Nonlinear Studies - www. nonlinearstudies.com
MESA - www.journalmesa.com

Preprint submitted to Nonlinear Studies / MESA

Sensitivity and Uncertainty Quantification of Random Distributed Parameter Systems

Jeff Borggaard[1,⋆], Vítor Leite Nunes[1], and Hans-Werner van Wyk[2]

[1] Interdisciplinary Center for Applied Mathematics, Virginia Tech, Blacksburg, VA
[2] Department of Scientific Computing, Florida State University, Tallahassee, FL

⋆ *Corresponding Author*, e-mail: jborggaard@vt.edu

Abstract. As simulation continues to replace experimentation in the design cycle, the need to quantify uncertainty in model parameters and its effect on simulation results becomes critical. While intelligent sampling methods, such as sparse grid collocation, have expanded the class of random systems that can be simulated with uncertainty quantification, the statistical characterization of the model parameters are rarely known. In previous works, we have proposed an optimization-based framework for estimating distributed parameters as well as a number of methods for identifying the most significant parametric variations. In this work, we consider the combination of these methods. We propose the use of Fréchet sensitivity analysis to determine the most significant parametric variations (MSPVs). These MSPVs inform the generation of low order Karhunen-Loève expansions of the unknown parameter. These expansions can be used to build smooth, finite noise approximations of the parameter identification problem based on the theory of infinite dimensional constrained optimization in the space of functions with bounded mixed derivatives. Alternatively, we can interrogate the identified parametric distributions to determine the tendency of the MSPVs. We illustrate our methods with a numerical example of identifying the distributed stochastic parameter in an elliptic boundary value problem.

1 Introduction

Quantifying the uncertainty in distributed parameter systems is important for making design decisions and for managing uncertainty under operation. When the stochastic representation of model parameters is known, a number of approaches can be used to approximate the stochastic output of a model. These range from sampling methods such as Monte Carlo [1] and centriodal Voronoi tessellations [2, 3] to Galerkin approaches based on parametric expansions such as polynomial chaos [4] and Karhunen-Loève expansions [5] to sparse grid stochastic collocation approaches [6–8].

[2010] **Mathematics Subject Classification**
Keywords: 60H25 Random Operators and Equations, 60H35 Computational Methods for Stochastic Equations

In [9, 10], the authors proposed an optimization-based methodology for computing stochastic representations of random distributed parameters in partial differential equations. One outcome of this approach is an algorithm for computing statistical moments of the parameters. However, for making design decisions, one would also like to take advantage of the spatial distribution of this information. For example, one may want to address questions such as: what is the "worst" distributed perturbation of this parameter? [11], and how likely is that perturbation? To partially answer the first question, we have developed a methodology based on Fréchet sensitivity analysis [12, 13] that simultaneously computes those parametric variations that give the most dominant (in the L_2-sense) changes to the solution to a PDE, as well as the sensitivity of the solution to that parametric change. In this paper, we develop a strategy to begin addressing the second question. In other words, we use Fréchet sensitivity analysis to interrogate stochastic descriptions of distributed parameters.

2 Fréchet Sensitivity Analysis of a Diffusion Coefficient

To provide a concrete illustration of our methodology we consider the effect of the diffusion coefficient $q \in H^2(D)$ on the solution $z \in H_0^1(D) \cap H^2(D)$ of the elliptic partial differential equation

$$\begin{aligned} -\nabla \cdot (q(x)\nabla z(x)) &= f(x) \quad \text{on } D \\ z(x) &= 0 \qquad \text{on } \partial D, \end{aligned} \tag{2.1}$$

where $D \subset \mathbb{R}^d$ is a bounded domain with C^0 boundary and $f \in L_2(D)$ is a fixed forcing term. We write $z = z(q)$ to emphasize the dependence of the model output z on the input parameter q. A local measure of the sensitivity of the model output $z(q)$ to perturbations in the underlying model parameter q is given by its Fréchet derivative. For any given parameter value q_0, the Fréchet derivative of z with respect to q is the unique bounded linear operator $D_q[z(q_0)] : H^2(D) \to H_0^1(D) \cap H^2(D)$ satisfying

$$z(q_0 + h) = z(q_0) + D_q[z(q_0)]h + o(\|h\|^2), \quad \text{for } h \in H^2(D). \tag{2.2}$$

The Fréchet derivative thus defines a local linear approximation of the model response $z(q_0 + h) - z(q_0)$ to small perturbations of q_0 in the direction h. It can be shown that the derivative $s_h := D_q[z(q_0)]h$ exists in all directions h and that it can be computed as the solution of the sensitivity equation

$$\begin{aligned} -\nabla \cdot (q\nabla s_h) &= \nabla \cdot (h\nabla z(q_0)) \quad \text{in } D \\ s_h &= 0 \qquad\qquad\qquad \text{on } \partial D. \end{aligned} \tag{2.3}$$

In [12] it was further shown that the Fréchet operator $D_q[z(q_0)]$ itself is in fact a Hilbert-Schmidt operator and therefore has a singular value decomposition of the form

$$D_q[z(q_0)]h = \sum_{n=1}^{\infty} \sigma_n \langle h, v_n \rangle u_n, \tag{2.4}$$

where $\{v_n\}_{n=1}^{\infty}$ and $\{u_n\}_{n=1}^{\infty}$ are orthonormal sets in $H^2(D)$ and $H_0^1(D)$ respectively and the sequence $\{\sigma_n\}_{n=1}^{\infty}$ is a positive, non-increasing sequence in ℓ_2. This spectral decomposition provides a hierarchical ordering of the local variations of q_0 to which the model output is most sensitive, namely the directions v_n corresponding to the dominant singular values σ_n. The decomposition simultaneously gives the directions u_n in which the model responds when q_0 is perturbated in direction v_n. The extent to which an arbitrary local variation h of q_0 affects the solution z can also be quantified by its components $\sigma_n |\langle h, v_n \rangle|$ in the span of $\{u_n\}_{n=1}^{\infty}$.

Of the numerical methods presented in [12] used to estimate the singular value decomposition of $D_q[z(q_0)]$, the most intuitive, though also the most expensive, relies on approximating both the parameter space $H^2(D)$ and the solution space $H_0^1(D)$ by finite dimensional spaces $H^N = \{\phi_i, 1 \le i \le N\}$ and $S^M = \{\psi_i, 1 \le i \le M\}$ respectively. A matrix representation $D^{M,N}$ of $D_q[z(q_0)]$, for which a matrix singular value decomposition can readily be computed, is then obtained by solving the sensitivity equation (2.3) in each direction $h = \phi_i$. It was shown [12] that the approximate singular value decomposition of $D^{M,N}$ tends towards the exact decomposition as $M, N \to \infty$. A truncated singular value decomposition that requires only the first $1 \le r < \min\{M,N\}$ singular values and singular vectors can mitigate the computational cost, while still remaining within an acceptable error tolerance, since the approximation error in the Frobenius (Hilbert-Schmidt) norm is controlled by the sum of the neglected singular values. A more efficient implementation, which amounts to application of the power method to the operator $D_q[z(q_0)]D_q[z(q_0)]^*$, is discussed in [12].

The linearization (2.2), together with a truncated form of (2.4) allows us to define a local reduced order model for the input-to-output mapping, defined for values of q close to q_0 by

$$z(q) \approx z(q_0) + \sum_{n=1}^{r} \sigma_n \langle q - q_0, v_n \rangle u_n \tag{2.5}$$

The following example illustrates the uses of the singular value decomposition in a deterministic setting.

Example 1. Lets consider the following 1D elliptic equation :

$$\begin{cases} -\nabla \cdot (q \nabla z) = x(1-x),\ x \in [0,1]; \\ z \in H_0^1([0,1]). \end{cases} \tag{2.6}$$

Evaluating the solution operator $z(q)$ at $q_0(x) = e^{-10x}$ we obtain the most sensitive directions and the associated unit responses shown in Fig. 1.

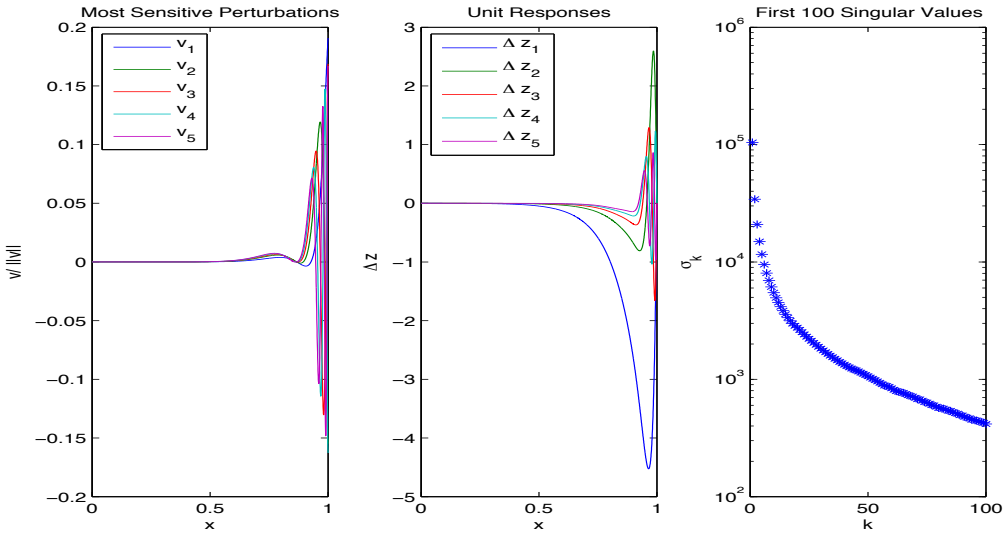

Fig. 1 Sensitive Directions

The most sensitive directions are almost zero in the interval $I_1 = [0, 0.5]$ and they oscillate away from zero in the interval $I_2 = [0.5, 1]$. This information can be translated as: $z(q_0)$ is not sensitive to perturbations on the interval I_1 whereas is highly sensitive in I_2. For instance perturbing $q_0(x)$ by $\delta_1(x) = \frac{1}{1000} e^{-14x}$ leads to an insignificant change in $z(q_0)$. On the other hand, perturbing $q_0(x)$ by $\delta_2(x) = \frac{1}{1000} e^{-14(1-x)}$ leads to a considerable change in $z(q_0)$. This can be visualized in Fig. 2.

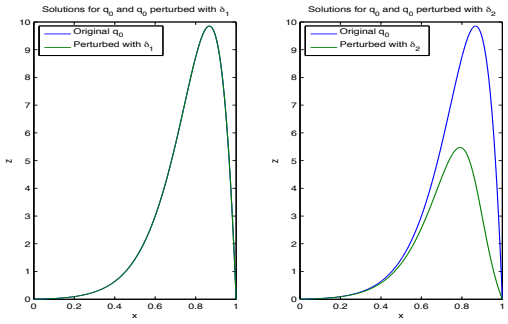

Fig. 2 Response of solutions to different perturbations

This example emphasizes the importance of determining those perturbations that produce the most significant changes in the solution. While both perturbations have the same L_2 norm $\|\delta_1\|_{L_2([0,1])} = \|\delta_2\|_{L_2([0,1])} = 1.916 \times 10^{-4}$, they have two significantly different responses. The relative change in $z(q_0)$ when q_0 is perturbed by δ_1 is

$$\frac{\|z(q_0 + \delta_1) - z(q_0)\|_{L_2([0,1])}}{\|z(q_0)\|_{L_2([0,1])}} = 9.4828 \times 10^{-5}$$

whereas an equivalent relative perturbation δ_2 leads to a substantially larger relative change in $z(q_0)$,

$$\frac{\|z(q_0(x) + \delta_2(x)) - z(q_0(x))\|_{L_2([0,1])}}{\|z(q_0(x))\|_{L_2([0,1])}} = 0.5122.$$

This can be easily explained by observing that the components of δ_2 in the first few MSPV's are substantially larger than those of δ_1. Therefore the study of the local sensitivity requires us to examine the Hilbert-Schmidt decomposition of $D_q[z(q_0)]$.

3 Elliptic Equation with a Random Diffusion Coefficient

In practice, the input parameter q is only known with a limited degree of certainty, due to the presence of unknown or unmodeled sources of variation, such as environmental factors or multiscale effects. In such cases q may be modeled as a random field $q(x, \omega)$ defined on some probability space $(\Omega, \mathcal{F}, d\omega)$ and $z = z(x, \omega)$ satisfies the stochastic equivalent of the partial differential equation (2.1), given by

$$\begin{aligned} -\nabla \cdot (q(x,\omega)\nabla z(x,\omega)) &= f(x) \quad \text{a.s. on } D \times \Omega \\ z(x,\omega) &= 0 \qquad \text{a.s. on } \partial D \times \Omega. \end{aligned} \tag{3.1}$$

Equation (3.1) can be interpreted as a family of deterministic forward problems, each corresponding to a certain 'state of the world' ω. Uncertainty in the model response z resulting from the stochastic

variability of q, is mostly quantified in terms of so-called statistical quantities of interest, such as the mean $\mathbb{E}[z(x,\cdot)]$, the covariance $\mathrm{cov}(x_1,x_2) := \mathbb{E}[z(x_1,\cdot)z(x_2,\cdot)] - \mathbb{E}[z(x_1,\cdot)]\mathbb{E}[z(x_2,\cdot)]$, or exceedance probabilities such as $\mathbb{P}(\max_{x\in D} z(x,\omega) \geq z_{\max})$, all of which can be written as an integral of the form

$$\mathrm{QoI}[z(x)] = \int_\Omega G(z(x,\omega),x)\, d\omega. \tag{3.2}$$

Monte Carlo sampling methods estimate such integrals by evaluating the integrand G at a random sample of model outputs $\{z(\cdot,\omega_i)\}_{i=1}^{n_{\mathrm{mc}}}$ corresponding to realizations $\{q(\cdot,\omega_i)\}_{i=1}^{n_{\mathrm{mc}}}$ of the stochastic input parameter and then aggregating, i.e.

$$\mathrm{QoI}^{MC}[z(x)] = \frac{1}{n_{\mathrm{mc}}} \sum_{i=1}^{n_{\mathrm{mc}}} G(z(x,\omega_i),x). \tag{3.3}$$

Often it is possible to express q as, or at least approximate it by, a function whose dependence on ω is mediated by a finite number of random variables, i.e. $q(x,\omega) = q(x,Y_1(\omega),...,Y_N(\omega))$, where $Y = (Y_1,...,Y_N) \in \Gamma$. We assume here for simplicity that the support Γ of the random vector Y is a hyper-rectangle in $\mathbb{R}^N$ and that $Y_1,...,Y_N$ have a joint density function $\rho : \Gamma \to [0,\infty)$. The stochastic forward problem (3.1) then takes the form of the 'finite noise' problem in which $z = z(x,y)$ satisfies

$$\begin{aligned} -\nabla\cdot(q(x,y)\nabla z(x,y)) &= f(x) \quad \text{a.s. on } D\times\Gamma \\ z(x,y) &= 0 \qquad \text{a.s. on } \partial D\times\Gamma, \end{aligned} \tag{3.4}$$

and related quantities of interest may be computed as

$$\mathrm{QoI}[z(x)] = \int_\Gamma G(z(x,y),x)\, \rho(y)\, dy. \tag{3.5}$$

Stochastic collocation methods provide an efficient alternative to Monte Carlo sampling when the model output $z(\cdot,y)$ depends smoothly on the random vector y and can be interpreted as a sampling-based strategy in which the sample points $y_i \in \Gamma$ are chosen to correspond with sparse grid quadrature points that are used to approximate the high dimensional integral in (3.5). A sparse grid approximation of the integral in (3.5) is therefore a weighted sum of the form

$$\mathrm{QoI}^{SC}[z(x)] = \sum_{i=1}^{n_{\mathrm{sc}}} w_i\, G(z(x,y_i),x), \tag{3.6}$$

where w_i are precomputed quadrature weights corresponding to quadrature points y_i.

Both Monte Carlo and stochastic collocation methods require a population of sample paths of the input parameter q and the response function z for the estimation of descriptive statistical quantities of interest. In this context, the information derived from the Fréchet derivative, and in particular its spectral decomposition, is more ambiguous, since it is based on a local linearization of the parameter-to-output mapping. Variations in the model output from a given sample path $z_0 = z(\cdot,\omega_{i_0})$ may have resulted from a slight perturbation of q in a direction of significant impact, or from a larger deviation where the linear approximation no longer holds. In order to obtain a description of the model sensitivity that is uniformly accurate over the parameter space, we compute a family of linearizations so that all output sample paths can be explained to within a given tolerance by one of the linear models thus obtained. This simple procedure is explained somewhat circuitously in Algorithm 1.

Compute a Family of Fréchet Operators for a Random Sample

Algorithm 1. **Input:** tol, $\{q(x,\omega_i)\}_{i=1}^{n_{\rm mc}}$, $\{z(x,\omega_i)\}_{i=1}^{n_{\rm mc}}$.
Choose $i_0 \in \{1,...,n_{\rm mc}\}$ and let $I = \{i_0\}$, $K = \text{size}(I)$
Compute the operator $D_q[z(q(\cdot,\omega_{i_0}))]$ and its singular value decomposition.
while There are paths $q(\cdot,\omega_i)$ so that

$$\|z(q(\cdot,\omega_i)) - z(q(\cdot,\omega_{i_k})) - D_q[z(q(\cdot,\omega_{i_k})](q(\cdot,\omega_i)) - q(\cdot,\omega_{i_k})\| \geq \text{tol}$$

for all $i_k \in I$ **do**
Choose $i_{K+1} \in \{1,...,n_{\rm mc}\}\setminus I$.
Compute the operator $D_q[z(q(\cdot,\omega_{i_{K+1}}))]$ and its singular value decomposition.
Let $I = \{I, i_{K+1}\}$, $K \leftarrow K+1$
end while

Note that although the samples appearing in Algorithm 1 are Monte Carlo samples indexed by $\omega \in \Omega$, this method applies equally to sparse grid samples in a stochastic collocation scheme. In the worst case, when none of the model outputs can be explained with sufficient accuracy by linearizations based at other sample paths (either due to an overly strict tolerance or a highly nonlinear parameter-to-output map) the number of linearizations required is $n_{\rm mc}$, or $n_{\rm sc}$ in the case of stochastic collocation. Generally, however, far fewer Fréchet derivatives are needed.

By combining a statistical description of the model with the spectral decomposition of the sensitivity operator (2.4) we now are able to study, as well as quantify, the interplay between parameter uncertainty and system sensitivity and hence obtain a more holistic understanding of the underlying mechanisms that cause a variation in the system response. The following example illustrates this interplay more concretely.

Example 2. Consider three stochastic systems of the form (3.1) with the same forcing term

$$f(x) = \pi^2 \sin(\pi x)(1+x) - \pi\cos(\pi x),$$

but with different, random diffusion coefficients, q_1, q_2, and q_3, given by

$$\begin{aligned}
q_1(x,\omega) &:= 1 + x + \frac{1}{2}\left(Y_1(\omega) - \frac{1}{2}\right)\cos(2\pi x) + \frac{1}{2}\left(Y_2(\omega) - \frac{1}{2}\right)\cos(3\pi x),\\
q_2(x,\omega) &:= q_1(x,\omega) + Y_3(\omega) e^{(\sqrt{1000}(x-\frac{1}{2}))^2}, \text{ and}\\
q_3(x,\omega) &:= q_1(x,\omega) + Y_3(\omega) e^{(\sqrt{1000}(x-\frac{1}{10}))^2},
\end{aligned}$$

where the random variables $(Y_1, Y_2) \sim \text{unif}([0,1]^2)$ and $Y_3 = 1 - \widetilde{Y}_3$, with $\widetilde{Y}_3 \sim \text{beta}(2,5)$ are mutually independent. The parameters q_2 and q_3 are obtained from q_1 by including a sharp increase in diffusivity at positions $x = 0.5$ and $x = 0.1$ respectively.

Fig. 3 shows the frequency plots of q_1, q_2, and q_3, while Fig. 4 shows frequency plots of the the deviations $z_i(x,\omega) - z_0(x)$ of the corresponding model outputs z_1, z_2, and z_3 from the deterministic reference output $z_0(x) = \sin(\pi x)$. Interestingly, $z_1 - z_0$ and $z_2 - z_0$ have almost identical distributions, both of which differ considerably from that of $z_3 - z_0$. It seems that a sharp increase in diffusion near the center $x = 0.5$ has almost no influence on the solution, while a similar increase at $x = 0.1$ has a considerable effect. We use a family of spectral decompositions of $D_q[z_i(q_i(\cdot,\omega_j))]$, based on a random sample of $10,000$ sample paths for each diffusion coefficient q_i, to explain this behavior. To approximate the parameter-to-output mapping to within a relative tolerance $\text{tol} = 10^{-3}$ for the

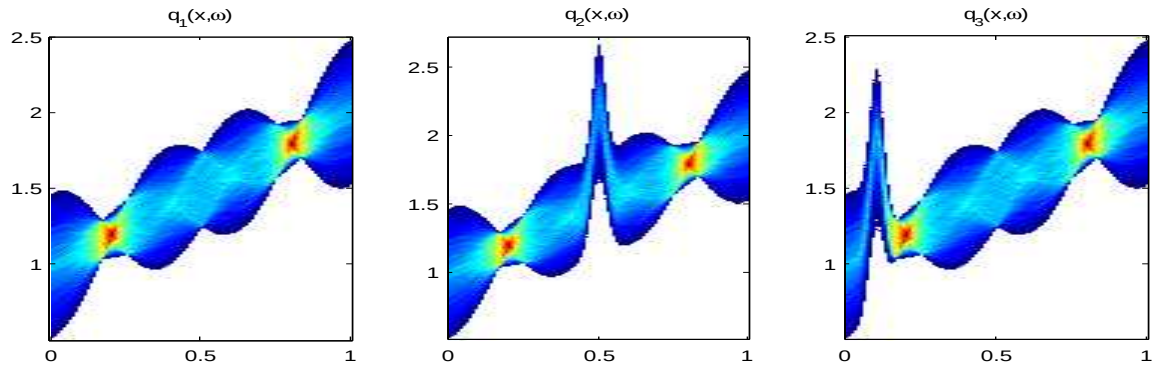

Fig. 3 Frequency plots of the densities of $q_1(x,\omega), q_2(x,\omega)$ and $q_3(x,\omega)$. The height of the density at the point (x,y) gives the likelihood that $q_i(x) = y$.

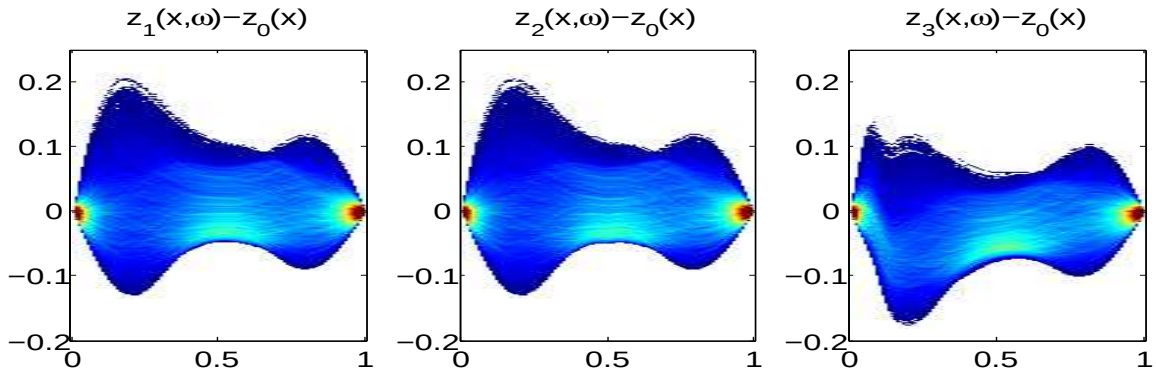

Fig. 4 Frequency plots of the densities of $z_1(x,\omega) - z_0(x), z_2(x,\omega) - z_0(x)$ and $z_3(x,\omega) - z_0(x)$.

given sample size required 7 linear models for q_1 and q_2 and 11 models for q_3. According to the decomposition (2.4), we expect sample paths whose deviations $q_i(x,\omega) - q_i(x,\omega_j)$ from their nearest linearization centers have significant components in the dominant directions $\{v_n\}_{n\approx 1}$, to have a greater effect on the model output. Those aligned with less dominant directions are likely to have a smaller effect. This holds, provided the singular values have sufficiently rapid decay. For convenience, we denote these deviations simply by $dq_i(x,\omega)$, for $i = 1,2,3$. Fig. 5 plots the first 10 right singular vectors $\{v_n\}_{n=1}^{10}$ for each of the 7 singular value decompositions of the Fréchet operators linearizing the model for q_2. Remarkably, all of the dominant singular vectors vanish at $x = 0.5$, although this is the spatial location at which q_2 shows the most variation. Thus, these perturbations are likely to have no significant effect on the output. Fig. 6 describes the cumulative contribution of the components $|\langle dq_i(\cdot,\omega), v_n\rangle|^2$ of dq_i onto the first k right basis vectors $\{v_n\}_{n=1}^{k}$ as a percentage of the total sum $\sum_{n=1}^{N_{\text{tot}}} |dq_i(\cdot,\omega), v_n|^2$, i.e. the ratio

$$\sum_{n=1}^{k} |\langle dq_i(\cdot,\omega), v_n\rangle|^2 / \sum_{n=1}^{N_{\text{tot}}} |\langle dq_i(\cdot,\omega), v_n\rangle|^2$$

for $i = 1,2,3$. For both q_1 and q_3, this ratio shows a strong initial increase, attesting to the fact that most perturbations are essentially aligned with the first few basis functions. In contrast, the ratio in the middle graph shows a much larger spread, with a significant number of sample paths whose

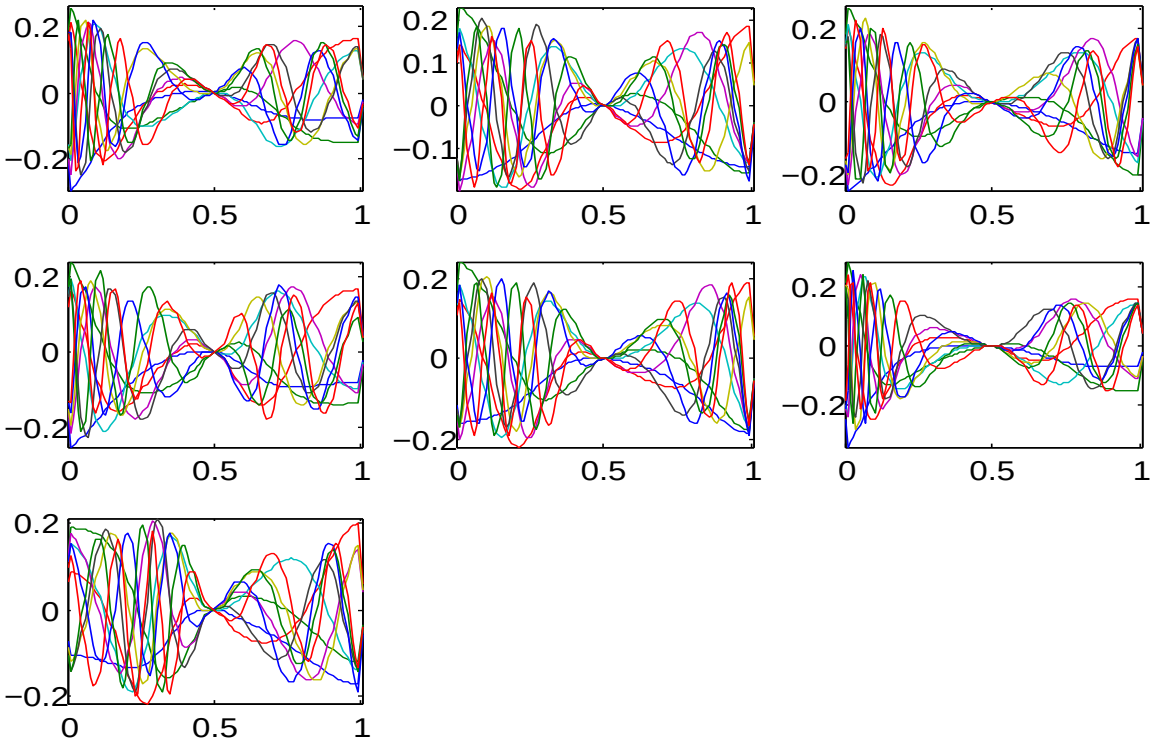

Fig. 5 First 10 right singular vectors $\{v_n\}_{n=1}^{10}$ of the spectral decompositions of the family of Fréchet operators for each of the 7 linear models of $q_2(x,\omega_j)$

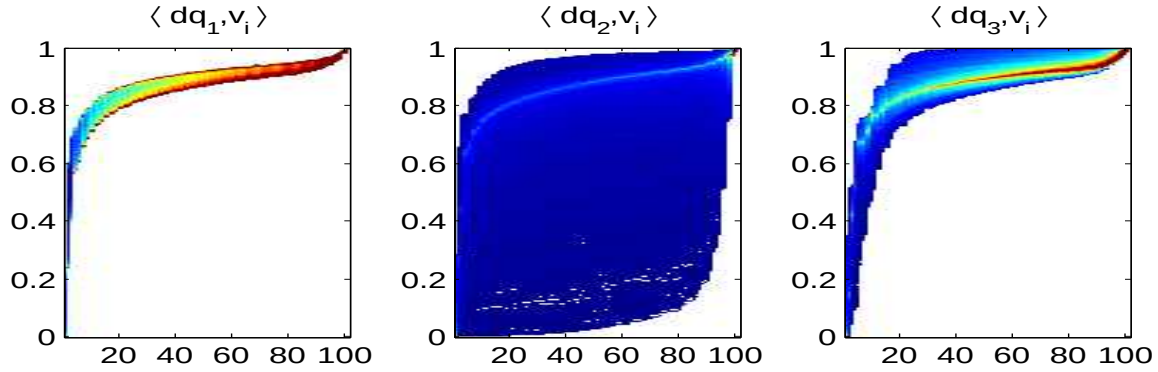

Fig. 6 Frequency distribution of $\sum_{n=1}^{k}|\langle dq_i,v_n\rangle|^2/\sum_{n=1}^{N_{\text{tot}}}|\langle dq_i,v_n\rangle|^2$ for $i=1,2,3$.

components onto the first 60 basis functions have a cumulative contribution of less than 60% to the total sum, confirming that the significant basis functions are incapable of capturing the smooth bump in q_2 at $x=0.5$.

To confirm that the model response is indeed determined by the first couple of basis functions, we show frequency plots similar to those in 6, but for deviations $dz_i = z_i(x,\omega)-z_i(x,\omega_j), i=1,2,3$ of the model output sample paths from their linearization centers (see Fig. 7). Evidently, variations in the model output are overwhelmingly determined by the first 10-20 right singular vectors v_n.

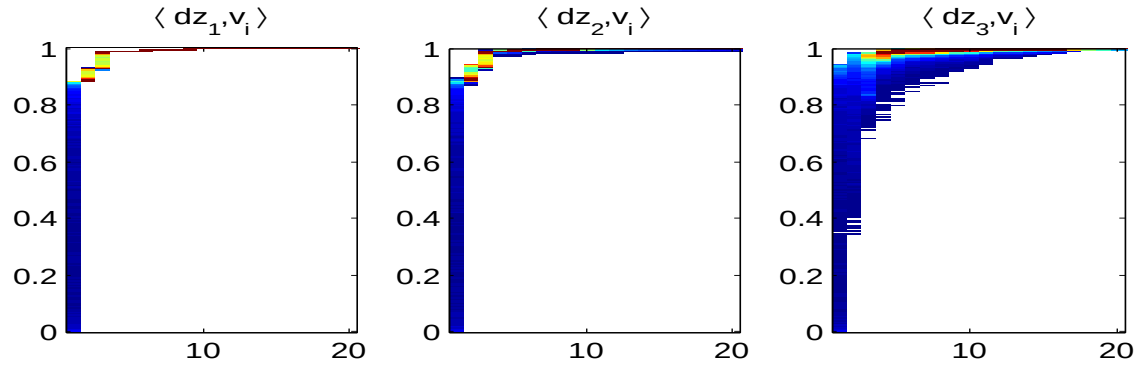

Fig. 7 Frequency distribution of $\sum_{n=1}^{k}\sigma_n^2|\langle dq_i,v_n\rangle|^2/\sum_{n=1}^{N_{\text{tot}}}\sigma_n^2|\langle dq_i,v_n\rangle|^2$ for $i=1,2,3$.

4 Sensitivity Analysis and the Identification of Uncertain Parameters

In certain cases the parameter q in (3.1) is not directly observable, and must be inferred from noisy measurements $\hat{z}$ of the model output z. If statistical information of $\hat{z}$ is available, the effect of these statistical variations on the estimated parameter q can be quantified. As in the case of the propagation of uncertainty through the forward model, the stochastic behavior of the input parameter q is usually quantified by means of statistical quantities of interest, obtained by aggregating either a Monte Carlo type sample, or the weighted sum of sample paths located at sparse grid collocation points .

In this section, we propose the use of the truncated Hilbert-Schmidt decomposition given in (2.4) to simultaneously construct a reduced order model for the stochastic input-output mapping corresponding to (3.1) and estimate a representative sample of Monte Carlo or stochastic collocation paths, $\{q(\cdot,\omega_i)\}_{i=1}^{n_{\text{mc}}}$, or $\{q(\cdot,y_i)\}_{i=1}^{n_{\text{sc}}}$ respectively. We begin by choosing an arbitrary measurement $\hat{z}(\cdot,\omega_{i_0})$ and using deterministic parameter identification techniques to obtain an estimate $\hat{q}(\cdot,\omega_{i_0})$ of the sample path $q(\cdot,\omega_{i_0})$, e.g. through minimization of the norm $\|z(\hat{q}(\cdot,\omega_{i_0})-\hat{z}(\cdot,\omega_{i_0})\|_{H_0^1}^2$ (see [14–16]). We then compute a linearization of the forward model, centered at the input-output pair $(q(\cdot,\omega_{i_0}),z(q(\cdot,\omega_{i_0})))$ as well as its (reduced) Hilbert-Schmidt decomposition characterized by the triple $\{v_n^{i_0}\}_{n=1}^r$, $\{u_n^{i_0}\}_{n=1}^r$, $\{\sigma_n^{i_0}\}_{n=1}^r$, where r can be chosen according to a desired approximation tolerance, as discussed earlier. We point out that scalar variable sensitivity analysis-enabled linear models have been used in conjunction with Monte Carlo sampling methods in earlier forward simulation studies [17]. Here, we take advantage of the fact that for sample paths $q(\cdot,\omega_i)$, $i\neq i_0$ that are 'close' to $q(\cdot,\omega_{i_o})$, the linear approximation

$$z(q(\cdot,\omega))\approx z(q(\cdot,\omega_{i_0}))+\sum_{n=1}^{r}\sigma_n\langle q(\cdot,\omega)-q(\cdot,\omega_{i_0}),v_n^{i_0}\rangle u_n^{i_0} \tag{4.1}$$

holds. We can therefore cheaply estimate these paths as expansions $\tilde{q}$ of the orthonormal set $\{v_n^{i_0}\}$ by setting

$$\hat{z}(\cdot,\omega)=z(q(\cdot,\omega_{i_0}))+\sum_{n=1}^{r}\sigma_n\langle\tilde{q}(\cdot,\omega)-q(\cdot,\omega_{i_0}),v_n^{i_0}\rangle u_n^{i_0} \tag{4.2}$$

and solving for the components $\langle\tilde{q}(\cdot,\omega),v_n^{i_0}\rangle$ to obtain

$$\tilde{q}(\cdot,\omega)=\sum_{n=1}^{r}\langle q(\cdot,\omega_{i_0}),v_n^{i_0}\rangle v_n^{i_0}+\sum_{n=1}^{r}\frac{1}{\sigma_n}\langle\hat{z}(\cdot,\omega)-z(q(\cdot,\omega_{i_0})),u_n^{i_0}\rangle v_n^{i_0}. \tag{4.3}$$

Although (4.3) defines a linear mapping from the data $\hat{z}$ to $\tilde{q}$, it nevertheless suffers from the same ill-posedness inherent in related nonlinear inverse problems. Since the terms in the second sum feature the reciprocals of the Fréchet operator's singular values $\sigma_n^{i_0}$, the components of $\hat{z}(\cdot,\omega) - z(\hat{q}(\cdot,\omega_{i_0}))$ onto the singular vectors $u_n^{i_0}$ corresponding to small singular values have a strong influence on the estimate $\tilde{q}$. Numerical errors or the presence of other, unknown sources of variation in these directions may thus have an unnecessarily significant effect on $\tilde{q}$, especially since these are in directions to which z is insensitive. This expansion, however, also suggests a straightforward regularization strategy. Restricting the number of terms r yields an approximation of (4.3) that is more stable as a mapping from $\hat{z}$ to $\tilde{q}$.

Since we do not know *a priori* for which sample paths $q(\cdot,\omega)$ this linear approximation is reasonable, it is necessary to compare the resulting model outputs $z(\tilde{q}(\cdot,\omega_i))$ with the measured data $\hat{z}$, keeping only the collection $\{\tilde{q}(\cdot,\omega_i)\}_{i\in I}$ of paths satisfying $\|\hat{z}(\cdot,\omega) - z(\tilde{q}(\cdot,\omega_i))\| < \text{tol}$. We then choose an arbitrary $i_1 \in \{1,\ldots,n_{\text{mc}}\}\setminus I$, construct a linearization at the new point and iterate the process, using the remaining sample paths, until all paths $\{q(\cdot,\omega_i)\}_{i=1}^{n_{\text{mc}}}$ have been estimated. This procedure is formalized in Algorithm 2.

As with the linearization of the stochastic forward model, the number of linearizations needed to estimate q so that the difference of the model output to $\hat{z}$ is within a given tolerance, can be n_{mc}, in the worst case. Our numerical studies, however, show this not to be the case.

Estimate the sample paths of the uncertain parameter q
based on a random sample of measurements $\hat{z}$ of the model output z.

Algorithm 2. **Input:** tol, $\{\hat{z}(x,\omega_i)\}_{i=1}^{n_{\text{mc}}}$.

Let $k=0$, $I=\emptyset$.
Choose $i_k \in \{1,\ldots,n_{\text{mc}}\}\setminus I$ and let $I=\{I,i_k\}$.
Compute the estimate $\hat{q}(\cdot,\omega_{i_k})$ of $q(\cdot,\omega_{i_k})$ from $\hat{z}(\cdot,\omega_{i_k})$.
Compute the operator $D_q[z(\hat{q}(\cdot,\omega_{i_k}))]$ and its singular value decomposition.
Use (4.3) to obtain estimates $\{\tilde{q}(\cdot,\omega_i)\}$ of $\{q(\cdot,\omega_i)\}$ for all $i \in \{1,\ldots,n_{\text{mc}}\}\setminus I$.
while there are paths $\hat{z}(\cdot,\omega_i)$ so that

$$\|z(\tilde{q}(\cdot,\omega_i)) - \hat{z}(\cdot,\omega_{i_l})\| \geq \text{tol} \tag{4.4}$$

for all i_l, $l=0,\ldots,k$. **do**
  if (4.4) doesn't hold for $i \in \{1,\ldots,n_{\text{mc}}\}\setminus I$ **then** let $I=\{I,i\}$.
  else
    Let $i_{k+1}=i$, $I=\{I,i_{k+1}\}$. $k \leftarrow k+1$.
    Repeat steps ...
  end if
end while

Example 3 (Parameter Identification). Here we apply Algorithm 2 to identify the diffusion coefficient $q_1(x,\omega)$ defined in Example 2 from stochastic measurements $\hat{z}$ of the output, using the method discussed above. In contrast to the previous example, we use a set of 65 sample paths $\{\hat{z}(\cdot,y_i)\}_{i=1}^{n_{\text{sc}}}$ corresponding to the quadrature points $y_i \in \Gamma$ of a sparse grid stochastic collocation approach, based on a Clenshaw-Curtis scheme (see [6–8, 18]). For the estimates $\tilde{q}$, we use a truncation level of $r=99$ and $r=20$ respectively. To approximate all sample paths $q(\cdot,y_i)$ so that the corresponding model output differs from the data to within an L_2-error tolerance of $\text{tol}=0.001$ requires 9 linear models for both truncation levels $r=99$ and $r=20$, the linearization centers of which are depicted in Fig. 8. Although a lower truncation level r results in a less accurate reconstruction $\tilde{q}$ of q (see Fig. 9), this

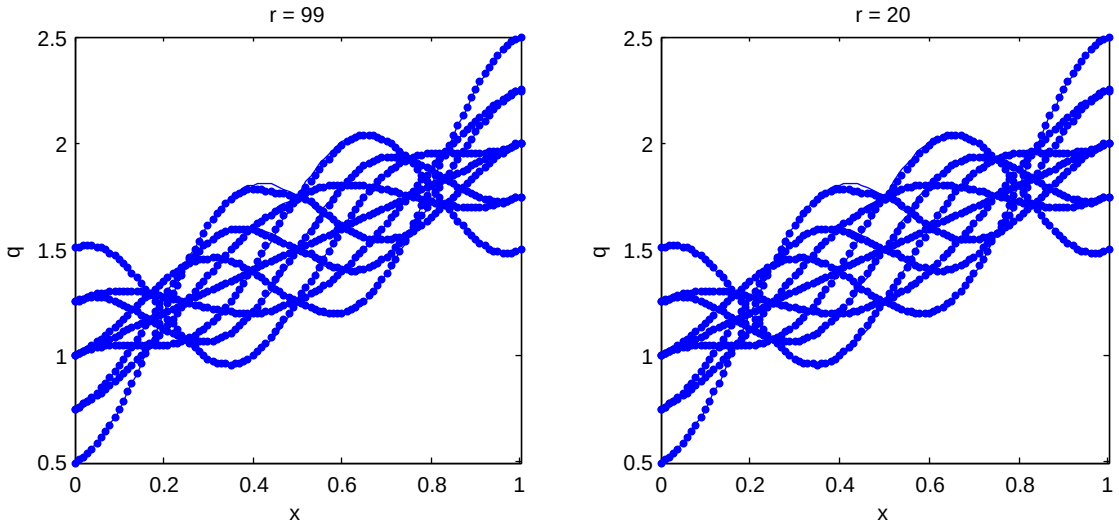

Fig. 8 Estimates $\hat{q}(\cdot, y_{i_k})$ of the linearization centers $q(\cdot, y_{i_k})$, $k = 1, \ldots, 9$ (dotted lines) together with their true values (solid lines) for both truncation levels.

doesn't seem to affect the validity of the linear model, attested to by the fact that 9 linear models with similar linearization centers (see Fig. 8) are sufficient to explain the input-output map for both $r = 99$ and $r = 20$ to within the required relative error tolerance (Fig. 10).

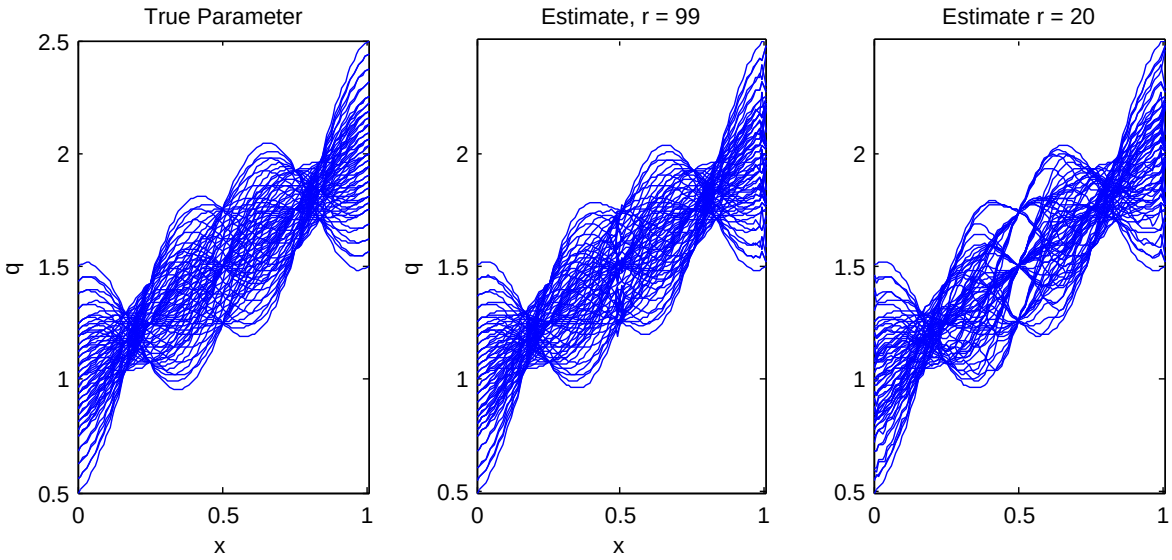

Fig. 9 Sample paths of the true parameter q as well as its estimate $\tilde{q}$.

For this example we used the same truncation levels r for the Hilbert-Schmidt decompositions of all linearizations. A more efficient implementation would adjust r for each linearization, according to some predetermined approximation tolerance.

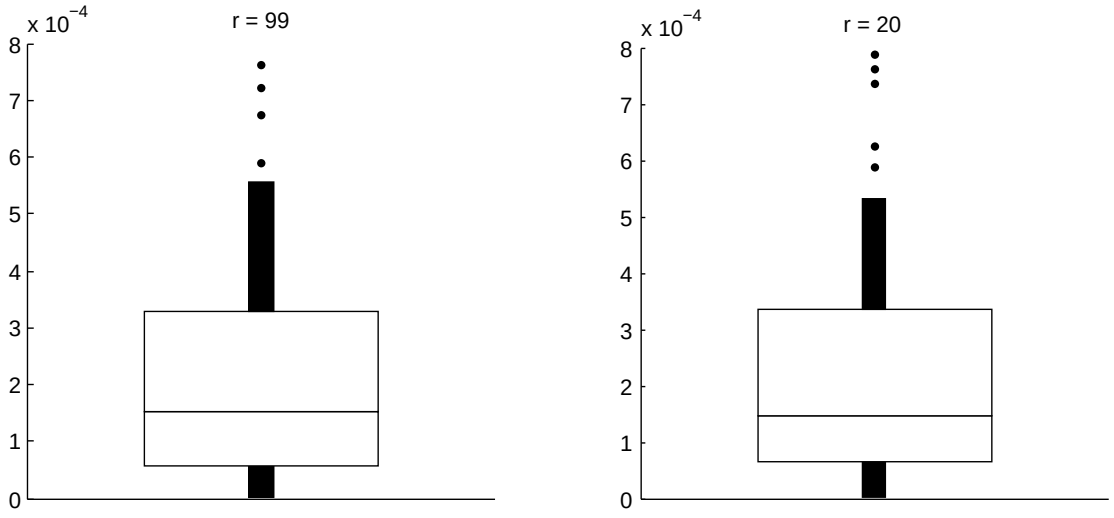

Fig. 10 Boxplot of the relative L_2-error in the model output for the parameter estimate based on truncated Hilbert-Schmidt decompositions with $r = 99$ and $r = 20$ expansion terms respectively.

5 Conclusion

In this paper, we presented two algorithms that exploit rapidly decaying Hilbert-Schmidt (singular) values of the Fréchet derivative operator in performing uncertainty quantification. We first demonstrate the ability of the Schmidt pairs to select those parametric perturbations that are significant and those that can be neglected. The first algorithm computes a family of Fréchet operators for a random sample such that the generated linearizations cover the sample space. This family of operators allow one to efficiently interrogate the sample space.

The remaining algorithm efficiently reduces the number of samples required to perform identification of random parameters. Only nine linear models were required to characterize the input-output map within a relative tolerance of 0.001. These algorithms require minimal dependence on the smoothness of the input-output map. However, when the map is not as smooth, both the number of linearizations required will increase and the rate of decay of the singular values will likely not be as sharp. The applicability of these strategies for more complex parameter estimation problems is the subject of further study.

Acknowledgements

The authors gratefully acknowledge partial support from the Air Force Office of Scientific Research under grants FA9550-10-1-0201 and FA9550-12-1-0173, the National Science Foundation under grants DMS-1016450 and EAR-0943415, and the Department of Energy under contract DE-EE0004261. The authors also appreciate comments from reviewers that improved the original manuscript.

References

[1] R. E. Caflisch. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, 7:1–49, January 1998.

[2] Y. Cao, Z. Chen, and M. Gunzburger. ANOVA expansions and efficient sampling methods for parameter dependent nonlinear PDES. *International Journal of Numerical Analysis and Modeling*, 6(2):256–273, 2009.

[3] V. J. Romero, J. V. Burkardt, M. D. Gunzburger, and J. S. Peterson. Comparison of pure and "Latinized" centroidal Voronoi tessellation against various other statistical sampling methods. *Reliability Engineering and System Safety*, 91:1266–1280, 2006.

[4] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer–Verlag, 1991.

[5] I. Babuška, R. Tempone, and G. E. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM Journal on Numerical Analysis*, 42(2):800–825, 2004.

[6] Ivo Babuška, Fabio Nobile, and Raúl Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.

[7] F. Nobile, R. Tempone, and C. G. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.

[8] C. G. Webster. *Sparse grid stochastic collocation techniques for the numerical solution of partial differential equations with random input data*. PhD thesis, The Florida State University, 2007.

[9] J. Borggaard and H.-W. van Wyk. Optimization-based estimation of random distributed parameters in elliptic partial differential equations. In *Proc. 51st IEEE Conference on Decision and Control*, pages 2926–2933, 2012.

[10] H.-W. van Wyk. *A Variational Approach to Estimating Uncertain Parameters in Elliptic Systems*. PhD thesis, Virginia Tech, 2012.

[11] M. A. Demetriou and J. Borggaard. Optimization of an integrated actuator placement and robust control scheme for distributed parameter processes subject to worst case spatial disturbance distribution. In *Proc. 2003 American Control Conference*, Denver, CO, 2003.

[12] J. Borggaard and V. Leite Nunes. Fréchet sensitivity analysis for partial differential equations with distributed parameters. In *Proc. 2011 American Control Conference*, pages 1789–1794, 2011.

[13] P. R. McGillivray and D. W. Oldenburg. Methods for calculating Fréchet derivatives and sensitivities for the non-linear inverse problem: A comparative study. *Geophysical Prospecting*, 38:499–524, 1990.

[14] H. T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems*, volume SC1 of *Systems & Control: Foundations & Applications*. Birkhäuser, 1989.

[15] K. Ito and K. Kunisch. The augmented Lagrangian method for parameter estimation in elliptic systems. *SIAM Journal on Control and Optimization*, 28(1):113–136, 1990.

[16] K. Ito, M. Kroller, and K. Kunisch. A numerical study of an augmented Lagrangian method for the estimation of parameters in elliptic systems. *SIAM Journal on Scientific and Statistical Computing*, 12(4):884–910, 1991.

[17] Y. Cao, M. Y. Hussaini, and T. A. Zang. On the exploitation of sensitivity derivatives for improving sampling methods, 2004.

[18] F. Nobile, R. Tempone, and C. G. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2411–2442, 2008.