Probabilistic Engineering Mechanics 24 (2009) 382–396

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

journal homepage: www.elsevier.com/locate/probengmech

Stochastic optimization using a sparse grid collocation scheme

Sethuraman Sankaran *

Department of Mechanical and Aerospace Engineering, EBU II 467, University of California San Diego, La Jolla, CA 92093-0411, USA

ARTICLE INFO

Article history:
Received 3 October 2007
Received in revised form
5 November 2008
Accepted 14 November 2008
Available online 27 November 2008

Keywords:
Stochastic optimization
Sparse grid collocation
Smolyak algorithm
Inverse problems
Stochastic sensitivities
Robust design

ABSTRACT

In computational sciences, optimization problems are frequently encountered in solving inverse problems for computing system parameters based on data measurements at specific sensor locations, or to perform design of system parameters. This task becomes increasingly complicated in the presence of uncertainties in boundary conditions or material properties. The task of computing the optimal probability density function (PDF) of parameters based on measurements of physical fields of interest in the form of a PDF, is posed as a stochastic optimization problem. This stochastic optimization problem is solved by dividing it into two problems—an auxiliary optimization problem to construct stochastic space representations from the PDF of measurement data, and a stochastic optimization problem to compute the PDF of problem parameters. The auxiliary optimization problem is solved using a downhill simplex method, whilst a gradient based approach is employed for solving the stochastic optimization problem. The gradients required for stochastic optimization are defined, using appropriate stochastic sensitivity problems. A computationally efficient sparse grid collocation scheme is utilized to compute the solution of these stochastic sensitivity problems. The implementation discussed, requires minimum intrusion into existing deterministic solvers, and it is thus applicable to a variety of problems. Numerical examples involving stochastic inverse heat conduction problems, contamination source identification problems and large deformation robust design problems are discussed.

1. Introduction

Stochastic optimization is the process of computing certain parameters of a system when either these parameters are random, some known properties of the system are random, and/or when measurement data are given in the form of a probability distribution function (PDF). Such stochastic problems are rife in the context of estimating boundary heat flux and heat transfer coefficients from temperature measurements, or computing initial concentration (contamination source) profiles from concentration measurements at a later time. The problem of stochastic optimization is also encountered in robust design problems (i.e. design problems wherein either known or unknown parameters of the system are stochastic).

The main goal of this work is to build an efficient computational model that can solve stochastic optimization problems. We are motivated by the following aspects, which have not been answered satisfactorily in literature—(i) How to account for data (in inverse problems) when the experiment/measurement technique is not repeatable or the data is random, (ii) How to perform estimation when the parameters are stochastic and (iii) How to design control parameters of a system when there is uncertainty in either the known or unknown parameters of a system. For example, consider the problem where we want to design the shape of the preform (raw workpiece) so that we minimize flash and underfill in a forged product. The forging velocity will be uncertain (since this is not exactly controllable) in addition to ambient temperatures, pressures etc. It is also possible that the shape of the preform cannot be exactly manufactured. We want the process to perform optimally in spite of such uncertainties. The same is the case when we try to infer probabilistic heat flux based on random temperature measurements at specific sensor locations. Essentially, we are interested in higher order statistics which have not been considered in previous research.

The transition from deterministic optimization problems to stochastic optimization problems has its share of problems and pitfalls – How will you find descent directions (in a gradient-based framework) at different iterative stages in the algorithm (which will be stochastic)? How will the random dimensions be resolved – using Monte-Carlo or more advanced techniques? The answers to these questions are quite pertinent with respect to computational efficiency, as well as to the amount of coding required to overhaul existing deterministic solvers. To simplify the discussion in the paper, we will be referring frequently to the Stochastic Inverse Heat Conduction Problem (SIHCP), though the concept is by no means restricted to this particular problem (this can be applied to a broad class of robust design problems as well). Also, we use the

* Tel.: +1 607 2274154.
E-mail address: sesankar@ucsd.edu.

0266-8920/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.probengmech.2008.11.002

terms stochastic optimization, stochastic inverse and robust design problem interchangeably.

Deterministic techniques for solving inverse problems involving heat conduction applications are detailed in [1]. Numerical techniques for inferring heat fluxes from temperature measurements using iterative regularization schemes are discussed in [2,3]. In [4], an adjoint based approach is employed to compute PDFs of heat fluxes based on PDFs of temperatures. A polynomial chaos approach is employed to resolve the stochastic dimensions present in the problem. In [5], a similar problem is resolved using a Bayesian approach.

The simplest and most intuitive means to deal with randomness in any system is the use of Monte-Carlo techniques—compute optimal parameters for many deterministic realizations of the system and thereby, compute the optimal PDF of parameters. However, the convergence of Monte-Carlo schemes is extremely slow, and in systems where it is computationally expensive to obtain solutions for deterministic algorithms, the task of stochastic optimization becomes increasingly burdensome. The idea used in this paper is based on stochastic collocation, where one carefully chooses realizations of the system where computations will be made such that the convergence with respect to the number of stochastic dimensions is significantly better than Monte-Carlo schemes. The mathematical formalism for doing the same follows from the sparse-grid stochastic collocation scheme [6].

We develop a generic mathematical framework incorporating the sparse-grid collocation framework for solving high-dimensional SIHCPs. In this aspect, we highlight the drawbacks of the stochastic adjoint approach in [4] wherein a similar problem has been dealt with—(a) inability to perform well under increasingly higher-order of stochastic dimensions (b) assumption that the data is provided directly in the stochastic space rather than constructing this stochastic space and (c) its inability to be extensible to non-linear problems wherein adjoint operators cannot be derived (refer to Problem 4 in this paper). In [4], evaluations of the temperatures done using the Generalized Polynomial Chaos approach (GPCE) involves a set of coupled equations. To overcome this disadvantage, a sparse grid stochastic collocation technique is utilized in [7] wherein the direct problem is solved at specific collocation points in the stochastic space. Each direct problem is decoupled and hence, computational complexity is significantly reduced. In addition, we develop a framework for computing stochastic sensitivities from a series of deterministic sensitivity problems. This is needed for computing stochastic gradients in the steepest gradient descent scheme. In order to compute these, we derive a scheme where we only compute deterministic sensitivities at specific collocation points. Another distinct advantage to using such an approach is the minimum coding effort to overhaul existing deterministic solvers and its applicability for systems with non-linear governing equations.

The paper is divided as follows: In Section 2, we provide some mathematical background necessary for the mathematical formalism discussed in this paper. In Section 3, we provide a working definition of stochastic inverse problems. Stochastic sensitivities are dealt with in Section 4. We discuss the mathematical procedure for performing the task of stochastic optimization in Section 5 and follow it up with some numerical examples for inverse heat conduction, inverse concentration reconstruction and robust design problems.

2. Mathematical background

This paper requires some background into probability theory, such as the definition of probability spaces, probability measures, random variables and space–time stochastic processes. There are several texts available in the literature. [8] is an excellent and comprehensive primer on probability theory. In the framework employed here, we define a stochastic space $\boldsymbol{\xi} = [\xi^1, \xi^2, \ldots, \xi^N]$ where ξ^i may represent either uniform or normally distributed random variables. Any construct on the stochastic space has a unique PDF associated with it, and we frequently work with stochastic spaces as opposed to PDFs for the analysis of random fields.

In practice, data (such as temperature measurements) is available only as a PDF. Since we work in the stochastic space for mathematical convenience, the following algorithm was developed to convert a PDF, say $p_{\text{given}}(f)$, into its representation in the stochastic space. In this paper, we restrict $\xi's$ to have either normal or uniform distribution, though the algorithm is generic in nature. Potentially, CDFs derived from the PDF can be utilized to construct the stochastic space. However, this means that an independent random variable is chosen for each sensor location, which is both impractical as well as computationally inefficient. Hence, the following algorithm is carried out for constructing the stochastic space from $p_{\text{given}}(f)$ using two choices-ξ has normal distribution and ξ has uniform distribution:

(1) Set $k = 1$
(2) Set $N = k$ and $P = 0$.
   (a) The representation in stochastic space is given by $f(\boldsymbol{\xi}) = \sum_{i=0}^{P} f_i \psi_i(\boldsymbol{\xi})$ where $\boldsymbol{\xi} = [\xi^1, \xi^2, \ldots, \xi^N]$
   (b) Compute the set of parameters f_i by optimizing: $\int (pdf(f(\boldsymbol{\xi})) - p_{\text{given}}(f))^2 \mathrm{d}f$ using the Nelder–Mead algorithm.
   (c) If the optimal objective function is greater than tolerance, set $P = P + 1$. If $P \leq 6$, continue with Step 2. If $P > 6$, go to Step 3. If the optimal objective function is less than tolerance, terminate.
(3) Set $k = k + 1$ and go to Step 2.

Nelder–Mead is a commonly used nonlinear optimization algorithm for extremely complex search spaces which is the situation herein. The technique employs the simplex technique at successive iterations, and gradually approaches the optimal solution. For further details, the interested reader is referred to [9] and *MATLAB*^{TM} is employed for the solution of this problem. This does not affect the time involved in the main algorithm, since this is used as a pre-processing step.

As part of the solution of the stochastic inverse problems, one needs to utilize computational tools for the solution of direct SPDEs. The governing SPDEs are represented in general as $\mathcal{L}u = 0$. The boundary conditions are represented as $\mathcal{B}u = 0$. The uncertainty in the input variables is commonly represented through a Karhunen–Loève (KL) expansion. This expansion can be used only if the covariance function is known a priori and hence, is only suitable for input random fields. The idea behind PCE is to perform the spectral expansion of a random process in terms of polynomial functions. The reader is referred to [10–12] for the solution to SPDEs computed using the GPCE technique. The main disadvantage of this method is that the number of terms grows combinatorially with the number of stochastic dimensions, and the equations for computing the unknown coefficients are coupled. Since the inverse problem requires repeated solutions to such direct problems, we utilize a recently developed technique, the sparse grid stochastic collocation method.

In the stochastic collocation method, we compute solutions at certain fixed locations in the multidimensional stochastic space and use interpolating functions to represent the solutions at other stochastic points [6,13,14]. The Smolyak's algorithm is a way to reduce the number of collocation points necessary for the interpolation in the multi-dimensional random space while simultaneously ensuring that the error does not increase significantly. This has been explained in [7] wherein the sparse grid interpolant is employed. Further details of the algorithm described

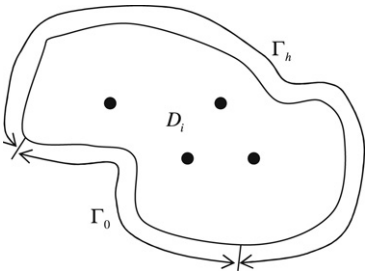

Fig. 1. The figure shows a schematic of the SIHCP problem. Γ_0 represents the boundaries where the heat flux is to be computed, Γ_h represents the domain with known heat flux and D_i represents specific points where PDF of temperature is provided. For robust design problems, only stages 2 and 3 are employed.

in this paper are given in [15–17]. Algorithms for integrations based on sparse grids are provided in [18,19]. Such a sparse grid algorithm was first used for stochastic applications recently in [7] for tackling natural convection problems and in [20] for tackling diffusion problems in random heterogeneous media.

3. The stochastic inverse heat conduction problem (SIHCP)

Let $\mathcal{D}$ be a bounded region in $\mathbb{R}^{\mathrm{d}}$, $d = 1, 2, 3$ with boundary Γ. Let the thermal conductivity $k(\mathbf{x}, \theta)$ and heat capacity $C(\mathbf{x}, \theta)$ be random fields. Let the boundary Γ be divided into Γ_h and Γ_0 with $\Gamma_h \cap \Gamma_0 = \emptyset$, where Γ_h is the part of the boundary Γ with known thermal boundary conditions (here, heat flux). There is no loss of generality in this assumption since problems where there is a boundary condition of the form $T = T_0$ can also be dealt with using the methodology given below. The PDF of heat flux on the boundary Γ_0 is considered unknown. We have to compute the PDF of the unknown stochastic flux on the boundary Γ_0 that yields the PDF of the measured stochastic temperature $Y(\mathbf{x}(D_i), t, \theta)$ at specific points D_i where $i = 1, 2, \ldots, s$, s represents the number of sensors where data is measured (ref. Fig. 1).

The stochastic partial differential equations involved in the direct heat conduction problem are summarized below:

$$
\begin{aligned}
&C\frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T), \quad (\mathbf{x}, t, \theta) \in (\mathcal{D}, \mathcal{T}, \Omega),\\
&T(\mathbf{x}, 0, \theta) = \hat{T}(\mathbf{x}, \theta), \quad (\mathbf{x}, t, \theta) \in (\mathcal{D}, \Omega),\\
&k\frac{\partial T}{\partial n}(\mathbf{x}, t, \theta) = q(\mathbf{x}, t, \theta), \quad (\mathbf{x}, t, \theta) \in (\Gamma_0, \mathcal{T}, \Omega),\\
&k\frac{\partial T}{\partial n}(\mathbf{x}, t, \theta) = \hat{f}(\mathbf{x}, t, \theta), \quad (\mathbf{x}, t, \theta) \in (\Gamma_h, \mathcal{T}, \Omega).
\end{aligned} \tag{1}
$$

The heat flux q on the boundary Γ_0 is used here as a parameter. It is apparent that for any given q, one can compute the solution $T(\mathbf{x}, t, \theta; q)$.

In the inverse problem, we are seeking a heat flux that minimizes the L_2-error norm between the measured and actual temperatures as computed at the sensor locations. In particular, one looks for a flux $\bar{q}(\mathbf{x}, t, \theta) \in L_2(\Gamma_0 \times \mathcal{T} \times \Omega)$ such that:

$$
\mathcal{J}(\bar{q}) \leq \mathcal{J}(q), \quad \forall\, q \in L_2(\Gamma_0 \times \mathcal{T} \times \Omega), \tag{2}
$$

where, $L_2(\Gamma_0 \times \mathcal{T} \times \Omega)$ is the space of all mean square integrable stochastic processes defined over the spatial and temporal domain Γ_0 and $\mathcal{T}$.

$$
\begin{aligned}
\mathcal{J}(q) &= \frac{1}{2}\|T(\mathbf{x}, t, \theta; q) - Y(\mathbf{x}, t, \theta)\|^2_{L_2(D_i \times \mathcal{T} \times \Omega)}\\
&= \frac{1}{2}\int_{\mathcal{T}}\int_{\Omega}\sum_{i=1}^{s}\{T(\mathbf{x}(D_i), t, \theta; q) - Y(\mathbf{x}(D_i), t, \theta)\}^2 \mathrm{d}P\mathrm{d}t,
\end{aligned} \tag{3}
$$

where $T(\mathbf{x}, t, \theta; q) \equiv T(\mathbf{x}, t, \theta; q(\Gamma_0, t, \theta))$ is the solution of the parametric direct stochastic heat conduction problem and $\int_{\Omega} \bullet \, \mathrm{d}P$ denotes an integral with respect to the probability measure on $(\Omega, \mathcal{F}, \mathcal{P})$.

The task of computing the solution to the inverse problem is divided into the following tasks (see also Fig. 2):

(1) *Pre-processing*—Obtaining measurement of temperature in a discrete form and computing its probability mass function.
(2) *Stochastic optimization*—
   (a) Auxiliary optimization problem defined to represent the input data in stochastic space.
   (b) Solution of direct SPDEs and sensitivity SPDEs.
   (c) Solution to the stochastic optimization method using stochastic gradient based algorithms.
(3) *Post-processing*—Conversion of discrete problem parameters to their respective PDFs.

4. Stochastic sensitivities

4.1. Definition

The main difficulty in solving the optimization problem defined in Eq. (3) is the calculation of the gradient $\mathcal{J}'(q)$ of the cost functional in the function space $L_2(\Gamma_0 \times \mathcal{T} \times \Omega)$. Stochastic sensitivities are interpreted as the change in PDF of the temperature at the sensor locations when the PDF of the heat flux q is perturbed. The sensitivity temperature field (directional derivative) $\Theta(\mathbf{x}, t, \theta; q, \Delta q) \equiv D_{\Delta q}T(\mathbf{x}, t, \theta; q)$ is defined as the linear part in Δq in the Taylor expansion of the process $T(\mathbf{x}, t, \theta; q + \Delta q)$ i.e.

$$
\begin{aligned}
T(\mathbf{x}, t, \theta; q + \Delta q) &= T(\mathbf{x}, t, \theta; q) + D_{\Delta q}T(\mathbf{x}, t, \theta; q)\\
&\quad + \mathcal{O}(\|\Delta q\|^2_{L_2(\Gamma_0 \times \mathcal{T} \times \Omega)})
\end{aligned} \tag{4}
$$

where $\Delta q \equiv \Delta q(\mathbf{x}, t, \theta)$.

4.2. Governing equations

The stochastic sensitivity problem is defined by linearization of the system of Eq. (1). The governing equations for computing the sensitivity of the temperature with respect to the heat flux are summarized below (refer [4]).

$$
\begin{aligned}
&C\frac{\partial \Theta}{\partial t} = \nabla \cdot (k \nabla \Theta), \quad (\mathbf{x}, t, \theta) \in (\mathcal{D}, \mathcal{T}, \Omega),\\
&\Theta(\mathbf{x}, 0, \theta; q, \Delta q) = 0, \quad (\mathbf{x}, \theta) \in (\mathcal{D}, \Omega),\\
&k\frac{\partial \Theta}{\partial n}(\mathbf{x}, t, \theta; q, \Delta q) = \Delta q(\mathbf{x}, t, \theta), \quad (\mathbf{x}, t, \theta) \in (\Gamma_0, \mathcal{T}, \Omega),\\
&k\frac{\partial \Theta}{\partial n}(\mathbf{x}, t, \theta; q, \Delta q) = 0, \quad (\mathbf{x}, t, \theta) \in (\Gamma_h, \mathcal{T}.\Omega).
\end{aligned} \tag{5}
$$

It is to be noted that $\Delta q(\mathbf{x}, t, \theta)$ drives the sensitivity problem and hence, it is important to define it numerically.

4.3. Numerical definition of perturbations

There are a number of issues regarding the numerical implementation of stochastic sensitivities which include—(i) What is Δq?, (ii) How do you ensure that $q + \Delta q$ remains normalized and (iii) Which technique is used to solve the set of Eq. (5). Stochastic sensitivities are essential to solve stochastic optimization problems using gradient-based approaches. Since Δq is a stochastic field, we explore if this perturbation can be defined by either perturbing specific coefficients in its spectral expansion or values of q at collocation points.

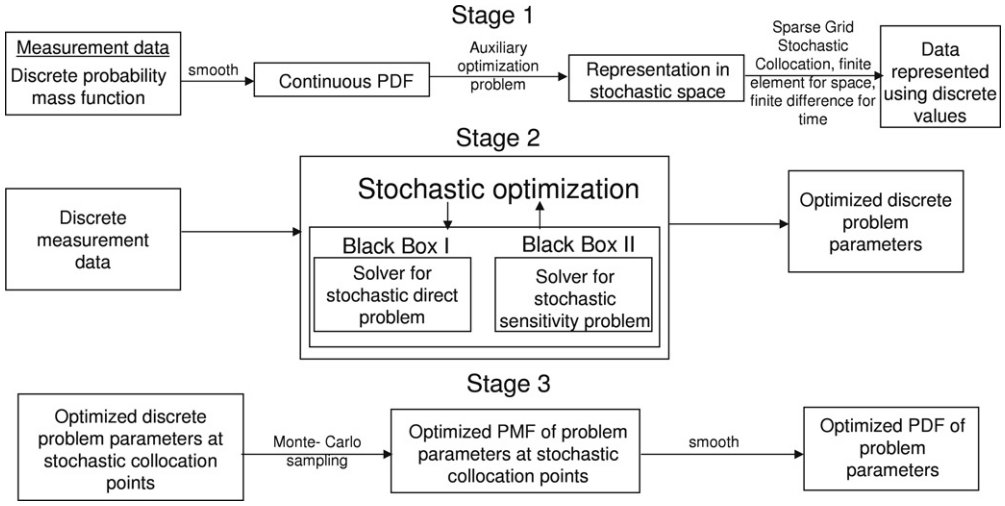

Fig. 2. The figure shows a schematic of the technique used to solve the SIHCP problem.

4.3.1. Method I: Perturbation of spectral coefficients

The unknown random parameters are represented as:

$$q(\mathbf{x}, t, \theta) = \sum_{i=0}^{N} q_i(\mathbf{x}, t)\psi_i(\boldsymbol{\xi}). \tag{6}$$

The perturbation Δq can be defined using perturbations to q_i. This is because the resultant PDF will still be normalized, as justified below.

Normalization: Let *pdf* (.) be denoted as $h(.)$ and let *cdf* (.) be denoted as $H(.)$. We have:

$$h(q) = \frac{\partial H}{\partial q}, \tag{7}$$

where $H(\hat{q}) = Prob(\sum_{i=0}^{N} q_i\psi_i(\boldsymbol{\xi}) \leq \hat{q})$. Hence,

$$\int_{-\infty}^{\infty} h(q)\mathrm{d}q = \int_{-\infty}^{\infty} \frac{\partial H}{\partial q}\mathrm{d}q = H(\infty) - H(-\infty) = 1. \tag{8}$$

Note that the GPCE should not be interpreted as a direct representation of the PDF and hence, normalization is not an explicit constraint while using GPCE. The spectral coefficients themselves do not lend any constraints to make the PDF normalized.

4.3.2. Method II: Perturbation of heat flux at collocation points

Perturbation in a collocation coefficient of a stochastic variable implies a valid perturbation of the PDF of the stochastic variable by ensuring its normality. Let us consider interpolation of the heat flux q with Lagrange polynomials using the Smolyak quadrature rule:

$$q(\mathbf{x}, t, \boldsymbol{\xi}) = \sum q_i(\mathbf{x}, t)\mathcal{L}_i(\boldsymbol{\xi}). \tag{9}$$

The proof directly follows from the relations given in the previous part. The only dependence on $\boldsymbol{\xi}$ lies in the interpolating function $\mathcal{L}_i(\boldsymbol{\xi})$. Hence, each coefficient $q_i(\mathbf{x}, t)$ may be perturbed independently and arbitrarily.

We can associate a specific $\Delta q(\mathbf{x}, t, \theta)$ with either perturbations to its GPCE coefficients or perturbations at specific collocation points as shown in Fig. 3. We also show how a perturbation in the collocation space can be converted into a corresponding perturbation in the GPCE coefficients. If $\Delta q(\mathbf{x}, t, \theta) = \Delta q_s(\mathbf{x}, t)\psi_s(\boldsymbol{\xi})$ for some s, then $\Delta q_s(\mathbf{x}, t)\psi_s(\boldsymbol{\xi}) = \sum_i \Delta q(\mathbf{x}, t, \boldsymbol{\xi}_i)L_i(\boldsymbol{\xi})$ where the $\boldsymbol{\xi}_i$'s are the cubature points and $\boldsymbol{\xi}$ is any arbitrary point in the random support space. From the last equation we can then derive that the perturbation at each collocation point is given as:

$$\Delta q(\mathbf{x}, t, \boldsymbol{\xi}_i) = \Delta q_s(\mathbf{x}, t)\psi_s(\boldsymbol{\xi}_i). \tag{10}$$

4.3.3. Implementation in spatial and temporal coordinates

In both the schemes dealt with in the previous section, the perturbation of the coefficients are in the form: $\Delta q_i(\mathbf{x}, t)$. We show how this perturbation will be implemented numerically. $q(\mathbf{x}, t)$ will be represented using its value at specific points in space and time $q(\mathbf{x}_j, t_k)$. The continuous field $q(\mathbf{x}, t)$ is extracted from $q(\mathbf{x}_j, t_k)$ by using linear interpolants in space and time. Finite element and finite-difference schemes are utilized for space and time, respectively. We sequentially choose $\Delta q_i(\mathbf{x}, t) \equiv \Delta q_i(\mathbf{x}_j, t_k)$ for each i. Hence, the number of sensitivity problems computed for each i is the product of the number of spatial and temporal degrees of freedom (DOF). Also, $\Delta q_i(\mathbf{x}_j, t_k) = \delta$, if $\mathbf{x} = \mathbf{x}_j$ and $t = t_k$ and 0 otherwise. δ is chosen to be 0.001 in all our computations.

4.4. Solution to the sensitivity equations

Using the stochastic collocation technique, the sensitivity field as defined in Eq. (5) is here represented as $\Theta(\mathbf{x}, t, \boldsymbol{\xi}; q, \Delta q) = \sum_i \Theta(\mathbf{x}, t, \boldsymbol{\xi}_i; q, \Delta q)L_i(\boldsymbol{\xi})$ where each of the $\Theta(\mathbf{x}, t, \boldsymbol{\xi}_i; q, \Delta q)$ is defined from the solution of a deterministic sensitivity problem as given in Eq. (11).

$$\begin{aligned}
&C\frac{\partial\Theta(x, t, \boldsymbol{\xi}_i; q, \Delta q)}{\partial t} = \nabla\cdot(k(\boldsymbol{\xi}_i)\nabla\Theta(\mathbf{x}, t, \boldsymbol{\xi}_i; q, \Delta q)),\\
&\Theta(\mathbf{x}, 0, \boldsymbol{\xi}_i; q, \Delta q) = 0 \quad (\mathbf{x}, \theta) \in (\mathcal{D}, \Omega),\\
&k(\boldsymbol{\xi}_i)\frac{\partial\Theta(\mathbf{x}, t, \boldsymbol{\xi}_i; q, \Delta q)}{\partial n} = \Delta q(\mathbf{x}, t, \boldsymbol{\xi}_i) \quad \mathbf{x} \in \Gamma_0\\
&k(\boldsymbol{\xi}_i)\frac{\partial\Theta(\mathbf{x}, t, \boldsymbol{\xi}_i; q, \Delta q)}{\partial n} = 0 \quad \mathbf{x} \in \Gamma_h.
\end{aligned} \tag{11}$$

Similarly, using perturbations to the GPCE coefficients, the sensitivity field can be represented as $\Theta(\mathbf{x}, t, \boldsymbol{\xi}; q, \Delta q) = \sum_i \Theta(\mathbf{x}, t, \boldsymbol{\xi}_i; q, \Delta q)\psi_i(\boldsymbol{\xi})$ where each of the $\Theta(\mathbf{x}, t, \boldsymbol{\xi}_i; q, \Delta q)$ is defined from the solution of a deterministic sensitivity problem as given in Eq. (12).

$$\begin{aligned}
&C\frac{\partial\Theta_k(\mathbf{x}, t, \theta; q, \Delta q)}{\partial t}\\
&\quad = \sum_j \nabla\cdot(\langle k(\mathbf{x}, \theta)\psi_j(\theta)\psi_k(\theta)\rangle\nabla\Theta_j(\mathbf{x}, t, \theta; q, \Delta q)),\\
&\Theta_k(\mathbf{x}, 0, \theta; q, \Delta q) = 0, \quad (\mathbf{x}, \theta) \in (\mathcal{D}, \Omega),\\
&\langle k(\mathbf{x}, t, \theta)\psi_j(\theta)\psi_i(\theta)\rangle\frac{\partial\Theta_j(\mathbf{x}, t, \theta; q, \Delta q)}{\partial n} = \Delta q_i(\mathbf{x}, t), \quad \mathbf{x} \in \Gamma_0\\
&\langle k(\mathbf{x}, t, \theta)\psi_i(\theta)\psi_j(\theta)\rangle\frac{\partial\Theta_i(\mathbf{x}, t, \theta; q, \Delta q)}{\partial n} = 0, \quad \mathbf{x} \in \Gamma_h.
\end{aligned} \tag{12}$$

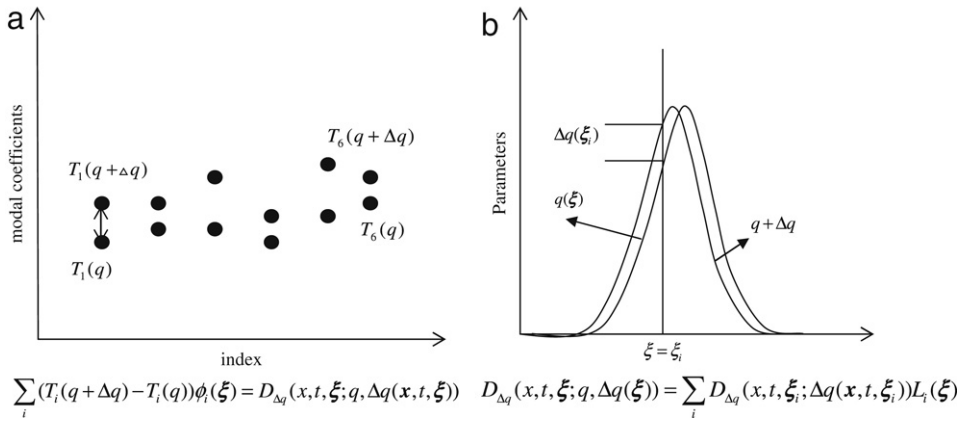

Fig. 3. The figure shows how a specific Δq can be associated with (a) on the left, the perturbation of GPCE coefficients of the heat flux and (b) on the right, the perturbations of heat flux at specific points in the stochastic space.

Note that the equations for Θ are uncoupled in Eq. (11) while they are coupled in Eq. (12). While the stochastic collocation technique was employed for solution of stochastic sensitivity PDE in Eq. (11), the GPCE technique was employed in Eq. (12).

4.5. Discrete optimization problem

It is apparent that the stochastic optimization problem explained in the previous section has to be posed discretely so that numerical schemes can be used to compute the solution. We have shown how each DOF – the spatial, temporal as well as stochastic dimensions – will be discretized. Hence, we are interested in finding the tuple, $q_v = [q_0 q_1 \dots q_{N-1} q_N]$ so as to minimize the objective function, $\mathcal{J}(q)$ (q is trivially constructed from q_v by using the corresponding interpolant).

As a result, we reduce the problem to the following form in the stochastic space (q_i here refers to the values of q at the cubature points):

$$q_v^* = \underset{q_i}{\operatorname{argmin}}\, \mathcal{J}\left(\mathbf{x}, t, \boldsymbol{\xi}; \sum_i q_i L_i(\boldsymbol{\xi})\right). \tag{13}$$

A similar problem can also be defined by using the spectral expansion on q (q_i here refers to coefficients of q in its GPCE expansion).

$$q_v^* = \underset{q_i}{\operatorname{argmin}}\, \mathcal{J}\left(\mathbf{x}, t, \boldsymbol{\xi}; \sum_i q_i.\psi_i(\boldsymbol{\xi})\right). \tag{14}$$

In this work, we will concentrate on the problem of Eq. (13). For a gradient optimization approach to this problem, we will need to utilize sensitivities of the form: $\mathcal{S}_{jki}(\mathbf{x}, t) = \frac{\Theta(\mathbf{x},t,\boldsymbol{\xi}_i;q,\Delta q(\mathbf{x}_j,t_k,\boldsymbol{\xi}_i))}{\Delta q(\mathbf{x}_j,t_k,\boldsymbol{\xi}_i)}$, where the continuum sensitivity Θ is here defined as the solution of Eq. (11) evaluated for the specific Δq as defined.

5. Optimization scheme

As we had discussed earlier, the aim is to compute a flux $\bar{q}(\mathbf{x}, t, \theta) \in L_2(\Gamma_0 \times \mathcal{T} \times \Omega)$ such that:

$$\mathcal{J}(\bar{q}) \le \mathcal{J}(q), \quad \forall\, q \in L_2(\Gamma_0 \times \mathcal{T} \times \Omega). \tag{15}$$

The strategy employed here is built on (a) the ability to solve stochastic optimization problems with minimal overhaul of existing deterministic codes and (b) be versatile enough to work for a wide range of PDEs. To satiate this need, we employ only direct and sensitivity PDE's for computing the optimal solution. It is to be noted that the efficiency of the optimization algorithm can be improved for specific problems by solving an auxiliary set of equations. For instance, the use of adjoint equations may be utilized in a SIHCP setting wherein conjugate gradient algorithms could be employed. However, the derivation of adjoint operators may not be feasible in complex fluid flow problems, and it is in this spirit that we stick to employing steepest descent schemes in this work. This ensures that the algorithm is generic in nature, while more sophisticated algorithms can be derived for specific problems.

The objective function can be written as:

$$\begin{aligned}\mathcal{J}(\mathbf{q}(\mathbf{x}, t, \boldsymbol{\xi})) = \frac{1}{2}\sum_{k=1}^{s}\int \sum_i \sum_j (T(\mathbf{x}(D_k), t, \boldsymbol{\xi}_i) \\ - Y(\mathbf{x}(D_k), t, \boldsymbol{\xi}_i))(T(\mathbf{x}(D_k), t, \boldsymbol{\xi}_j) \\ - Y(\mathbf{x}(D_k), t, \boldsymbol{\xi}_j))\mathrm{d}t \int L_i(\boldsymbol{\xi})L_j(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi},\end{aligned} \tag{16}$$

where $\int .\mathrm{d}\boldsymbol{\xi}$ implies that the integration is done as $\int .pdf(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}$. The integrals of the form $\int L_i(\boldsymbol{\xi})L_j(\boldsymbol{\xi})pdf(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}$ are computed using Monte-Carlo schemes.

The steps to be followed in performing the task of stochastic optimization are summarized below. (The measure $\mathcal{S}_{jki}$ drives the optimization procedure which indicates the variation of physical fields when parameters at specific points in space, time and stochastic space are perturbed.)

(1) Initialize values for $q(\mathbf{x}_j, t_k, \boldsymbol{\xi}_i)$, $q^0(\mathbf{x}_j, t_k, \boldsymbol{\xi}_i) = 0$. Set $k = 0$. The heat flux is $q^0(\mathbf{x}_j, t_k, \boldsymbol{\xi}) = \sum_i q^0(\mathbf{x}_j, t_k, \boldsymbol{\xi}_i)L_i(\boldsymbol{\xi})$.
(2) Solve the direct problem to compute the objective function $\mathcal{J}(q^k(\mathbf{x}, t, \boldsymbol{\xi}))$. Terminate if $k > 0$ and $\mathcal{J}(q^{k+1}) - \mathcal{J}(q^k) < tol$.
(3) Solve a set of $X \times M$ sensitivity problems where X represents the number of spatial and temporal discretizations and M denotes the stochastic discretizations of q. Compute $\mathcal{S}_{jki}(\mathbf{x}, t)$ (defined earlier) and $\mathbf{d}_{ijk} = \frac{\partial \mathcal{J}}{\partial q(\boldsymbol{\xi}_i, \tilde{\mathbf{x}}_j, \tilde{t}_k)} = \sum_{m=1}^{s}\sum_n \int (T(\mathbf{x}(D_m), t, \boldsymbol{\xi}_n; q) - Y(\mathbf{x}(D_m), t, \boldsymbol{\xi}_n))\mathcal{S}_{jki}(\mathbf{x}(D_m), t)\mathrm{d}t \int L_i(\boldsymbol{\xi})L_n(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}$ where $\mathbf{d}_{ijk}$ is written in the form of a vector, say $\mathbf{d}_s$ where s traverses the whole range of ijk.
(4) Update $k = k + 1$. $q_i^k(\mathbf{x}_j, t_k) = q_i^{k-1}(\mathbf{x}_j, t_k) + \alpha \mathbf{d}_{ijk}$ where $\alpha = -\frac{\mathbf{d}^T\mathbf{d}}{\mathbf{d}^T\mathbf{R}\mathbf{d}}$. If $s = ijk$ and $t = lmn$, then $R_{st} = \sum_{m=1}^{s} \int \mathcal{S}_{ijk}(\mathbf{x}(D_m), t)\mathcal{S}_{lmn}(\mathbf{x}(D_m), t)\mathrm{d}t$. Go to step 2.

NOTE: It is to be noted that a similar technique can also be established by utilizing the GPCE scheme for representing the unknown heat fluxes.

For the solution of eigenvalue problems while using Karhunen–Loeve expansion, the SLEPC parallel eigenvalue solver was utilized.

In addition, all computations utilized the PETSC library and were parallelized using MPI. The computing clusters available in Cornell Theory Center (CTC) were utilized for performing the computations.

6. Numerical examples

6.1. Example 1

A Gaussian triangular heat flux (see Eqs. (17) *and* (18)*) is applied at the left end of a one-dimensional heat conducting rod of length* $L = 1$ *units while the right end is insulated (motivated from* [4]*). The temperature is measured at a specific sensor location at* $x^+ = 0.3$ *in the time interval* $[0, 1]$ *and it is desired to reproduce the flux based on these temperature measurements.*

$$q^+_{\text{tri}} = \begin{cases} 2.5t^+, & 0 \le t^+ \le 0.4 \\ 2.0 - 2.5t^+, & 0.4 \le t^+ \le 0.8 \\ 0, & t^+ > 0.8 \end{cases} \tag{17}$$

$$q^+ = \mathcal{N}(q^+_{\text{tri}}, 0.1q^+_{\text{tri}}). \tag{18}$$

The system of direct governing equations are given in Eq. (1) and the system of sensitivity governing equations are given in Eq. (5).

The temperature measurements are taken at a point, $x^+ = d^+ = 0.3$ for all times $t^+ \in \mathcal{T} = [0, 1]$. Deterministic values of $k^+ = 1$ and $C^+ = 1$ were used. An explicit finite difference technique (central-difference in space and forward difference in time) with space discretization of $\Delta x^+ = 0.0025$ and time discretization $\Delta t^+ = 0.025$ was used along with Monte-Carlo technique for obtaining the random temperature 'measurements' (100 temperature measurements at each time). The auxiliary optimization scheme was run to convert the temperature measurements to its representation in the stochastic space. Since the randomness comes due to a Gaussian variation of the heat flux (Eq. (18)), a value of $N = 1$ was sufficient to represent the input randomness. The PDF of the temperature at the sensor location at a specific time and its representation in the stochastic space are shown in Figs. 4 and 5, respectively. It is to be noted that in Fig. 5, the stochastic space construct for measurement data was done using a linear polynomial where ξ is a normal variable. For visualization purposes, since the support is not finite, it was transformed into the uniform interval from 0 to 1. This is done by using the transformation $\hat{\xi} = \frac{erf(\xi)+1}{2}$ where $erf(.)$ is the error function.

The stochastic optimization scheme chalked out in this paper was utilized to compute the optimal PDF of heat flux based on the PDF of temperatures. For the solution of direct and sensitivity problems that are required during the optimization procedure, a space discretization, $\Delta x^+ = 0.025$ and a time discretization, $\Delta t^+ = 0.025$ were used in a finite element framework with linear two-noded elements. Results using the SC and GPCE based optimization algorithms are shown in Figs. 6 and 7 respectively. For performing stochastic optimization using the sparse grid collocation scheme, we used a depth of interpolation 8, and for solving using the polynomial chaos scheme, a third-order GPCE expansion was used.

In Figs. 6 and 7, it is clear that the first four moments of the heat flux is captured sufficiently. It is to be noted that there is a difference noticed at the peak value owing to a step change in the derivative. The same trend is also noticed in [1] and some problems in [4] as shown in Fig. 8 (where the solution is exactly the same everywhere except the peak values where a small error is noticed). It is clear that the stochastic optimization framework developed is accurate enough to capture randomness in the heat flux.

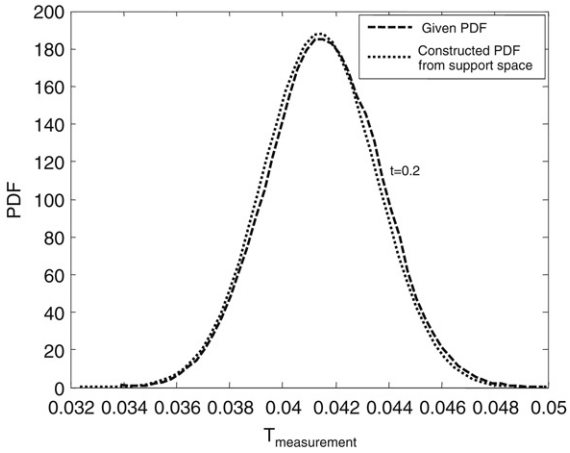

Fig. 4. The figure shows comparison of the measurement PDF with the PDF constructed by solving the auxiliary optimization problem at time $t^+ = 0.2$.

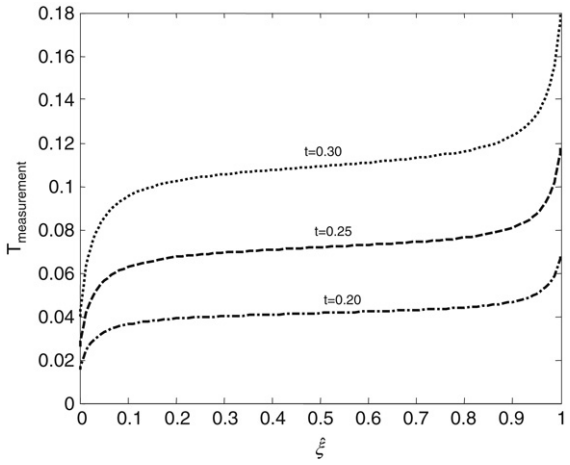

Fig. 5. The figure shows representation of measurement temperature in the support space (the corresponding PDF for $t^+ = 0.2$ is shown in Fig. 4).

NOTE: We show numerically that both the optimal heat flux as well as sensitivities computed at different stages in the spectral and collocation algorithms are identical. We construct a set of numerical tests wherein different Δq's are constructed by varying different coefficients in the GPCE expansion. Table 1 uses the measure $\mathcal{S}_{jki}(\mathbf{x}, t) = \frac{\Theta(\mathbf{x},t,\boldsymbol{\xi}_i;q,\Delta q(\mathbf{x}_j,t_k,\boldsymbol{\xi}_i))}{\Delta q(\mathbf{x}_j,t_k,\boldsymbol{\xi}_i)}$. In Table 1, we use $\mathcal{S}_{ijk}(0.3, 0.025)$ where the arguments mean that sensitivities are computed at the point $\mathbf{x} = 0.3$ and time $t = 0.025$. Also, the indices in the subscript denote how the heat flux was perturbed, as explained below:

(1) $\mathcal{S}_{jki}(.,.)$ indicates a heat flux which is perturbed at $x = x_j$, $t = t_k$ and $\boldsymbol{\xi} = \boldsymbol{\xi}_i$. Since this is a one-dimensional problem, $x_j = 0$ is the only point where we have a boundary heat flux and we specify the perturbation to be at $t_k = 0$.
(2) Each row in Table 1 indicates perturbation of a specific GPCE coefficient. d is used to indicate the dimensionality of the problem and δp indicates which term is perturbed (the perturbation has a magnitude of 0.001). For instance, a value of $\delta p = 3$ in the table means that the third GPCE term was perturbed, which is $0.5 * (3\xi^2 - 1)$ for Legendre polynomials. Using Eq. (10), the same was converted into perturbations at different stochastic collocation points (chosen as Chebychev points with a depth of interpolation, 8).
(3) The measure $\mathcal{S}_{jki}(.,.)$ is computed from $\Theta(.)$ which is computed using both the GPCE as well as collocation

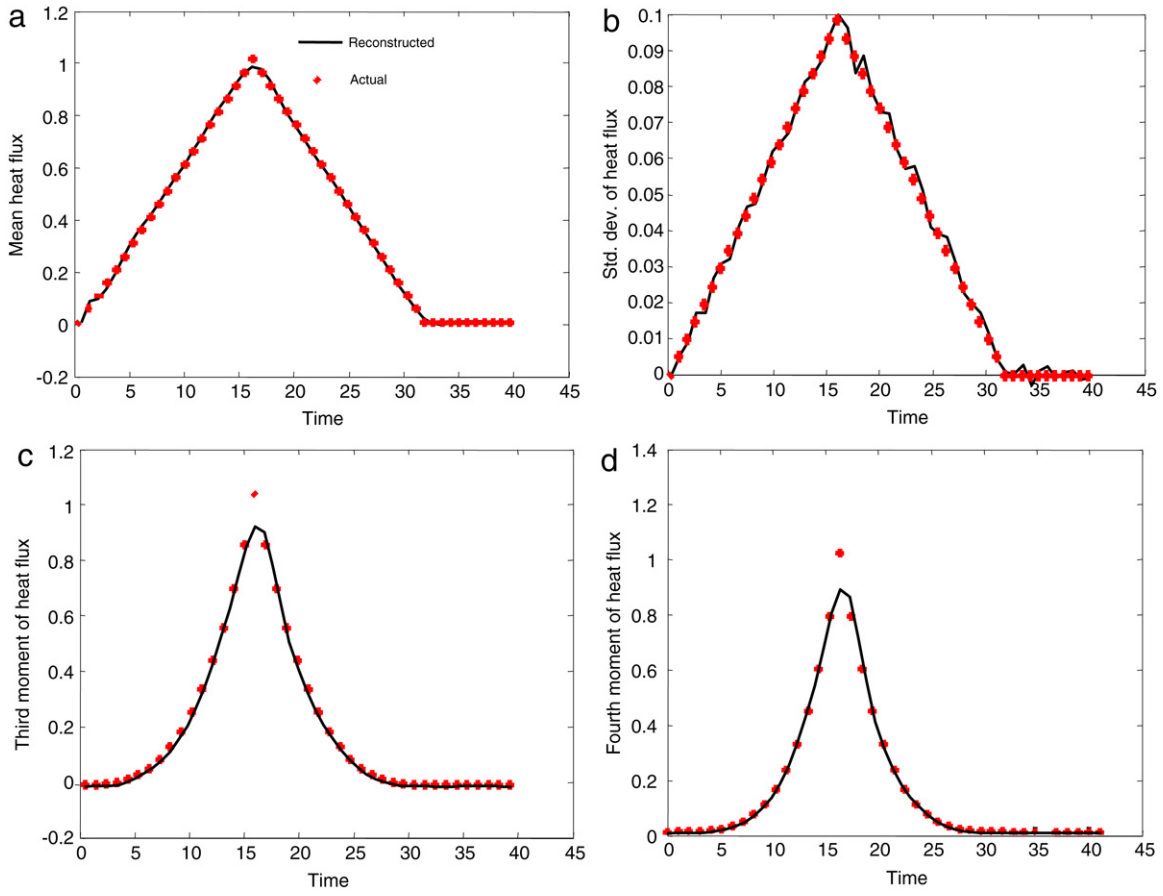

Fig. 6. The figure shows comparison of the optimal flux computed using the SC scheme with that of the actual heat flux (the first four moments). Note that nth moment means $\mathbf{E}(.)^n$ where $\mathbf{E}$ denotes expectation operator.

techniques, as defined in Eqs. (12) and (11) respectively. For stochastic collocation, statistics of $\mathcal{S}_{jki}(.,.)$ are constructed and shown in Table 1 (the statistics are taken over the index i since i represents $\boldsymbol{\xi}$). For GPCE, the values shown are statistics of $\frac{\Theta(\mathbf{x}_j,t_k,\boldsymbol{\xi};q,\Delta q)}{0.001}$. These computed statistics are denoted by $\langle x\rangle$ (mean), $\langle x^2\rangle$ (second moment) and so on in Table 1.

6.2. Problem 2: Two-dimensional SIHCP on a rolling body

In this section, a practical application of the SIHCP is discussed. Here, we consider rotating bodies that are subject to a boundary heat flux (Fig. 9). It is desired to reconstruct the heat flux based on certain temperature sensors within the rolling body.

Problem definition: A two-dimensional rolling body (Fig. 9(a)) *is subject to a unknown heat flux in one-quadrant on the outer boundary* ($-3\pi/4 \le \theta \le -\pi/4$ *where* θ *is measured from positive x-axis). The inner-boundary is insulated while the rest of the outer boundary is subject to a constant temperature of* $T = 25$ °C. *The problem is to recompute the PDF of heat flux given temperature measurements at four points as shown in* Fig. 9(b). *The inner radius is* 1.5 m *while the outer radius is* 3.0 m. *The temperature measurements are made at* $r =$ 2.8 m *and* $\theta = -135°, -45°, 45°, 135°$. *The thermal conductivity is taken to be random with a given exponential correlation function. The physical problem is treated as quasi-static.*

The following are the problem parameters used for this problem: Number of sensor nodes = 4, number of stochastic dimensions for thermal conductivity, which are known = 4 (obtained using eigenvalue decomposition and extracting eigenmodes with 99.5% energy), number of unknowns in the equivalent deterministic optimization problem = 35360, number of spatial nodes where the heat flux is unknown = 32.

We assume a quasi-steady state problem. Hence $\frac{\partial T}{\partial t} = 0$ and we do not consider boundary conditions containing time as a parameter. Naturally, the stochastic optimization algorithms detailed, are reduced to a simpler form owing to the absence of a temporal dimension.

Such problems occur frequently in processes such as rolling, where it is desired to compute heat fluxes at the roll contact regions. During the rolling process for manufacturing components, many practical problems such as the wear of the rolls, amount of coolant required at the contact zone, and stresses induced in the workpiece, as well as the rolls, is determined by the amount of heat generated at the contact zone. However, it is impractical to directly place thermo-couples in the contact zone, since they can easily wear away. Hence, a practical solution to this problem is to embed thermocouples within the rolls and recompute the heat flux on the boundary using temperature measurements within the body.

We use a random thermal conductivity defined by an exponential correlation of correlation length, 10, $f(r) = \exp(-r/10)$. The thermal conductivity, k, has the following expansion:

$$k(r,\theta,\boldsymbol{\xi}) = k_0(\theta) + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\xi_i(\theta)k_i(\mathbf{x}). \tag{19}$$

The resultant eigenvalue problem was solved and the first four modes of k were used in the final analysis, based on the eigenvalue decomposition. Some specific samples of conductivity are shown in Fig. 10.

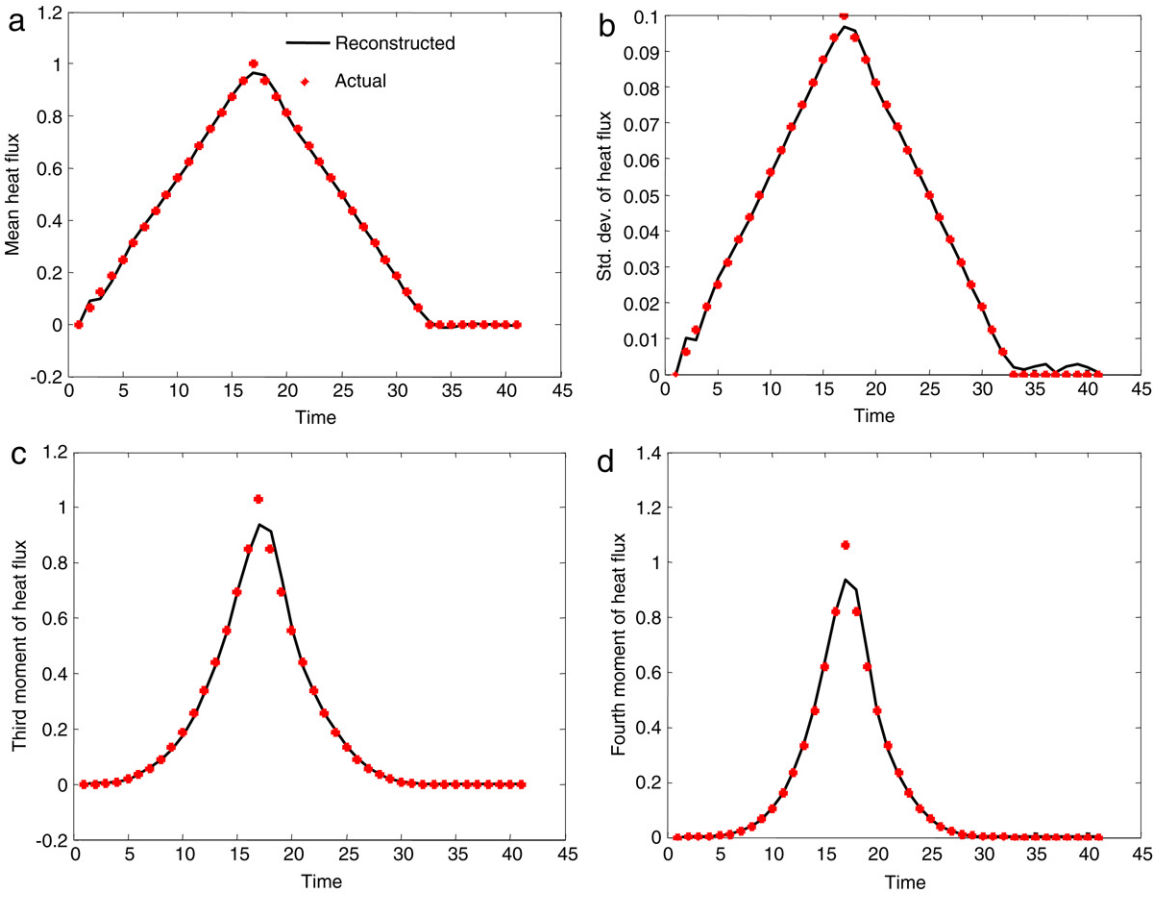

Fig. 7. The figure shows comparison of the first four moments of the optimal flux computed using the GPCE scheme with that of the actual heat flux.

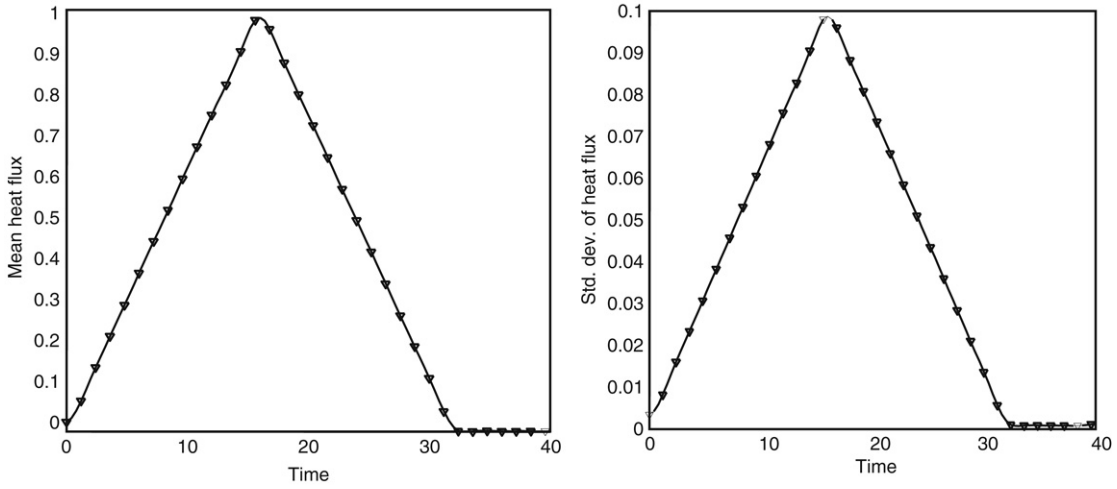

Fig. 8. The figure shows the mean and standard deviation of the heat flux computed using the adjoint method [4].

The temperature measurements obtained using the direct problem are solved using the computational method given in [21]. For the measurement data, the heat flux at the boundary is taken as Gaussian for obtaining PDF of temperature measurements. This is of the form $q_{mean}(\theta) = 100((\frac{\pi}{4})^2 - (\theta - \frac{3\pi}{2})^2)$ and $q \sim \mathcal{N}(q_{mean}, 0.1 \times q_{mean})$. A central finite difference scheme in a two-dimensional grid was employed with a Monte-Carlo scheme (1000 measurement samples) for computing PDF of temperature data. The auxiliary optimization problem converged within one additional dimension over that of the thermal conductivity (which is assumed known in the inverse problem too) due to the Gaussian nature of the heat flux.

The task of recomputing the heat flux is posed as an optimization problem. During optimization, a finite element scheme with linear triangular elements is used for the solution of direct and sensitivity problems. The governing equations of the problem in cylindrical coordinate systems are given by:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\alpha\frac{\partial T}{\partial r}\right) + \frac{1}{r^2}\frac{\partial}{\partial \theta}\alpha\frac{\partial T}{\partial \theta} - \omega\frac{\partial T}{\partial \theta} = 0, \tag{20}$$

Table 1
The table shows a comparison of stochastic sensitivities computed using two methods—the GPCE technique and the collocation technique for a perturbation in a specific term of the GPCE expansion, δp. SGSC refers to sparse grid stochastic collocation and d refers to stochastic dimension.

Case	$\langle x\rangle_{GPCE}$	$\langle x\rangle_{SGSC}$	$\langle x^2\rangle_{GPCE}$	$\langle x^2\rangle_{SGSC}$	$\langle x^3\rangle_{GPCE}$	$\langle x^3\rangle_{SGSC}$	$\langle x^4\rangle_{GPCE}$	$\langle x^4\rangle_{SGSC}$
$d=1, \delta p=1$	0.2870	0.2870	0.0824	0.0824	0.0236	0.0236	0.0068	0.0068
$d=1, \delta p=2$	0.0000	0.0000	0.0272	0.0274	0.0000	0.0000	0.0013	0.0013
$d=1, \delta p=3$	0.0000	0.0000	0.0164	0.0164	0.0013	0.0013	0.0006	0.0006
$d=2, \delta p=1$	0.2870	0.2870	0.0824	0.0824	0.0236	0.0236	0.0068	0.0068
$d=2, \delta p=4$	−0.0007	−0.0007	0.0916	0.0916	0.0000	0.0000	0.0027	0.0027
$d=2, \delta p=8$	−0.0000	0.0000	0.0055	0.0055	0.0000	0.0000	0.00012	0.00012
$d=3, \delta p=7(\times 10^{-1})$	−0.0000	0.0000	0.0915	0.0915	0.0000	0.0000	0.00271	0.00272
$d=3, \delta p=16$	−0.0007	−0.0007	0.0916	0.0916	0.0000	0.0000	0.0027	0.0027
$d=3, \delta p=20$	0.0000	−0.0001	0.0118	0.0118	0.0000	0.0000	0.0003	0.0005

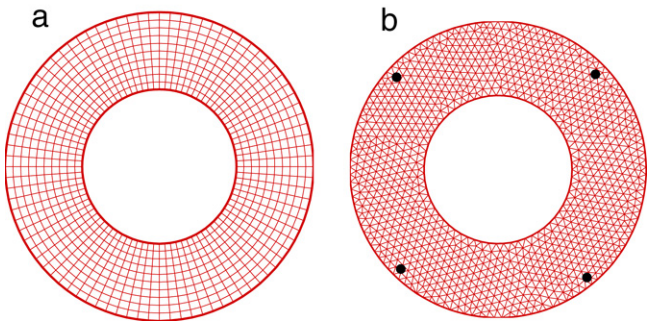

Fig. 9. (a) Mesh used for generating temperature data using forward differences and (b) Mesh generated for the optimization problem. The sensors are numbered, starting from the fourth-quadrant (clockwise) as: 1, 2, 3 and 4.

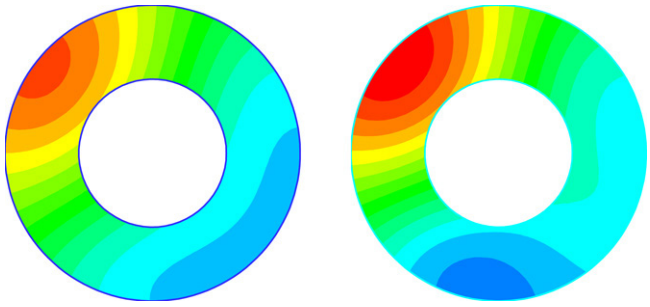

Fig. 10. Some samples of the thermal conductivity within the rolling specimen.

with the boundary conditions: $k\frac{\partial T}{\partial r}=0$ when $r=r_i, k\frac{\partial T}{\partial r}=q(r)$ when $r=r_0, -\frac{\pi}{4}\le\theta\le\frac{\pi}{4}$ and $T=T_{cons}$ for all other θ. w denotes the angular velocity and α is the thermal diffusivity.

The objective function is plotted as a function of iterations, as shown in Fig. 11 wherein a collocation scheme with depth of interpolation 5 was utilized. The convergence shown is also good (see Figs. 12 and 13).

Once the optimized values of the heat flux were obtained, we compared the temperature statistics at the four sensor locations of the optimized heat flux with the one used originally for the direct problem. Table 2 shows the comparison of these values until the fourth moment. It is apparent that the detailed methodology captures the stochastic nature of the temperature profiles. Figs. 19 and 20 show the comparison between the first four moments of the temperature profile for the actual and reconstructed temperature profile.

Further, we compare the mean heat flux in Fig. 14 and the second moment of heat flux in Fig. 15. The reason for such a large variation, as shown in the figures, is due to the fact that we do not have sufficient sensors along the contact region. However, it is clear that the objective function for this heat flux is almost 0.

To test if the effect shown above was due to an insufficient number of sensors alone, we solved the same problem, but now

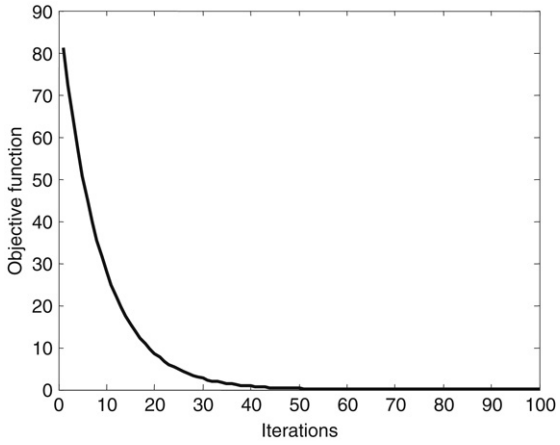

Fig. 11. Objective function computed at different iterations for IHCP problem on a rolling body.

with an increased number of twelve sensors (eight additional sensors in the contact region). Figs. 16 and 17 show the results with an increased number of sensors. It is clear from the results that an important reason for the inadequate reconstruction using four sensors was the insufficient number of sensors, and as we obtain more information, we can reconstruct the actual profile of the heat flux better.

6.3. Problem 3: 2D concentration reconstruction in porous media

In this section, stochastic inverse problems for 2D reconstruction of concentration profiles is undertaken. There is a porous medium and some contaminant is injected into it. Based on the concentration profile of the contaminant at a later time, it is desired to reconstruct the initial contaminant concentration. Here, we deal with a stochastic version of this problem, which is defined as: *Compute the PDF of concentrations at a specific time based on concentration PDF's at certain sensor locations at a later time*. Herein, we deal with uncertainties in initial source concentration. This stochastic inverse problem is posed as a stochastic optimization problem, and the methodology detailed in this paper is utilized for computing the solution. For similar problems which were tackled before, the reader is referred to [22–25].

The equations which govern flow in porous media (originating from conditions of mass balance) are given below

$$\phi\frac{\partial c}{\partial t}+\nabla\cdot(c\mathbf{u})-\nabla\cdot(\mathbf{D}\nabla c)=\tilde{c}q,\quad c\in(\mathcal{D},\mathcal{T},\Omega), \tag{21}$$

$$\nabla\cdot\mathbf{u}=q,\quad t\in[0t_l] \tag{22}$$

$$\mathbf{u}=-\frac{K(\mathbf{x})}{\mu(c)}\nabla p,$$

$$c(\mathbf{x},0,\boldsymbol{\xi})=c_0(\mathbf{x},\boldsymbol{\xi}),$$

$$\mathbf{D}\nabla c\cdot\mathbf{n}=0,\quad t\in[0\ t_l],$$

$$\nabla\mathbf{u}.\mathbf{n}=0,\quad t\in[0\ t_l].$$

Table 2
The figure shows the comparison of temperature statistics at the four sensor locations shown before.

Sensor location	$\langle T(x_d)\rangle$ computed	$\langle T(x_d)\rangle$ actual	$\langle (T(x_d))^2\rangle - \langle T(x_d)\rangle^2$ computed $\times$ 1e−2	$\langle (T(x_d))^2\rangle - \langle T(x_d)\rangle^2$ actual $\times$ 1e−2	$\langle (T(x_d))^3\rangle - \langle T(x_d)\rangle^3$ computed $\times$ 1e−3	$\langle (T(x_d))^3\rangle - \langle T(x_d)\rangle^3$ actual $\times$ 1e−3	$\langle (T(x_d))^4\rangle - \langle T(x_d)\rangle^4$ computed $\times$ 1e−4	$\langle (T(x_d))^4\rangle - \langle T(x_d)\rangle^4$ actual $\times$ 1e−4
1	45.6594	47.3247	21.154	22.804	99.791	112.32	481.201	567.39
4	25.86	25.975	6.687	6.7475	17.30	17.53	44.73	45.54
3	25.109	25.125	6.305	6.313	15.830	15.861	39.748	39.85
2	26.559	26.703	7.056	7.133	18.756	19.059	49.870	50.943

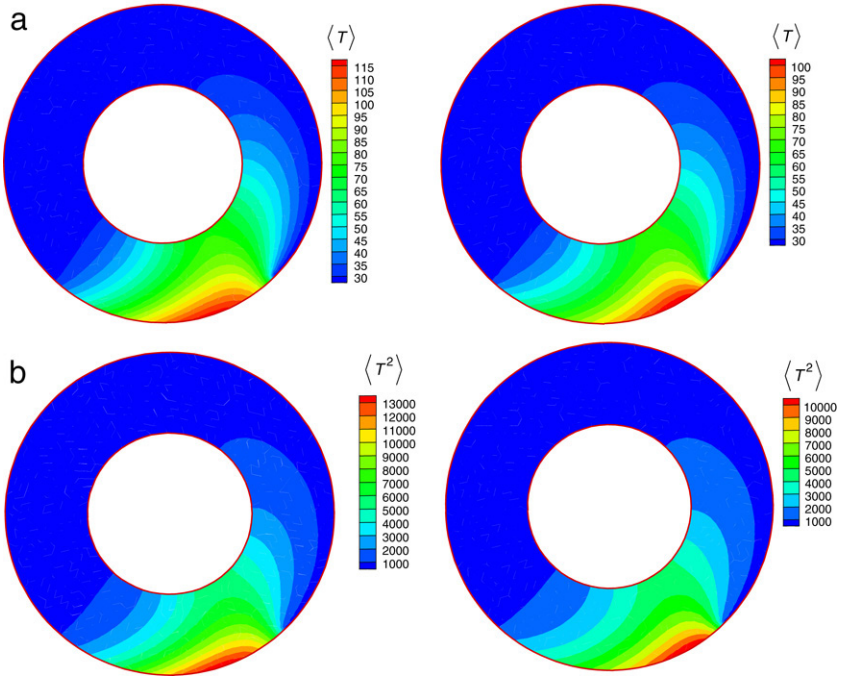

Fig. 12. The figure shows (a) top, the comparison of mean temperature profiles using the optimized heat flux with that using the original heat flux and (b) bottom, comparison of the second moments.

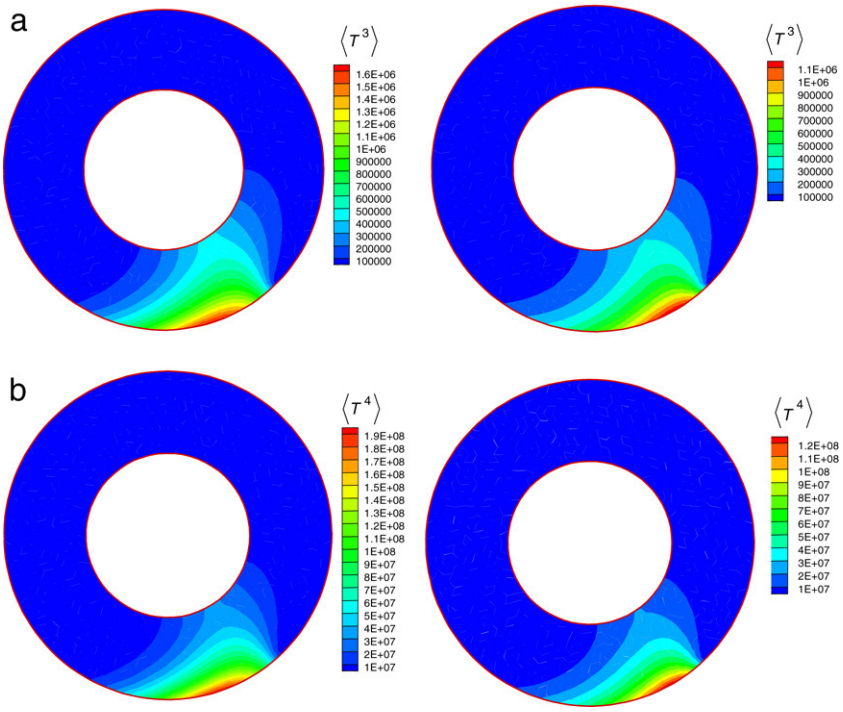

Fig. 13. The figure shows (a) top, the comparison of third and (b) bottom, fourth moments of temperature profiles of the optimized heat flux versus the actual value of the heat flux used.

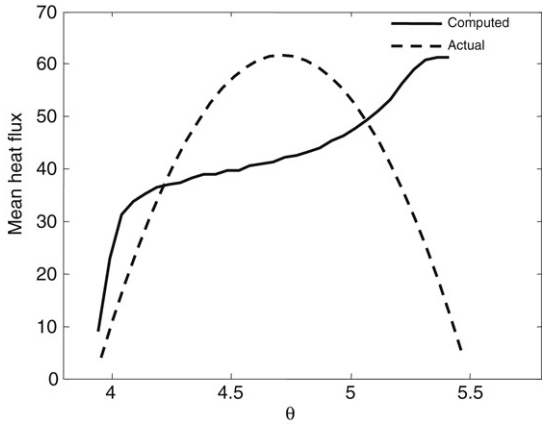

Fig. 14. The figure shows comparison of mean heat flux recomputed using the stochastic optimization algorithm.

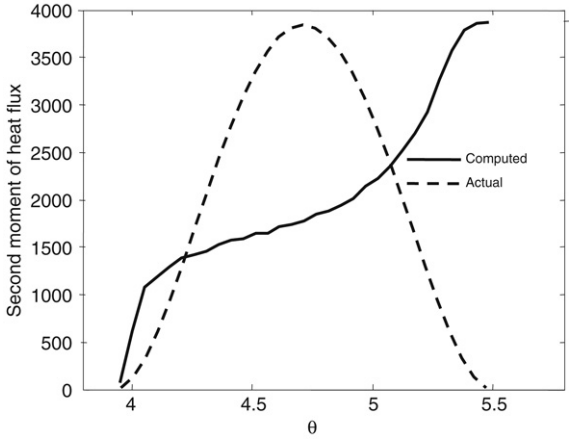

Fig. 15. The figure shows comparison of second moment of heat flux recomputed using the stochastic optimization algorithm.

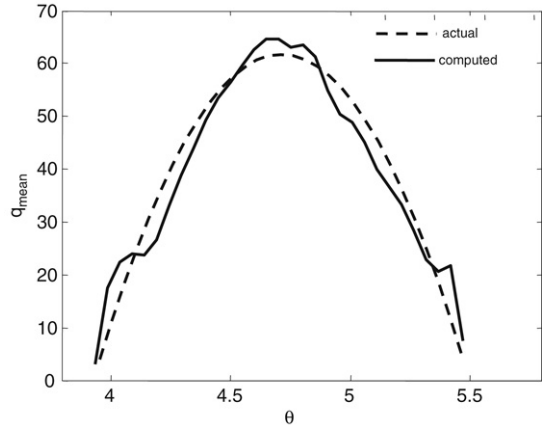

Fig. 16. The figure shows comparison of mean heat flux recomputed using the stochastic optimization algorithm using twelve sensors around the rolling specimen.

In the problem definitions, c represents concentration, $\mathbf{u}$ represents velocity, q denotes volume flux rate at the wells, and c_0 represents the initial concentration profile. The anisotropic dispersion coefficient, $\mathbf{D}$, is given by: $\mathbf{D} = \phi\{\alpha_m\mathbf{I} + \parallel \mathbf{u} \parallel [\alpha_l E(\mathbf{u}) + \alpha_t(\mathbf{I} - E(\mathbf{u}))]\}$ where $E(\mathbf{u}) = \frac{1}{\|u\|^2}\mathbf{u} \otimes \mathbf{u}$. α_m, α_l and α_t represent the molecular diffusivity, longitudinal dispersion coefficient

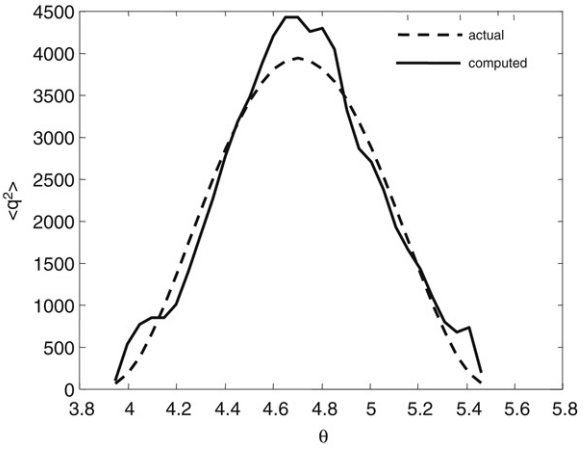

Fig. 17. The figure shows comparison of second moment of heat flux recomputed using the stochastic optimization algorithm using twelve sensors around the rolling specimen.

and transverse dispersion coefficient, respectively. μ represents the dynamic viscosity of the resident fluid, and the variable ϕ represents the porosity of the medium and is utilized to account for the fact that the effective area for flow is reduced in a porous medium. The initial concentration $c_0(x, y)$ is taken of the form: $c_0(x, y) = \mathrm{e}^{(-((x-x_c(\xi))^2+(y-y_c(\xi))^2)/2/\sigma^2)}$ for simulating measurements. The set of sensitivity equations can be derived from the set of direct Eqs. (21) and (22) as follows:

$$\begin{aligned}
&\phi\frac{\partial \mathcal{C}}{\partial t} + \nabla\cdot(\mathcal{C}\mathbf{u}) - \nabla\cdot(\mathbf{D}\nabla\mathcal{C}) = \tilde{\mathcal{C}}q, \quad \mathcal{C} \in (\mathcal{D}, \mathcal{T}, \Omega),\\
&\mathcal{C}(\mathbf{x}, 0, \boldsymbol{\xi}) = \Delta c_0(\mathbf{x}, \boldsymbol{\xi}),\\
&\mathbf{D}\nabla\mathcal{C}\cdot\mathbf{n} = 0, \quad t \in [0\ t_l].\\
&\nabla\mathbf{u}.\mathbf{n} = 0.
\end{aligned} \tag{23}$$

$\mathcal{C}$ represents the stochastic sensitivity of concentrations with respect to perturbations in $c_0(\mathbf{x}, \boldsymbol{\xi})$. In both the problems that follow, the values of some parameters are chosen as: $q = 0.04, \alpha_m = 0, \alpha_l = 0.04, \alpha_t = 0.004, \sigma = 0.1$. Stabilized SUPG finite element formulation is utilized for the solution of the direct and sensitivity problems. Numerical details for solving these equations are provided in [26,27].

Problem definition: We consider the task of reconstruction of initial concentration profile in a porous medium $t_0 = 0$ *based on concentration measurements at a different time,* $t_l = 0.1$. *The governing equations are given in* Eq. (22). *The domain is of size* $[0, 8] \times [0, 8]$. $x_c(\boldsymbol{\xi})$ *and* $y_c(\boldsymbol{\xi})$ *were chosen as Unif* [0.15 0.35] *and Unif* [0.275 0.475], *respectively.*

Physically, the source of initial concentration is not deterministic due to small perturbations or noise in its location. This is because the contaminant cannot be injected at a point and, depending on the inlet size of the injector, the location of injected fluid has some uncertainty associated with it. We model this as uniform randomness in x_c and y_c.

We generate measurement data using a 64×64 grid (with a time step of 0.0025) and choose data at points corresponding to the sensor nodes of a 32×32 grid (i.e. information at every other nodal point is extracted). This generated data is available as a PDF of concentration at sensor locations by obtaining 500 measurements using Monte-Carlo sampling technique. The task is to recompute the concentration at time, $t = 0$, and hence the concentrations at any other time using the concentration data at $t = 0.1$ (see Fig. 18).

As a first step, the auxiliary optimization problem was solved to construct the stochastic space, based on measurement concentrations. The auxiliary problem converged at two stochastic dimensions using Hermite polynomials. Though the trends between

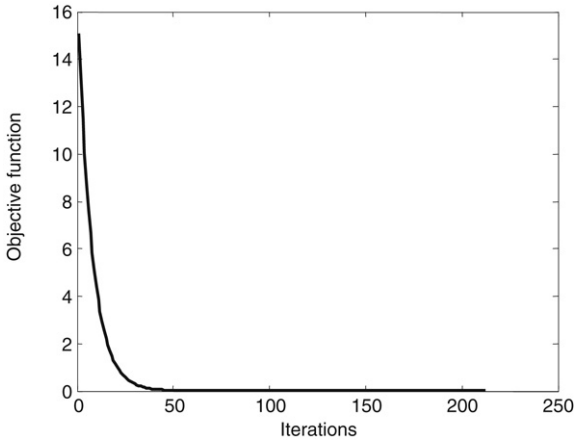

Fig. 18. The figure shows convergence of the objective function for the inverse problem of estimation of initial concentrations based on concentration measurements at a later time.

the PDF of measurement data and that constructed using the support space representation are almost the same, the values are different at certain locations in the collocation space. This can be improved only by choosing non-polynomial schemes for representation of the support space such as representation using wavelet bases. For instance, the mean and standard deviation at the location 0.25, 0.375 was 0.6004 and 0.0414, respectively while that of the reconstructed distribution was 0.6054 and 0.0480, respectively. The stochastic optimization problem is driven by this support space representation.

The inverse problem of computing the PDF's of $c_o(\mathbf{x}, \boldsymbol{\xi})$ based on concentration measurements at a later time instance, is posed as a stochastic optimization problem to minimize the objective function:

$$\begin{aligned}&\mathcal{J}(c_0(\mathbf{x}, \boldsymbol{\xi}))\\ &= \frac{1}{2}\|c(\mathbf{x}, t, \boldsymbol{\xi}; c_0(\mathbf{x}, \boldsymbol{\xi})) - c_m(\mathbf{x}, t, \boldsymbol{\xi})\|^2_{L_2(D_i\times\mathcal{T}\times\Omega)}\\ &= \frac{1}{2}\int_t\int_\Gamma\int_\Omega \{c(\mathbf{x}, t, \boldsymbol{\xi}; c_0(\mathbf{x}, \boldsymbol{\xi})) - c_m(\mathbf{x}, t, \boldsymbol{\xi})\}^2 \mathrm{d}P\mathrm{d}\mathbf{x}\,\mathrm{d}t. \end{aligned} \quad (24)$$

Here c_m represents the PDF of measured concentrations at the sensor locations. Also, note that for the current example, the integral over t becomes superfluous since we measure concentrations at just one time instance. This objective function is rewritten in the following form:

$$\begin{aligned}&\mathcal{J}(c_0(\mathbf{x}, \boldsymbol{\xi}))\\ &= \int\left[\sum_i\sum_j\int\int (c(\mathbf{x}, t, \boldsymbol{\xi}; c_0(\mathbf{x}, \boldsymbol{\xi}))\right.\\ &\quad \left. - c_m(\mathbf{x}, t, \boldsymbol{\xi}; c_0(\mathbf{x}, \boldsymbol{\xi})))^2\mathrm{d}t \times \int L_i(\boldsymbol{\xi})L_j(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}\right]\mathrm{d}x. \end{aligned} \quad (25)$$

The actual values of the concentration at initial time is computed using:

$$\tilde{c}_0(\mathbf{x}, \boldsymbol{\xi}) = \underset{c_0(\mathbf{x},\boldsymbol{\xi})}{\operatorname{argmin}}\, \mathcal{J}(c_0(\mathbf{x}, \boldsymbol{\xi})). \quad (26)$$

The stochastic optimization technique using a fifth level sparse grid collocation scheme was employed. The problem reduces to a deterministic optimization problem of dimensions 148 480. The direct and sensitivity equations are solved during the optimization routine by using a spatial and temporal discretization of 32×32 and 0.005 respectively. The mean and the standard deviation of the reconstructed concentrations at $t = 0.1$ and $t = 0.8$ are shown in Fig. 19. The results show a good comparison of the first two moments with the measurement data.

A comparison of the PDF of the computed and the measured concentrations is shown in Fig. 20. The difference is mainly due to the polynomial approximation of the support space for the measurement data. To further improve upon the accuracy, a more general basis such as wavelet functions should be used in the support space. However, it is clear that the profile of the PDF is captured.

6.4. Problem 4: Robust design of deformation processes

We adapt the methodology, stated above, to perform robust design of deformation processes. In large deformation manufacturing processes such as forging, it is necessary to ensure that the quality of the final product is good, and that the cost involved in the production of the product is minimized. This has to be done in the presence of diverse sources of uncertainties such as friction coefficient, forging velocity, initial preform shape etc. These processes are governed by constrained stochastic partial differential equations (the constraints originate due to the presence of forming dies and the P.D.E.'s are stress equilibrium equations).

$$\sigma_{ij,j} + b_i = 0; \quad u \in \mathcal{U} \quad (27)$$

where $\mathcal{U}$ is the allowable space of deformations, arising due to the presence of dies. The reader is referred to [28] for further details on the mathematical formalism to solve such equations. We now define a robust design problem in such a context, as follows:

$$\mathcal{J}(\beta) = \int (\alpha_1(\mathit{flash}(\boldsymbol{\xi})) + \alpha_2(\mathit{underfill}(\boldsymbol{\xi})))\, p(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}. \quad (28)$$

The objective is to minimize $\mathcal{J}$. "Flash" represents the material that is wasted during the forging process and "underfill" corresponds to the volume of die that is not filled. It is to be noted that the integral is taken over a probability space, defined by the uncertain sources. β represents the controllable parameters which are usually the initial shape parameters for the forging problem. In the mathematical formalism, we allow for uncertainties in β. However, this is only used in checking the convergence of the stochastic direct and sensitivity problems, whilst in the robust design problem these are taken to be deterministic. The objective function can now be written as:

$$\begin{aligned}\mathcal{J}(\beta) &= \alpha_1\int\left(\frac{1}{N_1}\sum_{i=1}^{N_1}\sum_{j=1}^{3}\left(x_j^i(\beta(\boldsymbol{\xi}), \boldsymbol{\xi}) - x_j^d\right)^2\right)p(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}\\ &\quad + \alpha_2\int\left(\frac{1}{N_2}\sum_{i=1}^{N_2}\sum_{j=1}^{3}\left(x_j^i(\beta(\boldsymbol{\xi}), \boldsymbol{\xi}) - x_j^d\right)^2\right)p(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi} \end{aligned} \quad (29)$$

where N_1 represents the number of finite element nodes which are yet to fill the die, and N_2 represents the number of nodes responsible for flash. x_j^d represents the desired value of the coordinate which is given by the shape of the die. The objective function can be written discretely as:

$$\begin{aligned}\mathcal{J} &= \frac{\alpha_1}{N_1}\sum_{i=1}^{N_1}\sum_{j=1}^{3}\int\left(\left(\sum_{k=1}^{N_c}x_j^i(\beta(\boldsymbol{\xi}^k), \boldsymbol{\xi}^k)L_k(\boldsymbol{\xi}) - x_j^d\right)^2\right)p(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}\\ &\quad + \frac{\alpha_2}{N_2}\sum_{i=1}^{N_2}\sum_{j=1}^{3}\int\left(\left(\sum_{k=1}^{N_c}x_j^i(\beta(\boldsymbol{\xi}^k), \boldsymbol{\xi}^k)L_k(\boldsymbol{\xi}) - x_j^d\right)^2\right)p(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi}. \end{aligned} \quad (30)$$

We assume that the primary design variables are the shape parameters of the workpiece that is undergoing deformation.

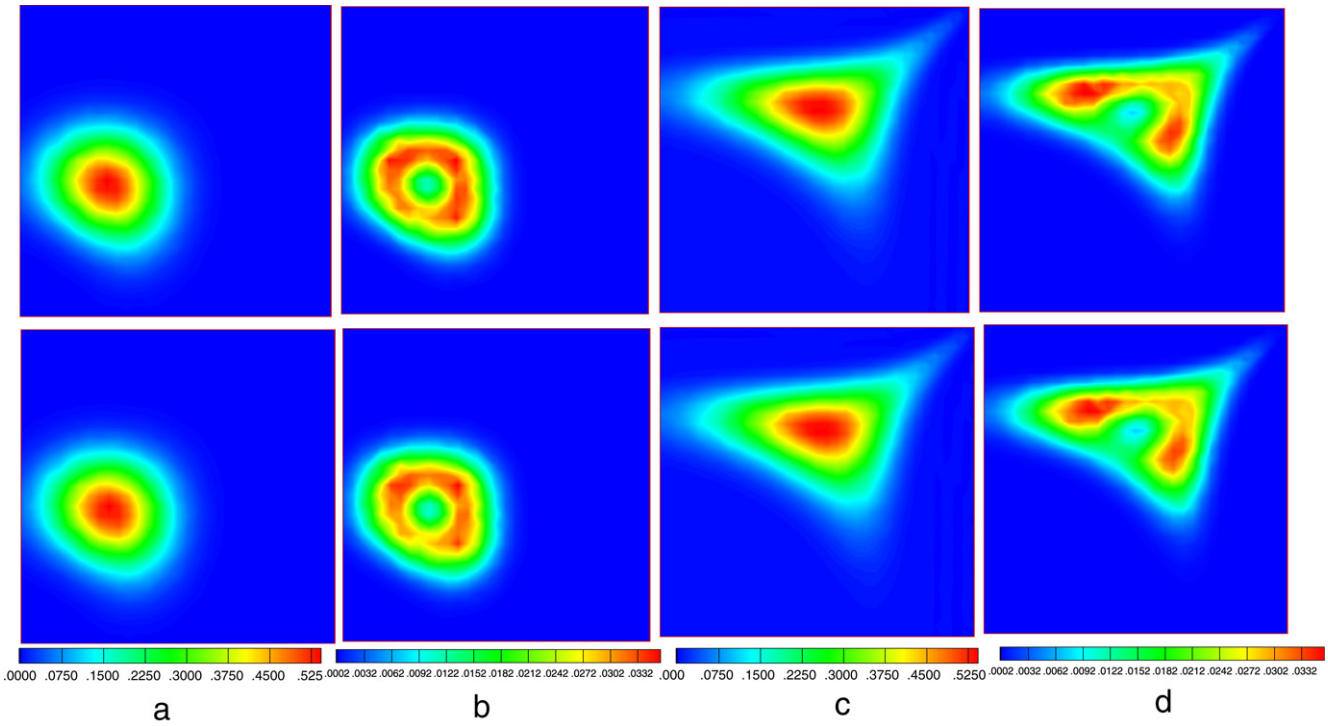

Fig. 19. Some statistics of the reconstructed concentrations (a) Comparison of the actual mean concentration and recomputed mean concentration at $t = 0.1$ (b) Comparison of actual variance of concentration with the recomputed variance at $t = 0.1$ (c) Comparison of the actual mean concentration and recomputed mean concentration at $t = 0.8$ and (d) Comparison of actual variance of concentration with the recomputed variance at $t = 0.8$.

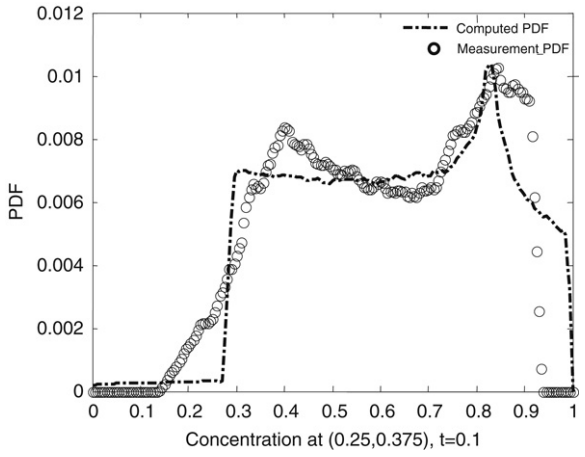

Fig. 20. The figure shows comparison of the concentration of the injected fluid at $t = 0.1$ at the location (0.25 0.375) computed using the stochastic optimization scheme. It also shows the PDF of measurement data at the same point.

These variables are designed, based on the diverse sources of uncertainties present. Taking the sensitivity of the objective function with respect to a particular realization of the shape parameter, we have:

$$\frac{\partial \mathcal{J}}{\partial \beta(\boldsymbol{\xi}^m)} = \frac{2\alpha_1}{N_1}\sum_{i=1}^{N_1}\sum_{j=1}^{3}\int I_{ij}(\boldsymbol{\xi}^m)\mathrm{d}\boldsymbol{\xi} + \frac{2\alpha_2}{N_2}\sum_{i=1}^{N_2}\sum_{j=1}^{3}\int I_{ij}(\boldsymbol{\xi}^m)\mathrm{d}\boldsymbol{\xi} \tag{31}$$

where $I_{ij}(\boldsymbol{\xi}^m) = \left(x_j^i(\beta(\boldsymbol{\xi}^m), \boldsymbol{\xi}^m)L_k(\boldsymbol{\xi}) - x_j^d\right)\frac{\partial x_j^i(\beta(\boldsymbol{\xi}^m),\boldsymbol{\xi}^m)}{\partial \beta(\boldsymbol{\xi}^m)}L_m(\boldsymbol{\xi})p(\boldsymbol{\xi})$.

It is to be noted that during the robust design procedure, we assume β's are deterministic, since the contrary does not make physical sense. However, we assume they are uncertain only to show comparison and convergence of stochastic sensitivities.

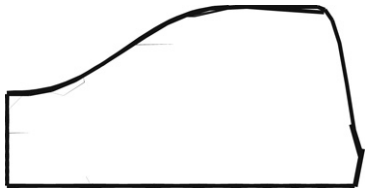

Fig. 21. The figure shows the shape of a rigid die to which the workpiece has to be deformed.

6.4.1. Problem convergence

A simple large-deformation problem with the objective of forming the workpiece to the shape shown in Fig. 21 is chosen. Details of the finite element implementation of such a process is shown in [29]. It is to be noted that a PetSc implementation of the same algorithm was performed, which resulted in significant improvement in computational efficiency, and employed here. A stochastic analysis is performed, with the main uncertainty coming from the initial shape of the workpiece and the forging velocity. The stochastic collocation technique is employed by repeatedly solving the direct problem for computing the PDF of problem solution. We first check the convergence and accuracy of the stochastic direct problem by comparing with Monte-Carlo results. The plots shown in Figs. 22 and 23 show that the solutions using increasingly higher level of stochastic collocation grids match with the Monte-Carlo solution.

Further, a comparison of the moments of stochastic shape sensitivity $\frac{\partial x_j^i(\beta(\boldsymbol{\xi}),\boldsymbol{\xi})}{\partial \beta_1(\boldsymbol{\xi})}$ is shown in Fig. 24. The results show a good convergence of the stochastic sensitivities as the level of interpolation increases. Finally, the objective function is minimized for uniform variability in the forging velocity of the process between 0.09 mm/s and 0.11 mm/s. The best design occurs for the shape parameter

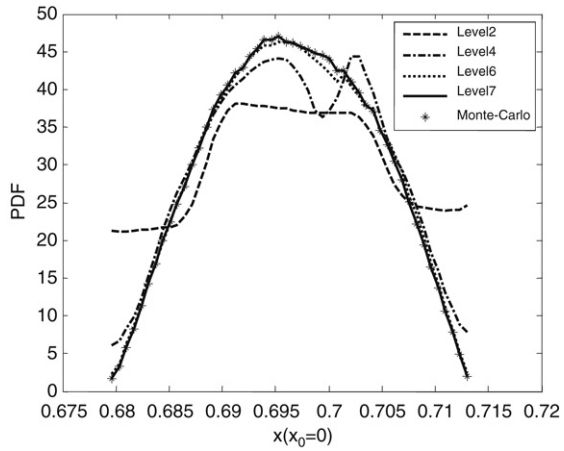

Fig. 22. The figure shows the convergence of the PDF of spatial location of a point with initial location $x_0 = 0$ at a time $t = 10$ s. The plot shows a convergence of the PDF towards that estimated using Monte-Carlo simulation.

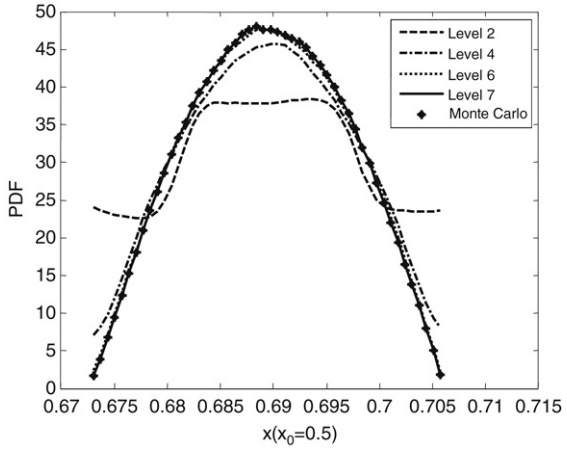

Fig. 23. The figure shows the convergence of the PDF of spatial location of a point with initial location $x_0 = 0.5$ at a time $t = 10$ s. The plot shows a convergence of the PDF towards that estimated using Monte-Carlo simulation.

Statistic	Level 2	Level 4	Level 6	Level 7	M.C
$\langle x \rangle$	-0.3020	-0.3019	-0.3016	-0.3017	-0.3017
$\langle x^2 \rangle$	0.0912	0.0912	0.0910	0.0910	0.0910
$\langle x^3 \rangle$	-0.0274	-0.0276	-0.0275	-0.0275	-0.0275
$\langle x^4 \rangle$	0.0081	0.0084	0.0083	0.0083	0.0083

Fig. 24. The table shows a comparison of the stochastic sensitivity of spatial location (occupying $x_0 = 0.5$) with respect to a perturbation in the shape parameter, β_1. Statistics of the stochastic sensitivities, computed using different levels of the collocation technique, are compared against those obtained using Finite Difference and Monte-Carlo scheme.

values of $\beta = [1.00137; 0.995292; 1.01149; 1.04842]$. The mean and extremal solutions are plotted in Fig. 25. While the objective is clearly satisfied for the mean forging velocity, it cannot be assured for all samples, since the shape parameter is assumed to be deterministic.

7. Discussion and future work

In this work, we have discussed a novel technique for solving stochastic inverse and robust design problems which only require deterministic solvers. With repeated calls to direct solvers, a full PDF of the input variability can be captured. To test the technique, we employed the SIHCP, and the results show that the developed methodology works well. We went on to employ successfully the technique on concentration reconstruction problems, as well as robust design for large deformation problems. In addition, the new technique has the ability to work well for large stochastic dimensions also, which is not the case if we employ the GPCE scheme. The generic framework that was developed in this paper can be incorporated into any stochastic design/inverse problem with a solver for computing direct and sensitivity sub-problems.

However, the most important limitation of this work is the use of polynomials to construct the support space which drives the stochastic optimization problem. It is because of this, that we are not precisely able to capture the PDFs, though the effect is not so prominent. For instance, in problems involving shocks or bifurcations, polynomials may not be the most ideal way to approximate the stochastic space, and we may need to use schemes such as wavelets. The work that is showcased here is exploratory in answering how stochastic measurements can be captured using stochastic problem parameters. In this sense, this can be considered as a problem of reverse uncertainty propagation—how does the stochastic nature of the measurements affect the computed problem parameters? This is done in a precise mathematical setting, using the stochastic collocation technique.

Acknowledgments

The computing was conducted using the resources of the Cornell Theory Center, which receives funding from Cornell University, New York State, federal agencies, and corporate partners.

References

[1] Beck JV, Blackwell B, St Clair CR. Inverse heat conduction-Ill posed problems. Wiley Interscience Publications; 1985.
[2] Alifanov OM. Inverse heat transfer problems. Berlin: Springer-Verlag; 1994.
[3] Zabaras N, Kang S. On the solution of an ill-posed inverse design solidification problem using minimization techniques in finite and infinite dimensional spaces. International Journal for Numerical Methods in Engineering 1993;36:3973–90.
[4] Narayanan VAB, Zabaras N. Stochastic inverse heat conduction using a spectral approach. International Journal for Numerical Methods in Engineering 2004;60:1569–93.
[5] Wang J, Zabaras N. A Bayesian inference approach to the stochastic inverse heat conduction problem. International Journal of Heat and Mass Transfer 2004;47:3927–41.
[6] Xiu D, Hesthaven JS. High order collocation methods for the differential equation with random inputs. SIAM Journal of Scientific Computing 2005;27:1118–39.

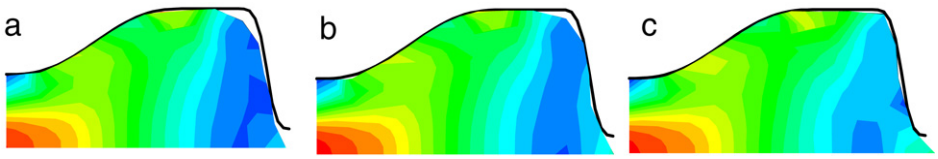

Fig. 25. The figure shows (a) the shape of the final workpiece at the left tail of the forging velocity (0.09 mm/s), (b) the shape of the final workpiece at the right tail of the forging velocity (0.11 mm/s) and (c) the shape of the final workpiece at the mean forging velocity (0.1 mm/s).

[7] Ganapathysubramaniam B, Zabaras N. Sparse grid collocation schemes for stochastic natural convection problems. Journal of Computational Physics 2007;225:652–85.
[8] Loève M. Probability theory. 4th edition. Berlin: Springer-Verlag; 1977.
[9] Nelder JA, Mead R. A simplex method for function minimization. Computer Journal 1965;7:308–13.
[10] Ghanem RG, Spanos PD. Stochastic finite elements: A spectral approach. New York: Springer Verlag; 1991.
[11] Xiu D, Karniadakis GE. Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. Computer Methods in Applied Mechanics and Engineering 2002;191:4927–48.
[12] Xiu D, Karniadakis GE. Modeling uncertainty in flow simulations via generalized polynomial chaos. Journal of Computational Physics 2003;187:137–67.
[13] Babuska I, Nobile F, Tempone R. A stochastic collocation method for elliptic partial differential equations with random input data. ICES report. p. 05-47; 2005.
[14] Xiu D. Efficient collocational approach for parametric uncertainty analysis. Communications in Computational Physics 2007;2:293–309.
[15] Klimke A. Uncertainty modeling using fuzzy arithmetic and sparse grids. Ph.D. Thesis. Aachen: Universitt Stuttgart, Shaker Verlag, 2006.
[16] Klimke A, Wohlmuth B. Algorithm 847: Spinterp: Piecewise multilinear hierarchical sparse grid interpolation in MATLAB. ACM Transactions on Mathematical Software 2005;31.
[17] Klimke A. Sparse grid interpolation toolbox users guide. IANS report 2006/001. University of Stuttgart; 2006.
[18] Gerstner T, Griebel M. Numerical integration using sparse grids. Numerical Algorithms 1998;18:209–32.
[19] Barthelmann V, Novak E, Ritter K. High dimensional polynomial interpolation on sparse grids. Advances in Computational Mathematics 2000;12:273–88.
[20] Ganapathysubramaniam B, Zabaras N. Modelling diffusion in random heterogeneous media: Data-driven models, stochastic collocation and the variational multi-scale method. Journal of Computational Physics 2007;226:326–53.
[21] Hsu P-T. A three-dimensional inverse problem of estimating the surface thermal behavior of the working roll in rolling process. Journal of Manufacturing Science and Engineering 2000;122:76–82.
[22] Atmadja J, Bagtzoglou AC. Pollution source identification in heterogeneous porous media. Water Resources Research 2001;37:2113–25.
[23] Gorelick SM, Evans BE, Remson I. Identifying sources of groundwater pollution: An optimization approach. Water Resources Research 1983;19:779–90.
[24] Wagner BJ. Simultaneously parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modeling. Journal of Hydrology 1992;135:275–303.
[25] Wang J, Zabaras N. A Markov random field model to contamination source identification in porous media flow. International Journal of Heat and Mass Transfer 2006;49:939–50.
[26] Coutinho ALGA, Alves JLD. Parallel finite element simulation of miscible displacement in porous media. SPE Journal 1996;1:487–500.
[27] Coutinho ALGA, et al. Stabilized methods and post-processing techniques for miscible displacements. Computer Methods in Applied Mechanics and Engineering 2004;193:1421–36.
[28] Zabaras N, Srikanth A. An object-oriented programming approach to the Lagrangian FEM analysis of large inelastic deformations and metal forming processes. International Journal for Numerical Methods in Engineering 1999;45:399–445.
[29] Acharjee S, Zabaras N. The continuum sensitivity method for the computational design of three-dimensional deformation processes. Computer Methods in Applied Mechanics and Engineering 2006;195:6822–42.