z^2/2}$  on  $\mathbb{R}$  and the Hermite polynomials form such an orthogonal family. For  $\mathcal{Z} = (z_1, z_2, \dots, z_D)$  define the space of  $D$ -variate

polynomials of degree  $N$  or less as

$$\Pi_N = \text{span} \left\{ z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_d^{\alpha_d} : \alpha_j \geq 0, \sum_{j=1}^d \alpha_j \leq N \right\}.$$

Let  $\mathcal{P}_N^{(\nu)}$  be the  $L^2(\nu)$ -projection operator defined by

$$\langle u - \mathcal{P}_N^{(\nu)} u, v \rangle_\rho = 0, \quad v \in \Pi_N,$$

where  $u \in L^2(\nu)$ .

## 2.2 Problem setup

Assume a physical system is modeled by a differential operator

$$\mathcal{L}(u; x; \mathcal{Z}) = 0, \quad (2.1)$$

where  $u$  is the unknown quantity,  $x$  is the independent spatial/temporal variable (it is not necessary for  $\mathcal{L}$  to depend on time, and we do not distinguish between time and space in our notation). We assume that  $\mathcal{L}$  is well-posed in both the  $x$  and  $\mathcal{Z}$  variables. Therefore, for each realization  $z$  of  $\mathcal{Z}$  we can solve the problem (2.1) and obtain the solution  $u(x; z)$ . We denote the QoI  $\xi$ , and it is usually obtained from  $u$  via integration over  $\mathcal{Z}$ . For example,  $\xi$  can be the mean or variance of  $u$ . For some generality assume

$$\xi_\nu(u) = \int f(u(z)) d\nu(z) = E_\nu f(u), \quad (2.2)$$

where  $f$  itself may integrate over  $t$  and/or  $x$ . In the specific cases where  $\xi$  is the mean or variance, we make use of the notation  $\mu$  and  $\sigma^2$ , respectively:

$$\mu_\nu(u) = E_\nu u, \quad \sigma_\nu^2(u) = E_\nu [u - E_\nu u]^2.$$

Many quantities of interest are computable once an approximation for the solution  $u$  has been formed. However, there are situations when one is interested in estimating how quantities of interest change with respect to the underlying distribution of the random parameter  $\mathcal{Z}$ . In addition, sometimes the distribution of  $\mathcal{Z}$  may not *a priori* be known with certainty. Such an example is the class of Bayesian inference problems when the density function is continually updated based on newly computed data [10]. Having already spent considerable computational expense in determining an approximation to  $u$ , one is interested in quantifying the deviation in the approximation caused by incremental changes to the density function. In addition, the input distribution to a numerical model affects the computed quantities of interest; we propose below a quantitative method for evaluating sensitivity of the model with respect to the underlying distribution.

### 2.3 Distributional sensitivity

In order to quantify sensitivities with respect to distributional changes, we require a metric on probability distributions. Such metrics are plentiful, we denote them as follows:

**Definition 2.1.** For measures  $\nu$  and  $\eta$ , we let  $d(\nu, \eta)$  denote any (possibly nonsymmetric) distance over the measures.

Our allowance of nonsymmetry means that  $d$  is not necessarily a proper metric. Since we assume that both  $\nu$  and  $\eta$  are dominated by  $\lambda$ , we can use measures of distance between the densities as well. Some examples for  $d(\cdot, \cdot)$  are

$$\begin{aligned} d(\nu, \eta) &= \left\| \frac{d\eta}{d\lambda} - \frac{d\nu}{d\lambda} \right\|_1 && (L^1(\lambda) \text{ distance between densities}), \\ d(\nu, \eta) &= \frac{1}{\sqrt{2}} \left\| \sqrt{\frac{d\eta}{d\lambda}} - \sqrt{\frac{d\nu}{d\lambda}} \right\|_2 && (\text{Hellinger distance}), \\ d(\nu, \eta) &= E_\eta \log \frac{d\eta}{d\nu} && (\text{Unsymmetrized Kullback-Leibler divergence}). \end{aligned}$$

In the numerical results of this paper, we only use the  $L^1$  metric; our experiments show that the choice of metric is relatively unimportant. Various relations between the different metrics can also be used to determine which is the appropriate choice for a given problem [7, 9]. We are now in a position to define the distributional sensitivity:

**Definition 2.2.** Let  $u$  be the solution to Eq. (2.1). Consider  $\zeta \cdot (u) : U \rightarrow \mathbb{R}^N$ , where

$$U \doteq \{ \nu : \nu \text{ is a probability measure and } \nu \ll \lambda \}.$$

Then the continuous Distributional Sensitivity of  $\zeta$  with respect to the measure  $\nu$  in the direction  $\Delta\nu \ll \lambda$  is the modulus of the Gateaux derivative of  $\zeta$  at  $\nu$  along  $\Delta\nu$ :

$$\widetilde{DS}_\zeta[\nu; \Delta\nu] = \left\| \lim_{\epsilon \rightarrow 0} \frac{\zeta_{\nu+\epsilon\Delta\nu}(u) - \zeta_\nu(u)}{\epsilon} \right\|, \quad (2.3)$$

where  $\|\cdot\|$  is the Euclidean norm.

See, for example, the same definitions in Section 3 of [9]. We elect to define the sensitivity by not taking the derivative inside the integral in (2.2) as is done in the score function method; in order to pass the derivative under the integral we require some additional assumptions for mathematic rigor, and we could not keep the simplistic form dependent explicitly on  $\zeta$ .

With the unrestrictive definition of  $\zeta$  in (2.2), it is possible for the distributional sensitivity to not exist in some direction(s)  $\Delta\nu$ . Indeed, it is not necessarily the case that  $\zeta$  satisfies the stronger notion of Frechét differentiability at  $\nu$ ; such considerations are common in Bayesian sensitivity studies [3, 5]. The examples shown in Section 5 suggest that  $\zeta$  is frequently not Frechét differentiable. The above definition serves as inspiration for a more easily computable, discrete version of the distributional sensitivity that we define below:

**Definition 2.3.** The discrete Distributional Sensitivity of the quantity of interest  $\xi$  between measures  $\nu$  and  $\eta$  is defined to be

$$DS_{\xi}[v,\eta](u) = \frac{\|\xi_{\nu}(u) - \xi_{\eta}(u)\|}{d(v,\eta)}. \tag{2.4}$$

This finite-difference notion of sensitivity appears in many other contexts including estimates of stochastic derivatives [2] and in obtaining Lipschitz stability results for solutions to optimization problems [11, 12].

We emphasize that we only require the two measures  $\eta$  and  $\nu$  to be members of the same Banach space so that the continuous definition (2.3) makes sense (and therefore the discrete definition (2.4) is meaningful).  $\eta$  and  $\nu$  need not share a common parameterization. In particular, if  $\nu$  is some "nominal" measure, the collection of perturbations  $\{\eta\}$  over which we measure the DS could be characterized by complicated topology.

The remainder of this paper deals with the distributional sensitivity defined in Definition 2.3. Hereafter, the term "distributional sensitivity" refers to the discrete distributional sensitivity. The distributional sensitivity is a function of  $\nu$  and  $\eta$  and the definition is not dependent on the particular numerical approximation of  $u$ . It is a measure of how a quantity of interest changes depending on the assumed distribution of the random parameter(s).

**Example 2.1.** Consider a system whose solution  $u$  is given by the model

$$u(z_1, z_2) = \exp(2z_1) \cos(3\pi z_2). \tag{2.5}$$

The unknown system parameters are  $z_1$  and  $z_2$  and we are interested in the mean and variance of the output. Suppose the nominal state of the system dictates that  $z_1$  and  $z_2$  are independent uniform random variables over  $[-1, 1]$ ; denote this measure  $\nu$ . Let us devise a prioritization of the two parameters, measured by the effect that each has on  $\xi$  (the mean or variance). To test the sensitivities of the system, now let  $z_i \sim \text{Beta}(\alpha_i, \alpha_i)$  for  $i = 1, 2$ . Let us call this new measure  $\eta$ . Using the  $L^1(\lambda)$  distance metric, the following are readily computable:

		$DS_{\mu}[v,\eta](u)$	$DS_{\sigma^2}[v,\eta](u)$
$\alpha_1 = \frac{1}{2},$	$\alpha_2 = 0$	0	4.2080
$\alpha_1 = -\frac{1}{3},$	$\alpha_2 = 0$	0	5.0884
$\alpha_1 = 0,$	$\alpha_2 = \frac{1}{2}$	0.2946	0.2198
$\alpha_1 = 0,$	$\alpha_2 = -\frac{1}{3}$	0.6787	0.7008

The results clearly show that if the mean is considered, then the parameter  $z_2$  affects the system more than  $z_1$ ; the reverse is true if the variance is the quantity of interest. This conclusion is logically consistent when the solution  $u$  in the form (2.5) is considered. We note finally that our choice of perturbed measures as being parameterized by the scalars  $\alpha_1, \alpha_2$  is done for convenience of exposition; there is no such restriction in the definition of DS.

We now discuss some useful facts about DS. If we choose the probability distance function  $d(\eta, \nu)$  as the  $L^1$  distance between the densities and use the definition of  $\zeta$  provided in (2.2), then the distributional sensitivity satisfies

$$DS_{\zeta}[\eta, \nu](u) \leq E_{\rho}|f(u)|, \quad (2.6a)$$

$$\rho = \frac{\left| \frac{d\eta}{d\lambda} - \frac{d\nu}{d\lambda} \right|}{E_{\lambda} \left| \frac{d\eta}{d\lambda} - \frac{d\nu}{d\lambda} \right|}, \quad (2.6b)$$

and  $\rho$  is a probability density with respect to  $\lambda$ . Even if a different distance function is used, the bound (2.6a) is correct for any  $\zeta$  of the form (2.2) (with the expectation changed to an integral); however  $\rho$  may not be a probability density and may not satisfy (2.6b).

In some applications the particular distributional perturbations of interest may fall on a one-dimensional parameterization, e.g.,

$$\frac{d\nu}{d\lambda}(Z) = g_{\alpha}(Z)$$

for a scalar parameter  $\alpha$ . In this case the discrete distributional sensitivity can be replaced by a more accurate functional derivative. We obtain a new quantity:

$$\widehat{DS}_{\zeta}[v, \alpha](u) = \left| \int f(u) \frac{\partial g_{\alpha}}{\partial \alpha} d\lambda \right|.$$

Then the above quantity coincides with the continuous distributional sensitivity  $\widetilde{DS}$ . The above case is mathematically equivalent to the modulus of the score function formulation [13]; a unique (and important) aspect of the score function method is to write

$$\frac{\partial g_{\alpha}}{\partial \alpha} = g_{\alpha} \frac{\partial}{\partial \alpha} \log g_{\alpha},$$

identifying sensitivity as an expectation, which is not done in our formulation. Of course, in such a case when the parameterization is specifically available, the above quantity, implemented using the score function approach, is preferred.

Finally, we remark that the definitions of distributional sensitivity extend to cases when  $\nu$  is not absolutely continuous with respect to  $\lambda$ . For the continuous case, one must define the appropriate enclosing Banach space on which to apply the Gateaux derivative. For the discrete case, only a choice of probability distance  $d(\cdot, \cdot)$  is required. In particular, the case of a discrete state space is an easy extension.

### 3 Numerical approximations of distributional sensitivity

The definition of distributional sensitivity is computable if the exact solution  $u$  is available. However in practice we have an approximation  $u_N$  obtained by assuming  $\mathcal{Z} \sim \nu$ , a

"prior" density. In applications, we can compute the distributional sensitivity from the approximation  $u_N$ . We would like to know if our computed  $DS(u_N)$  is comparable to the true  $DS(u)$ . We will characterize the computed  $DS$  in terms of the true  $DS$  by making assumptions on the accuracy of  $u_N$ .

### 3.1 Solutions converging in averaged norms

Let the solution  $u_N^{(v)}$  be a numerical approximation of  $u$ , obtained by, for example, the stochastic Galerkin solution of the original system:

$$\left\langle \mathcal{L}(u_N^{(v)} - u; x; \mathcal{Z}), \phi \right\rangle_\rho = 0, \quad \phi \in \Pi_N,$$

where  $\rho = d\nu/d\lambda$ . Due to the strong  $L^2$ -theory behind Galerkin approximations, it is often possible to state convergence estimates between  $u_N$  and  $u$  under an  $L^p$ -norm weighted by the pdf  $\rho$ . If the norms induced by the two measures  $\eta$  and  $\nu$  are related, then we can prove convergence of the quantity of interest  $\xi$  in the cases when  $\xi$  is the mean and variance. Here we discuss two widely encountered  $L^p$ -norms, the  $L^1$ - and  $L^2$ -norms.

**Proposition 3.1.** Assume that  $\|\cdot\|_{L^1(\eta)} \leq \|\cdot\|_{L^1(\nu)}$ , with

$$\|u - u_N^{(v)}\|_{L^1(\nu)} < \varepsilon.$$

Then

$$|DS_\mu[\eta, \nu](u_N^{(v)}) - DS_\mu[\eta, \nu](u)| \leq \frac{2\varepsilon}{d(\eta, \nu)}.$$

*Proof.* Due to the relationship between the norms, the true solution  $u$  with respect to measure  $\nu$  is also the solution with respect to  $\eta$ . Using the triangle inequality with the definition of  $DS_\mu[\eta, \nu](u_N^{(v)})$ , we obtain

$$d(\eta, \nu) DS_\mu[\eta, \nu](u_N^{(v)}) \leq |E_\eta(u_N^{(v)} - u) - E_\nu(u_N^{(v)} - u)| + |E_\eta u - E_\nu u|,$$

which implies that

$$d(\eta, \nu) (DS_\mu[\eta, \nu](u_N^{(v)}) - DS_\mu[\eta, \nu](u)) \leq E_\eta |u - u_N^{(v)}| + E_\nu |u - u_N^{(v)}|.$$

The same bound can be obtained for  $DS(u) - DS(u_N^{(v)})$  by using the triangle inequality on the definition of  $DS(u)$ . Now the dominance of the  $\nu$  measure along with the assumption of accuracy for the Galerkin solution establishes the result.  $\square$

The assumption of the relationship between the norms here plays a role both in identifying a common solution  $u$  and in translating a bound in the  $\eta$  norm into one for the  $\nu$  norm. If the  $L^2$  error norm is bounded then the  $L^1$  assumption in Proposition 3.1 is satisfied. In this case of stronger convergence, we have convergence of the variance:

**Proposition 3.2.** Assume that  $\|\cdot\|_{L^2(\eta)} \leq \|\cdot\|_{L^2(\nu)}$  and  $u \in L^2(\nu)$ , with

$$\|u - u_N^{(\nu)}\|_{L^2(\nu)} < \varepsilon.$$

Then

$$|DS_{\sigma^2}[\eta, \nu](u_N^{(\nu)}) - DS_{\sigma^2}[\eta, \nu](u)| \leq \frac{K\varepsilon(\sigma_\nu(u_N^{(\nu)}) + \mu_\nu(u_N^{(\nu)})^2 + \varepsilon^2)}{d(\eta, \nu)}.$$

*Proof.* As with the previous proposition, the triangle inequality applied to the definition of  $DS(u_N^{(\nu)})$  yields

$$\begin{aligned} d(\eta, \nu)DS_{\sigma^2}[\eta, \nu](u_N^{(\nu)}) &\leq d(\eta, \nu)DS_{\sigma^2}[\eta, \nu](u) \\ &\quad + \underbrace{\left| \mathbb{E}_\eta[u_N^{(\nu)}]^2 - \mathbb{E}_\eta u^2 \right|}_{(A)} + \underbrace{\left| (\mathbb{E}_\eta u_N^{(\nu)})^2 - (\mathbb{E}_\eta u)^2 \right|}_{(B)} \\ &\quad + \underbrace{\left| \mathbb{E}_\nu[u_N^{(\nu)}]^2 - \mathbb{E}_\nu u^2 \right|}_{(C)} + \underbrace{\left| (\mathbb{E}_\nu u_N^{(\nu)})^2 - (\mathbb{E}_\nu u)^2 \right|}_{(D)}. \end{aligned}$$

By using Hölder's inequality and the assumption on the norms, term (A) satisfies

$$\left| \mathbb{E}_\eta[u_N^{(\nu)}]^2 - \mathbb{E}_\eta u^2 \right| \leq 2K\varepsilon \|u_N^{(\nu)}\|_{L^2(\nu)} + \varepsilon^2,$$

where  $K$  is the constant of relating the norms. Term (C) also satisfies this bound with  $K = 1$ . Similarly, one can bound term (B):

$$\left| \mathbb{E}_\eta[u_N^{(\nu)}]^2 - \mathbb{E}_\eta u^2 \right| \leq 2K\varepsilon \|u_N^{(\nu)}\|_{L^2(\nu)} + \varepsilon^2,$$

and  $K = 1$  for term (D). This establishes the result when

$$DS(u_N) \leq DS(u).$$

The reverse bound can be obtained by using the triangle inequality on the definition of  $DS(u)$ .  $\square$

The assumptions and results of the previous two propositions are not directly tied to how the approximate solution  $u_N$  is obtained. However, the assumptions on the relationship between the true solution  $u$  and the approximate solutions  $u_N$  are frequently proved for many Galerkin approximations of differential equations. The critical requirement that may become restrictive is that the  $\eta$ -norm is dominated by the  $\nu$ -norm. In an extreme case where the underlying measures do not even have the same support, bounds on the error may not be possible.

### 3.2 Solutions converging uniformly

In practice one may be able to obtain an approximate solution that is convergent uniformly in the random space. Such a stronger approximation (stronger than  $L^p$ -norm) is usually a result of high regularity of the true solution in the random space. In this case, convergence for the distributional sensitivity of the mean and variance are easier to prove:

**Proposition 3.3.** Assume that  $\|\cdot\|_{L^1(\eta)} \leq \|\cdot\|_{L^1(\nu)}$  and

$$\|u_N - u\|_{L^\infty} < \varepsilon.$$

Then

$$\begin{aligned} |DS_\mu[\eta, \nu](u_N) - DS_\mu[\eta, \nu](u)| &\leq \frac{2\varepsilon}{d(\eta, \nu)}, \\ |DS_{\sigma^2}[\eta, \nu](u_N) - DS_{\sigma^2}[\eta, \nu](u)| &\leq K\varepsilon \frac{|\mu_\nu(u_N)| + \varepsilon}{d(\eta, \nu)}. \end{aligned}$$

*Proof.* Both proofs follow the methodologies of the previous propositions. The proof for the mean  $\mu$  is straightforward. For the proof of the variance, term (A) from the proof of Proposition 3.2 can be bounded by

$$|E_\eta[u_N^{(\nu)}]^2 - E_\eta u^2| \leq |E_\eta(u_N - u)(u_N + u)| \leq \varepsilon E_\eta(u_N + u) \leq \varepsilon(2|\mu_\nu(u_N)| + \varepsilon).$$

The bounds for the remaining terms are obtained analogously. □

If (a)  $u_N$  is obtained as a Monte-Carlo estimate and (b) the measures are parameterized so that

$$\widetilde{DS}_\mu[v; \Delta v_\alpha] = \left| \int u \frac{\partial g_\alpha}{\partial \alpha} d\lambda \right|$$

for some parameter  $\alpha$ , then precise estimates of the bias are possible [19]; for example,

$$E(\widetilde{DS}_\mu(u) - DS_\mu(u_N))^2 \sim N^{-\frac{2}{3}}.$$

These specialized results do not rely on the assumptions of the proposition.

In Proposition 3.3 we still assume that the  $\nu$ -norm dominates the  $\eta$ -norm to relate the  $\eta$ -mean to the  $\nu$  mean. However, as long as we can guarantee pointwise convergence to the common solution, computing the mean and variance with accuracy is always possible.

Although the convergence results of Propositions 3.1-3.3 are proven under the assumption that a continuous density exists for each measure, this can be relaxed. In particular, the results remain true for discrete distributions (in which case one must replace, e.g.,  $L^1$  with  $l^1$ ).

## 4 Computability of distributional sensitivity

Here we discuss numerical approaches for computing distributional sensitivity based on two widely adopted stochastic solvers, the generalized Polynomial Chaos (gPC) methods and sampling methods. A key feature of the proposed algorithms is that the computation of DS requires only a forward stochastic simulation, by either gPC or sampling, and therefore renders the DS computation as a post-processing step. The gPC and Monte-Carlo approaches are *examples* of how DSA can be used in computation; the formulation of DS is not dependent on these two simulation methods.

### 4.1 Algorithms based on generalized polynomial chaos

The generalized Polynomial Chaos (gPC) approximation, made popular by the work of [6, 16], is a function expansion over stochastic space to approximate the dependence of an unknown on the random parameters. Typically, the function expansion takes the form of a polynomial expansion. That is, we seek an approximation  $u_N$ , in a properly defined polynomial space  $\Pi$ , such that  $\|u_N - u\|$  converges in a proper norm. For example, let  $u(x; \cdot) \in L^2(\nu)$  for all  $x$ , and assuming  $\mathcal{Z} \sim \nu$ , we expand  $u(x; z)$ :

$$u(x; z) \simeq u_N(x; z) = \sum_{n=1}^N \Psi_n(z) \hat{u}_n(x), \quad (4.1)$$

where the  $\{\Psi_n\}_{n=1}^{\infty}$  form an orthogonal basis for  $L^2(\nu)$ . For practicality the state space is usually chosen to be a Euclidean hypercube and the basis is chosen as a tensor-product of orthogonal polynomials under the appropriate one-dimension marginal distributions of  $\nu$ . In the above expansion, an approximation  $u_N$  is said to be of gPC order  $P$  if  $\Pi_P = \text{span}\{\Psi_n : n = 1, 2, \dots, N\}$ .

Once the polynomial approximation is obtained, the solution statistics can be approximated by computing the statistics of  $u_N$ . Similarly,  $u_N$  can be used to compute an approximation of  $\zeta$ , the QoI, using (2.2). Our general form for  $\zeta$  amounts to an integration over  $\mathcal{Z}$ -space so that as long as  $f(u_N)$  can be expanded as in (4.1) then computing  $\zeta$  can be accomplished by integrating the basis functions  $\Psi_n$ , for which exact formulae or quadrature may be used.

The crucial step in all of this is the computation of the gPC expansion coefficients  $\hat{u}$ . This is what separates the different flavors of polynomial chaos approximation, of which the two most popular are:

- the *Galerkin* (or intrusive) approximation, in which an approximation  $u_N \in \Pi_N$  is sought such that  $\langle \mathcal{L}(u_N - u; x; z), v \rangle_\rho = 0$  for all  $v \in \Pi_N$ , where  $\rho = d\nu/d\lambda$ .
- the *Collocation* (or nonintrusive) approximation, in which the approximation  $u_N$  is obtained from numerous samples in stochastic space  $u(x; z_j)$ ,  $j = 1, 2, \dots, M$ . The gPC coefficients  $\hat{u}_n$  in (4.1) are usually computed using either interpolation or quadrature.

In either case, we call  $u_N(z)$  the resulting approximation to  $u$  in parameter space assuming  $\mathcal{Z} \sim \nu$ . For more details on the methods, we refer the interested readers to the recent literature, for example the review article [15].

When computing the distributional sensitivity of  $\zeta(u_N)$ , we need to compute three quantities:  $E_\nu[f(u_N)]$ ,  $E_\eta[f(u_N)]$ , and  $d(\nu, \eta)$ . Depending on the choice of metric,  $d(\nu, \eta)$  can be computed via straightforward quadrature. Calculating  $E_\nu[\zeta(u_N)]$  is likewise straightforward since we have obtained a gPC approximation for the solution under the assumption that  $\mathcal{Z} \sim \nu$ .

Using these coefficients, we can interpolate the approximate solution  $u_N$  to any value of  $\mathcal{Z}$ . Therefore, we can derive an accurate quadrature rule for the new measure  $\eta$  (e.g., a Gauss rule) and use the interpolated values of  $u_N$  to compute the integral in (2.2). This method requires a gPC expansion of the QoI  $f$ , such that  $f_N$  approximate  $f$  accurately. This is usually accomplished via a stochastic collocation based method. Nevertheless, once an accurate  $f_N$  is constructed, the evaluation of DS requires manipulations of the gPC polynomials and would incur no additional stochastic simulations.

## 4.2 Algorithms based on sampling methods

Another popular stochastic method is sampling, e.g., Monte Carlo sampling, or deterministic sampling. The goal of sampling methods is often to evaluate solution statistics such as mean, (co)variance, etc., and not to construct a polynomial approximation. In this case, we can still evaluate the DS based on the sampling results, without requiring additional simulations. We start with a collection of  $M$  realizations  $u_j$ ,  $j=1, 2, \dots, M$ . These realizations are taken from stochastic locations  $z_j$ . Without reconstructing the polynomial then we have

$$E_\nu[f(u_N)] \simeq \sum_{j=1}^M f(u_j) \omega_j,$$

where  $\omega_j$  are the weights associated with the locations  $z_j$ . For Monte Carlo sampling, the weights are equal. The same quadrature rule can be used to compute  $E_\eta[f]$ :

$$E_\eta[f(u_N)] \simeq \sum_{j=1}^M f(u_j) \omega_j \times \frac{\frac{d\eta}{d\lambda}(z_j)}{\frac{d\nu}{d\lambda}(z_j)}, \quad (4.2)$$

where the last term is the ratio of densities between  $\eta$  and  $\nu$ , or the Radon-Nikodym derivative of  $\eta$  with respect to  $\nu$ . Such a derivative exists as long as  $\eta$  is absolutely continuous with respect to  $\nu$ . This is not necessarily satisfied by our assumption that  $\lambda$  dominates both measures, but in practice it is rarely a problem. The ratio of densities can also be interpreted as a likelihood ratio between the two distributions. Implementing (4.2) if we have the point-evaluations  $u_j$  is trivial and can be applied to a broad class of perturbations with minimal effort. The notion of likelihood estimation for epistemic

uncertainty in stochastic systems [8] has established literature and is a well-understood method. Since sampling methods are an example of how DS can be employed, we do not discuss the alternative formulations of likelihood sensitivity estimators here.

If the measures are connected by a parameterization in this sampling-based setup, the established practices of e.g., infinitesimal perturbation analysis and the score function method [2] are well-understood and established and may well be preferred. In any case the method (4.2) is inexpensive: the values  $f(u_j)$  need only to be computed and stored once; then (4.2) can be evaluated many times for a great variety of different likelihood ratios. One caveat with this method is that if  $d\eta/d\nu$  is not a smooth function, then the quadrature approximation (4.2) will not be accurate.

## 5 Examples

In this section we provide numerical examples. The purposes of the examples are to (1) illustrate the procedure of computing DS numerically; and (2) demonstrate convergence properties of the numerical DS. Since the topic of DS appears to yet be introduced to the UQ field, our examples all take simple form. However, they include univariate cases vs. multivariate cases, as well as stochastic ODE vs. stochastic PDE, and cover all the key ingredients of computing DS. Again, for ease of exposition, almost all examples offer sensitivity analysis for parameterized perturbations of the density.

### 5.1 One-dimensional interpolation

Suppose the solution to problem (2.1) has the form

$$u(\mathcal{Z}) = \exp[\cos(15\mathcal{Z})],$$

where  $\mathcal{Z} \in [-1, 1]$  and  $\mathcal{Z} \sim \text{Beta}(3, 2)$ . We use the gPC approximation with  $\mathcal{Z} \sim \text{Beta}(3, 2)$  and we compute an  $N$ -term Galerkin expansion (i.e., a projection) using a high-order quadrature. Having computed a gPC approximation to  $u$ , let us examine the distributional sensitivity of the mean and variance of  $u$  assuming perturbations in the parameters of the Beta distribution. The perturbed density is then  $\eta \sim \text{Beta}(3 + \Delta\alpha, 2 + \Delta\beta)$  for quantities  $\Delta\alpha$  and  $\Delta\beta$ . We first report on the values of the distributional sensitivity by plotting it against small changes in the perturbation parameters in Fig. 1. We see that the distributional sensitivity in this example exhibits a saddle-like structure, approaching one value along  $\Delta\alpha + \Delta\beta = 0$  and a different value along  $\Delta\alpha - \Delta\beta = 0$ . This shows that if epistemic uncertainty is present in the distribution, the "direction" of the perturbation is quite important. Also, this indicates that the Gateaux differentiability in Definition 2.2 is perhaps more relevant than Frechét differentiability. For this example, the distributional sensitivities of the mean and variance have very similar qualitative characteristics.

We now consider the convergence of  $DS(u_N)$  to  $DS(u)$ , where  $u_N$  is either the Galerkin or collocation gPC expansion. In Fig. 2 we show the difference between  $DS(u_N)$

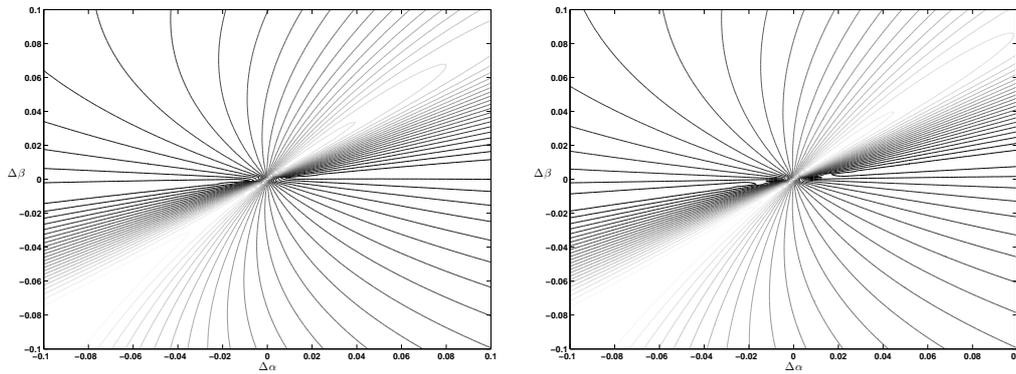


Figure 1: Contour plots of the distributional sensitivity of the mean (left) and variance (right) against the perturbation parameters  $\Delta\alpha$  and  $\Delta\beta$ . Darker lines indicates low-magnitude level sets.

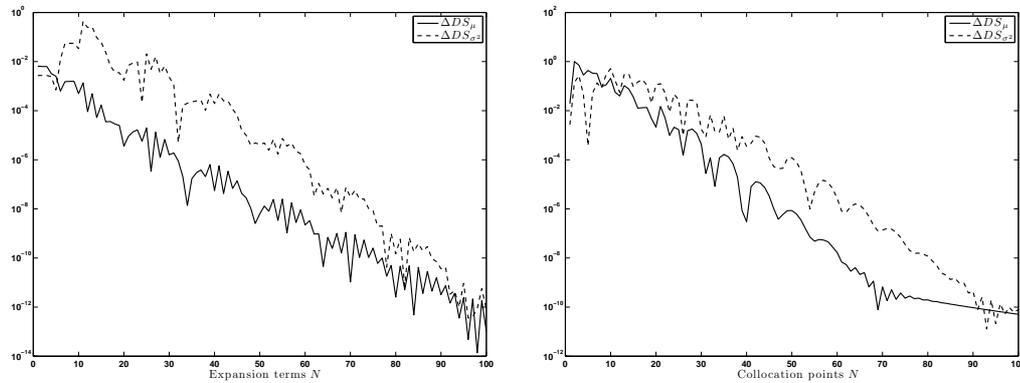


Figure 2: Plots of the mean and variance distributional sensitivity errors from the Galerkin (left) and collocation (right) expansions.

and  $DS(u)$  as a function of  $N$  when  $\Delta\alpha = -0.05$  and  $\Delta\beta = 0.05$ . The Galerkin expansion entails using a total of  $N$  gPC expansion coefficients to form the polynomial approximation  $u_N$  and to find the mean and variance of that function by interpolating it to an appropriate quadrature grid. The collocation expansion uses only  $N$  point-evaluations of the solution and computes the means and variances using (4.2). The results show that (spectral) convergence is obtained using both the Galerkin and collocation expansions; this is expected since both the projection and the interpolant converge exponentially. It is also worth noting with our choice of  $\eta \sim \text{Beta}(2.95, 2.05)$ , then

$$\|\cdot\|_{L^2(\eta)} \not\leq \|\cdot\|_{L^2(\nu)},$$

as required by the convergence proofs in Section 3. Nevertheless, we do see convergence in this example. This suggests that our requirements for the relationships between the  $\eta$  and  $\nu$  norms is a strong sufficient requirement and in practice weaker conditions are sufficient.

## 5.2 Truncation of infinite intervals

Consider now the differential equation

$$\frac{du}{dt} = -Zu, \quad t \in [0,1]; \quad u(0) = 1,$$

where our quantities of interest are the mean and variance of  $u(t=1)$ . The exact solution to this problem is  $u = e^{-Z}$ .

Using the definition we can compute the distributional sensitivity of the mean and variance between two measures: a Gaussian measure  $\nu$ ,  $Z \sim N(0,v)$  for a variance  $v$  and a second measure  $\eta$ ,  $Z \sim \text{Beta}(\alpha,\alpha)$ . We choose the state space of  $Z$  under  $\eta$  to be  $Z \in [-1,1]$ ; the variance of  $\nu$  is chosen to be  $v = 1/4\sqrt{2}$ .

The domain is not truncated in computations of the distance metric. We will compute distributional sensitivities between  $\eta$  and  $\nu$  by varying the parameter  $\alpha$ . Our method of computation will be to use an approximation  $u_N$  projected onto the  $N$ -dimensional space  $\Pi_N$  using  $\mathcal{P}_N^{(\nu)}$ . To ensure that we compute the correct quantities, Fig. 3 shows the error between the distributional sensitivities of the mean and the variance as a function of  $N$  for a fixed  $\alpha = 10$ . We see spectral convergence for our approximation of the distributional sensitivity. This is expected since this example falls under the assumptions of Proposition 3.1.

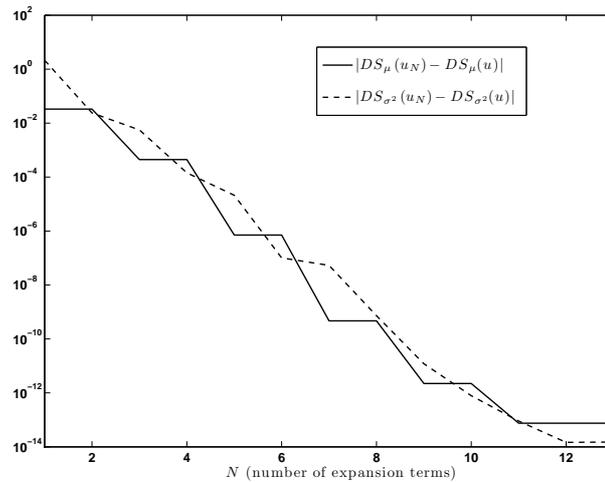


Figure 3: Error between the distributional sensitivity computed using  $u_N$  and that computed using  $u$ .  $u_N$  is the  $N^{\text{th}}$  order Galerkin approximation under the measure  $\nu$ . The exact solution  $u$  is taken as  $u_{50}$ .

Having established the accuracy for our computation of the distributional sensitivity, we use  $u_N$ ,  $N = 20$  to compute approximations to the distributional sensitivities of the variance, taking  $\alpha \in [10,20]$ . In Fig. 4 we show the  $L^1$  distance function  $d(\nu,\eta)$  and  $DS_{\sigma^2}(u_N)$  as functions of  $\alpha$ . The distance function shows a minimum near the value  $\alpha = 15$ , but  $DS_{\sigma^2}$  shows a minimum in a slightly different location. We omit the plot of the

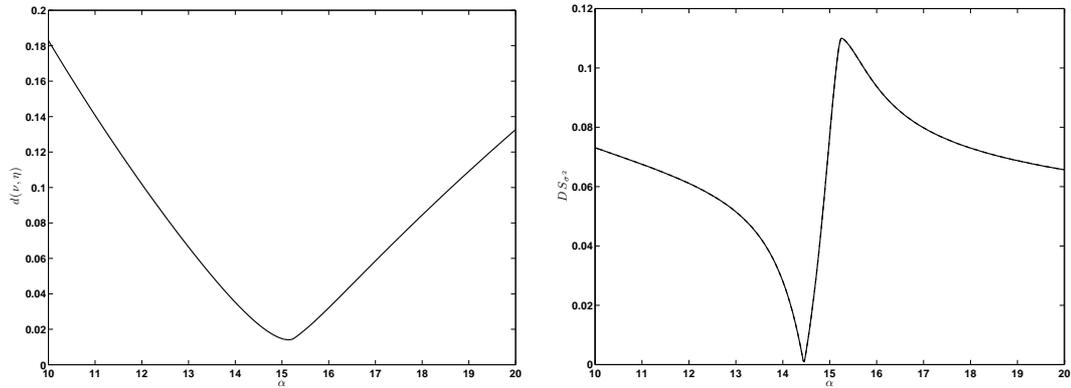


Figure 4: Plot of the  $L^1$  distributional distance  $d(v, \eta)$  (left) as a function of  $\alpha$ , the beta-distribution parameter for  $\eta$ ; also shown is the resulting distributional sensitivity of the variance  $\sigma^2$  (right).

distributional sensitivity of the mean, which shows similar qualitative features compared to the variance, and differs mainly in magnitude.

### 5.3 A diffusion example

We examine the distributional sensitivity from an elliptic problem where the diffusion is a random process [18]. The diffusion problem is given by

$$-\nabla_x(\kappa(x; \mathcal{Z}) \nabla_x u(x; \mathcal{Z})) = f(x; \mathcal{Z}). \tag{5.1}$$

Here  $x \in [0, 1]$  is the physical coordinate with homogeneous boundary conditions and  $\mathcal{Z} \in \mathbb{R}^N$  is the  $N$ -dimensional random parameter  $\mathcal{Z} = (z_1, z_2, \dots, z_N)$ . The diffusivity is given by

$$\kappa(x; z) = 1 + \sigma \sum_{k=1}^N \frac{z_k}{k^2 \pi^2} \cos(2\pi kx), \tag{5.2}$$

where  $\sigma$  measures the strength of the overall randomness. The constant term is present to ensure positivity of the diffusion. We do not allow ill-posedness of the problem and therefore restrict  $z_k$  to be bounded,  $z_k \in [-1, 1]$ . An analytic solution for (5.1) with diffusion (5.2) is not available, so we use the  $P = 8$  order gPC approximation as the true solution  $u$ . We take  $N = 4$  and  $z_k \sim \text{Beta}(\alpha, \alpha)$  for all  $k$ . Let  $\nu$  be the measure corresponding to the distribution  $\text{Beta}(0, 0)^4$  and  $\eta$  correspond to  $\text{Beta}(\alpha, \alpha)^4$ , where  $\text{Beta}(\alpha, \beta)^N$  denotes an  $N$ -variate componentwise-independent random variable with identical marginal distributions  $\text{Beta}(\alpha, \beta)$ . The quantity  $DS_\mu[v, \eta](u)$  is plotted as a function of  $\alpha$  and  $x$  in Fig. 5. In this case the mean is relatively insensitive to the distribution of the perturbed random variables. The sensitivity of the variance  $\sigma^2$  exhibits similar behavior.

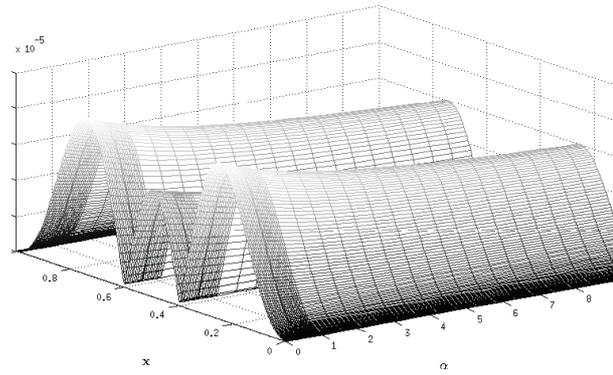


Figure 5: Plot of the distributional sensitivity  $DS_\mu(u)$  between measures  $\nu = \text{Beta}(0,0)^4$  and  $\eta = \text{Beta}(\alpha,\alpha)^4$ .

## 5.4 Supersensitivity

Finally we consider the viscous Burgers' equation

$$u_t + uu_x = \zeta u_{xx}, \quad (5.3a)$$

$$u(-1) = 1 + z_2, \quad u(1) = -1 + z_1, \quad (5.3b)$$

where  $z_2$  and  $z_1$  are random parameters. The steady-state solution has the form

$$u_\infty(x) = -A \tanh\left(\frac{A}{2\zeta}[x - \bar{\zeta}]\right),$$

for parameters  $A$  and  $\bar{\zeta}$  dependent on  $z_2$  and  $z_1$ . The steady-state solution exhibits a steep (but smooth) transition between positive and negative values at the location  $x = \bar{\zeta}$ , and  $A$  is proportional to the transition slope. By solving a nonlinear system, the exact values of  $A$  and  $\bar{\zeta}$  are computable using given values of  $\zeta$ ,  $z_2$ , and  $z_1$ . Details can be found in [17]. In this example, we fix  $\zeta = 0.1$ . We are interested in quantifying the uncertainty of the transition location  $\bar{\zeta}$ ; in particular, we want to quantify the sensitivity of the mean and variance of the transition layer with respect to the distributions of the uncertain boundary conditions, represented by  $z_2$  and  $z_1$ . Our quantities of interest are

$$\mu = E_\nu \bar{\zeta}, \quad \sigma^2 = E_\nu \bar{\zeta}^2.$$

We assume that the stochastic parameters  $z_1$  and  $z_2$  have unknown distribution, but are always restricted to the interval  $[-10^{-3}, 10^{-3}]$  and are roughly uniform. It is known that even with this small variation in the boundary conditions, the perturbation of  $\bar{\zeta}$  is  $\mathcal{O}(1)$ . We initially assume that  $z_1$  and  $z_2$  are distributed uniformly and repeatedly solve (5.3a) for multiple values of  $z_1$  and  $z_2$ , thereby obtaining a stochastic gPC approximation for  $u_\infty$  over the  $(z_1, z_2)$  parameter space. The gPC approximation can take the form of a true polynomial approximation that we can evaluate at any point in random variable space, or as a collocation approximation where we collect point evaluations and use (4.2) to compute expectations.

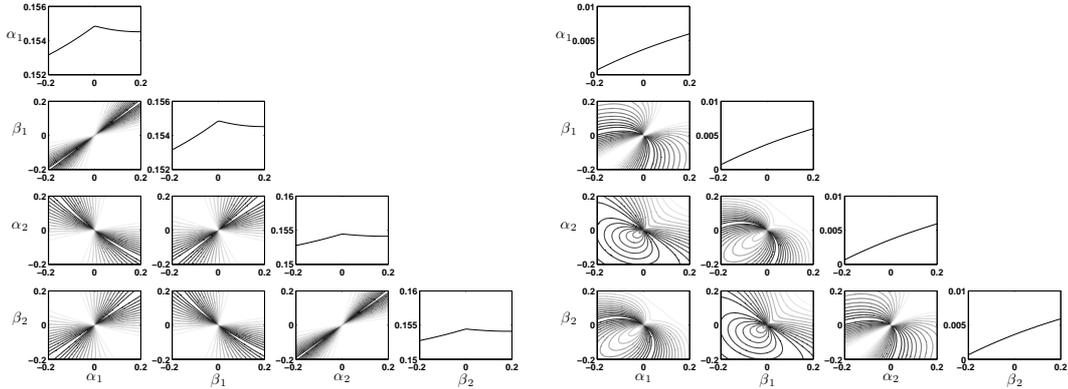


Figure 6: Contour plots of the distributional sensitivity of the mean and the variance of the transition layer location. The measures are the uniform measure  $\nu$  and the perturbed measure  $\eta = \text{Beta}(\alpha_1, \beta_1) \times \text{Beta}(\alpha_2, \beta_2)$ , where the perturbation takes the form of a change in the distribution shape parameters. The diagonal plots represent the value of the distributional sensitivity plotted against the value of the single perturbation parameter.

We denote the uniform distribution for  $(z_1, z_2)$  as the measure  $\nu$ . As a first test, we assume that the possible perturbations in the distribution of  $z_1$  takes the form  $z_1 \sim \eta_1 = \text{Beta}(\alpha_1, \beta_1)$  with both  $\alpha_1$  and  $\beta_1$  variable parameters. Similarly, we assume that  $z_2 \sim \eta_2 = \text{Beta}(\alpha_2, \beta_2)$ . Our perturbed measure is then the product  $\nu = \nu_1 \times \nu_2$ . This restricts the possible perturbations to a four-parameter family. Fig. 6 shows contour plots of the distributional sensitivity for this perturbation for various couplings of two of the four parameters; the two unused parameters are both set to 0 for each contour plot. The (unplotted) symmetry against the main diagonal corresponds to transposition of the  $x$ - $y$  axes, and some of the observed symmetry along the antidiagonal in the figure corresponds to symmetry in the parameters  $z_1$  and  $z_2$ . Our experiments show that for this example it makes little difference whether we use the approximate gPC projection by computing expansion coefficients, or the gPC collocation approach of (4.2) to compute the sensitivities.

For another example, we consider perturbation by a different four-parameter family. Now the measure  $\eta$  is defined by the density function

$$\frac{d\eta}{d\lambda^2}(z_1, z_2) = \frac{d\nu}{d\lambda^2} \times g(z_1, a_1, b_1) \times g(z_2, a_2, b_2), \tag{5.4}$$

where the parameters are  $(a_1, b_1, a_2, b_2)$  and the shape function  $g$  are defined as

$$g(x, a, b) = \begin{cases} 1, & a|x-b| \geq \frac{1}{2}, \\ 1 - \frac{1}{4}[t(x, a, b)^2 + 1][t(x, a, b)^2 - \frac{1}{2}] \exp[-t(x, a, b)^2], & a|x-b| < \frac{1}{2}, \end{cases}$$

$$t(x, a, b) = \tan[\pi a(x-b) \times 10^{-3}].$$

The parameters  $1/a$  and  $b$  correspond to affine scales and shifts, respectively, and the function  $g$  is a localized, analytic "bump" centered around  $x = b$  with approximate relative width  $1/a$ . We allow the  $a$  variables to exist over the range  $[2/3, 3/2]$  and the  $b$

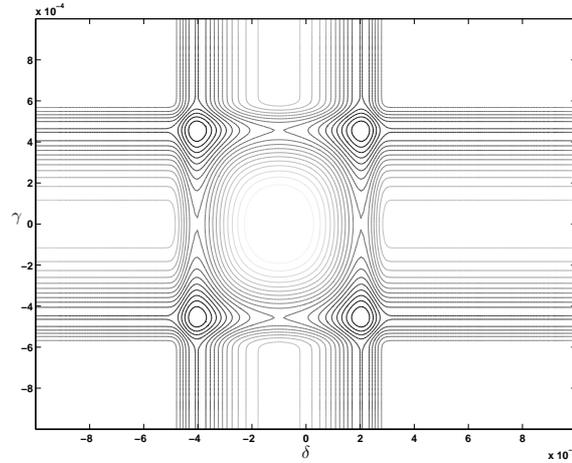


Figure 7: Contour plot of the perturbed probability density function  $d\eta/d\lambda^2$  given by (5.4). The perturbation parameters are  $a_1=0, b_1=2/3, a_2=-1 \times 10^{-4}, b_2=1$ . The axes labels correspond to  $\gamma=z_1$  and  $\delta=z_2$ .

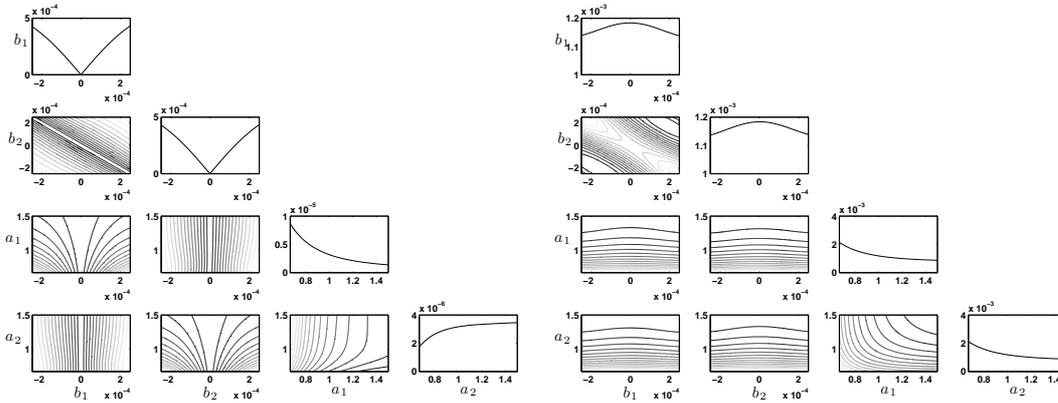


Figure 8: Contour plots of the distributional sensitivity of the mean and variance of the transition layer location as a function of four parameters defining the distributional perturbation of the boundary conditions. Darker lines indicates contours of smaller-value level sets. The diagonal plots represent the value of the distributional sensitivity plotted against the value of the single perturbation parameter.

variables to exist over the range  $[-0.25 \times 10^{-3}, 0.25 \times 10^{-3}]$ . A plot of  $dv/d\lambda^2$  (which is a probability density) for one realization of the 4-tuple  $(a_1, b_1, a_2, b_2)$  is shown in Fig. 7. With this perturbation family, the distributional sensitivity of the mean and variance are shown in Fig. 8. As before, we show contour plots when 2 of the 4 variables are allowed to vary, keeping the other two at their nominal values ( $a$  is 1, and  $b$  is 0).

Our parameters of interest here are not the mean and variance of the solution, so none of Propositions 3.1, 3.2, or 3.3 apply. Nevertheless, the gPC collocation expansion provides an inexpensive approximation for the stochastic variation of the transition location with uncertain underlying distribution.

## 6 Conclusions

We have explored the notion of distributional sensitivity (DS) for modelling systems where the underlying distribution is uncertain, where our definition of DS is extended to a collection of probability distributions that does not need to be parameterized. The concept of DS can help us to understand the relative importance (sensitivity) of the random inputs, in the sense of how change of their probability distribution functions effect the quantity of interest. We also discussed the numerical approximation of DS. When the quantity of interest is the mean or the variance of the solution, we have shown that a numerical algorithm that converges to the true solution under a measure will also yield approximations to the distributional sensitivity that converge to the true value of the distributional sensitivity.

The computation of the (discrete) distributional sensitivity can be accomplished in an efficient post-processing step that is applicable for a wide class numerical discretizations, including Monte-Carlo. With the generalized Polynomial Chaos discretization, we have shown examples that verify the proven accuracy of the computed distributional sensitivity.

## Acknowledgments

We are indebted to the anonymous referee for making us aware of the score function and Bayesian sensitivity methods.

The authors' research is supported by NSF awards DMS-0645035 and IIS-0914447, AFOSR award FA9550-08-1-0353, and DOE award DE-FC52-08NA28617.

## References

- [1] H. Arsham, A. Feuerverger, D. L. McLeish, J. Kreimer, and R. Y. Rubinstein, Sensitivity analysis and the "what if" problem in simulation analysis, *Math. Comput. Model.*, 12(2) (1989), 193–219.
- [2] S. Asmussen, and Peter W. Glynn, *Stochastic Simulation: Algorithms and Analysis*, Springer, 2007.
- [3] J. Berger, Discussion: on the consistency of bayes estimates, *Annal. Stat.*, 14(1) (1986), 30–37.
- [4] J. O. Berger, Robust bayesian analysis: sensitivity to the prior, *J. Stat. Plan. Infer.*, 25(3) (1990), 303–328.
- [5] P. Diaconis, and D. Freedman, On the consistency of bayes estimates, *Annal. Stat.*, 14(1) (1986), 1–26.
- [6] Roger G. Ghanem, and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, Inc., 1991.
- [7] A. L. Gibbs, and F. Edward Su, On choosing and bounding probability metrics, *Int. Stat. Rev.*, 70(3) (2002), 419–435.

- [8] P. W. Glynn, Likelihood ratio derivative estimators for stochastic systems, in Proceedings of the 21st Conference on Winter Simulation, pages 374–380, Washington D. C. United States, 1989, ACM.
- [9] P. Gustafson, Local sensitivity of posterior expectations, *Annal. Stat.*, 24(1) (1996), 174–195.
- [10] Y. Marzouk, and Dongbin Xiu, A stochastic collocation approach to bayesian inference in inverse problems, *Commun. Comput. Phys.*, 6(4) (2009), 826–847.
- [11] S. T. Rachev, and W. Romisch, Quantitative stability in stochastic programming: the method of probability metrics, *Math. Operat. Res.*, 27(4) (2002), 792–818.
- [12] W. Romisch, and R. Schultz, Lipschitz stability for stochastic programs with complete recourse, *SIAM J. Opt.*, 6(2) (1996), 531–547.
- [13] R. Y. Rubinstein, Sensitivity analysis and performance extrapolation for computer simulation models, *Operat. Res.*, 37(1) (1989), 72–81.
- [14] F. Ruggeri, and L. Wasserman, Infinitesimal sensitivity of posterior distributions, *Canadian J. Stat.*, 21(2) (1993), 195–203.
- [15] Dongbin Xiu, Fast numerical methods for stochastic computations: a review, *Commun. Comput. Phys.*, 5(2-4) (2009), 242–272.
- [16] Dongbin Xiu, and G. Em Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.*, 24(2) (2002), 619–644.
- [17] Dongbin Xiu, and G. Em Karniadakis, Supersensitivity due to uncertain boundary conditions, *Int. J. Numer. Methods. Eng.*, 61(12) (2004), 2114–2138.
- [18] Dongbin Xiu, and Jie Shen, Efficient stochastic galerkin methods for random diffusion equations, *J. Comput. Phys.*, 228(2) (2009), 266–281.
- [19] M. A. Zazanis, and R. Suri, Convergence rates of finite-difference sensitivity estimates for stochastic systems, *Operat. Res.*, 41(4) (1993), 694–703.