Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 227 (2008) 4697–4735
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

A scalable framework for the solution of stochastic inverse problems using a sparse grid collocation approach

N. Zabaras *, B. Ganapathysubramanian

Materials Process Design and Control Laboratory, Sibley School of Mechanical and Aerospace Engineering, 101 Frank H.T. Rhodes Hall, Cornell University, Ithaca, NY 14853-3801, USA

Received 26 September 2007; received in revised form 27 November 2007; accepted 14 January 2008
Available online 26 January 2008

Abstract

Experimental evidence suggests that the dynamics of many physical phenomena are significantly affected by the underlying uncertainties associated with variations in properties and fluctuations in operating conditions. Recent developments in stochastic analysis have opened the possibility of realistic modeling of such systems in the presence of multiple sources of uncertainties. These advances raise the possibility of solving the corresponding stochastic inverse problem: the problem of designing/estimating the evolution of a system in the presence of multiple sources of uncertainty given limited information.

A scalable, parallel methodology for stochastic inverse/design problems is developed in this article. The representation of the underlying uncertainties and the resultant stochastic dependant variables is performed using a sparse grid collocation methodology. A novel stochastic sensitivity method is introduced based on multiple solutions to deterministic sensitivity problems. The stochastic inverse/design problem is transformed to a deterministic optimization problem in a larger-dimensional space that is subsequently solved using deterministic optimization algorithms. The design framework relies entirely on deterministic direct and sensitivity analysis of the continuum systems, thereby significantly enhancing the range of applicability of the framework for the design in the presence of uncertainty of many other systems usually analyzed with legacy codes. Various illustrative examples with multiple sources of uncertainty including inverse heat conduction problems in random heterogeneous media are provided to showcase the developed framework.

Keywords: Stochastic partial differential equations; Inverse problems; Stochastic optimization; Collocation methods; Sparse grids; Scalable algorithms

1. Introduction

With the rapid advances in computational power and easier access to high-performance computing platforms, it has now become possible to computationally investigate realistic multiscale, multiphysics

* Corresponding author. Tel.: +1 607 255 9104; fax: +1 607 255 1222.
E-mail address: zabaras@cornell.edu (N. Zabaras).
URL: http://mpdc.mae.cornell.edu/ (N. Zabaras).

0021-9991/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2008.01.019

phenomena. As a direct consequence of this computational ability, there has been growing awareness that the tightly coupled and highly nonlinear systems (that such problems are composed of) are affected to a significant extent by the inherent (usually unresolvable) uncertainties in material properties, fluctuations in operating conditions and system characteristics. The scale at which these phenomena occur can range from the micro-scale (MEMS devices), macro-scale (devices/components made from polycrystalline or functionally graded materials) to the geological scale (geothermal energy systems).

It is clearly necessary to incorporate the effects of uncertainty in property variation and operating conditions for the design of components and devices that perform satisfactorily in a variety of conditions. The first step towards design of components and devices in the presence of uncertainty is a framework for the analysis of the direct stochastic problem, i.e. the ability to quantify, analyze and predict the effects of multiple sources of uncertainty on the performance of the system.

1.1. Scalable direct stochastic analysis

The presence of uncertainties can be modeled in the system through reformulation of the governing equations as stochastic partial differential equations (SPDEs). A recent approach to model uncertainty is based on the spectral stochastic finite element method (SSFEM) [1]. This method has been applied with great success to investigate numerous physical phenomena [2–5]. Error bounds and convergence studies [6–8] have shown that these methods exhibit fast convergence rates with increasing orders of expansions. Though the SSFEM method has been used successfully in a variety of scenarios, it has a few drawbacks like the 'curse-of-dimensionality' and the intense programming effort required to build the framework [9–11]. The SSFEM method cannot be easily applied to problems involving high stochastic dimension (there has been some recent progress in this regard, see [12]).

To solve problems in high-dimensional stochastic spaces (in a scalable way) and to allow non-smooth variations of the solution in the random space, there have been recent efforts to couple the fast convergence of the Galerkin methods with the decoupled nature of Monte-Carlo sampling [13,14]. The Smolyak algorithm has been used recently to build sparse grid interpolants in high-dimensional space [10,11,15]. Using this method, interpolation schemes (for the solution) can be constructed with orders of magnitude reduction in the number of sampled points to give the same level of approximation (up to a logarithmic factor) as interpolation on a uniform grid. In [9], we extended this methodology to adaptively sample more important dimensions, resulting in further computational gains. The sparse grid collocation and cubature schemes have been well studied and utilized in different fields [16–18]. The sparse grid collocation strategy provides a seamless way to scalably incorporate the effects of multiple sources of uncertainty in an embarrassingly parallel way. Furthermore, it extensively utilizes multiple solutions of deterministic problems to construct the stochastic solution. This allows the use of pre-existing deterministic legacy codes in a stochastic setting.

1.2. Stochastic inverse/design problems: state-of-the-art

In a typical inverse problem, one is interested in identifying the initial, boundary and/or material properties given sensor measurements or desired values of the dependant variable inside the domain. For example, the inverse heat conduction problem is of interest in a wide range of scientific and engineering areas including manufacturing process control, metallurgy, chemical, aerospace and nuclear engineering, and food science's among others. There are a number of solution techniques to solve the deterministic inverse/design problem [19,20]. Deterministic inverse techniques (usually based on exact matching or least-squares optimization) lead to point estimates of unknowns without rigorously considering system uncertainties and without providing quantification of the uncertainty in the inverse solution.

The Bayesian inference approach provides a robust means of taking system variabilities and parameter fluctuations into account. This framework formulates a complete probabilistic description of the unknowns and uncertainties given measurement data [21]. The Bayesian approach incorporates the known information regarding the unknown parameters and system uncertainties into a prior distribution model that is then combined with the likelihood to formulate the posterior probability density function (PPDF). This methodology

has been used with great success to solve a variety of problems [22–25]. Though Bayesian based inference is a powerful approach to estimate parameters in the presence of noisy measurements and uncertainties in the system, there are a few issues that have motivated the development of alternate means of solving stochastic inverse/design problems. The higher-order statistics of the unknown are highly-dependent on the form of the prior distribution chosen. Furthermore, the inputs to this framework are a finite number of deterministic sensor measurements, i.e. this methodology provides no means of incorporating statistics and/or PDF of the measured quantities. Though possible, this framework has not been applied to design problems with multiple sources of uncertainty.

An alternate approach that tries to resolve the issues raised above was proposed in [26]. In this method, a spectral stochastic formulation was utilized to represent the input uncertainties. This framework was applied to the stochastic inverse heat conduction problem. Since this framework utilized a GPCE based approach, it required extensive software developments to convert the validated deterministic direct, sensitivity and adjoint problems to their corresponding stochastic counterparts. The input measurement data in this framework was assumed to be given in the form of an unphysical spectral representation. Furthermore, the coupled nature of the solution strategy precluded the extension of this methodology to include multiple sources of uncertainty. In the current work, we extend our previous work in [26] into a scalable, decoupled parallel framework for solving stochastic inverse problems.

The keys properties that a widely usable stochastic optimization framework must possess are:

- The ability to seamlessly utilize available, validated deterministic simulators. This will significantly enhance the applicability of the stochastic optimization/inverse framework to areas of analysis that utilize complex legacy codes. Such complex legacy codes cannot be rewritten to include stochastic analysis without a great deal of effort.
- The framework must be highly scalable to provide the ability to solve complex problems in a reasonable amount of time.
- The framework must be able to seamlessly incorporate the effects of multiple sources of uncertainties.

In the present work, we develop a stochastic optimization framework where the uncertainty representation is based on a sparse grid collocation approach. Utilizing a sparse grid collocation strategy guarantees scalability, the ability to seamlessly incorporate multiple sources of uncertainty and more importantly relies entirely on *multiple calls to deterministic simulators* [9]. The stochastic inverse/design problem is posed as a stochastic optimization problem. Using a sparse grid representation of the design variable converts the stochastic optimization problem into a deterministic optimization problem in a larger-dimensional space. This deterministic optimization problem is subsequently solved using standard gradient based optimization algorithms. A novel stochastic sensitivity computation method to compute the sensitivity of the dependant stochastic variables with respect to the design stochastic variables is formulated. This formulation arises naturally from posing the problem in the sparse grid framework and involves multiple calls to deterministic sensitivity problems. To the best knowledge of the authors, this is the first framework that: (1) addresses inverse/design problems in the presence of multiple sources of uncertainty, (2) relies completely on deterministic simulators, (3) requires no a priori assumption on the design variables (as compared to the Bayesian framework), and (4) is based on off-the-shelf optimization algorithms. A subtle qualification of (3): no prior model representation needs to be assumed as is usually done in the Bayesian framework. However, the conventional sparse grid collocation strategy does implicitly assume that the function is piecewise smooth. Nevertheless, by suitably refining the depth of interpolation, arbitrary functions can be represented. Furthermore, adaptive sparse grid strategies can also remove this implicit assumption of smoothness completely.

The outline of the paper is as follows: In the next section, the direct problem of interest is posed, all the necessary variables and function spaces are defined and some preliminary issues clarified. In Section 3, the definition of the stochastic inverse problem is given and the scalable framework is presented. This is followed by Section 4, where numerical examples showcasing the proposed framework are presented. We finally conclude with a summary of this work and a discussion of future avenues of research in Section 5.

2. The stochastic direct problem: definition, solution strategy and physical interpretation

The developments detailed in the present paper are applicable to a general class of stochastic inverse problems. For the sake of clarity of presentation, we will henceforth limit our discussion to a particular class of problems (namely, the stochastic inverse heat conduction problem). Before defining the stochastic inverse problem of interest, the stochastic direct problem, that describes the physical phenomena, is represented first.

2.1. Definition of the stochastic direct problem

Denote by $\mathcal{D}$, a bounded connected region in $\mathbb{R}^{nsd}$, where $nsd = 1, 2, 3$ is the number of space dimensions over which $\mathcal{D}$ is defined. Denote the boundary of $\mathcal{D}$ as $\partial\mathcal{D}$. The boundary $\partial\mathcal{D}$ is divided into disjoint sub-boundaries $\partial\mathcal{D}_\theta$ and $\partial\mathcal{D}_h$ $(\partial\mathcal{D}_\theta \bigcup \partial\mathcal{D}_h = \partial\mathcal{D}, \partial\mathcal{D}_\theta \bigcap \partial\mathcal{D}_h = \emptyset)$ such that Dirichlet boundary conditions are applied on $\partial\mathcal{D}_\theta$ and Neumann boundary conditions are applied in $\partial\mathcal{D}_h$ (see Fig. 1). The material property that affects the thermal evolution in the domain is the thermal diffusivity, $\alpha = \frac{k}{\rho C_p}$, where k is the thermal conductivity, ρ is the density of the material and C_p is the specific heat of the material.

Assume that there is some uncertainty in the application of the thermal boundary conditions. Without loss of generality, we will assume that the Neumann boundary conditions are enforced with some uncertainty. i.e. the flux boundary conditions are stochastic while the Dirichlet boundary conditions are deterministic. The thermal boundary conditions can be understood to be realizations from a space of events. Let us denote this space as Ω_q. We can associate a σ-algebra $\mathcal{F}_q$ and a probability measure $\mathcal{P}_q : \Omega_q \to [0, 1]$ to construct a probability space $(\Omega_q, \mathcal{F}_q, \mathcal{P}_q)$ that describes the variability in the stochastic heat flux q [27]. Analogously, the thermal properties in the domain $\mathcal{D}$ (here the diffusivity α) are also assumed to be stochastic. Consequently, we can define a probability space $(\Omega_\alpha, \mathcal{F}_\alpha, \mathcal{P}_\alpha)$ that describes the variability in the diffusivity.

The temperature in the domain depends on the uncertainty in the boundary conditions and material properties as well as temporal and spatial coordinates of measurement. We denote this dependance by the function $\theta(\boldsymbol{x}, t, \omega_q, \omega_\alpha) : \mathcal{D} \times \mathcal{T} \times \Omega_q \times \Omega_\alpha \to \mathbb{R}$, where $\boldsymbol{x} \in \mathcal{D}$, $t \in \mathcal{T} \equiv [0, t_{\max}]$, $\omega_q \in \Omega_q$, $\omega_\alpha \in \Omega_\alpha$. Note that this stochastic process, $\theta(\cdot)$ lies in the *product space of the input uncertainties* (i.e. $\Omega_q \times \Omega_\alpha$). In fact, most dependant and independent variables one encounters in stochastic analysis are usually random processes [7,28–30] that are defined in product spaces. These variables (for instance $q(\boldsymbol{x}, \omega_q)$) usually have one structure in the stochastic space $q(\omega_q, \cdot)$ and another in the physical space $q(\cdot, \boldsymbol{x})$. The numerical analysis/approximation [7,28] of such functions can be performed by defining the appropriate tensor product spaces. The interested reader is referred to [7,28] for insightful discussions on the definitions of these product spaces. Following [7,28], we define appropriate function spaces that encode variations of the function in the physical domain $\mathcal{D} \times \mathcal{T}$ and in the stochastic space $\Omega_q \times \Omega_\alpha$.

Denote the Banach space $L^p(\mathcal{D})$ as the space of functions that are pth power integrable, i.e. $f \in L^p(\mathcal{D})$ if $\int_{\mathcal{D}} |f|^p d\boldsymbol{x} < \infty$. Let $W^{s,p}(\mathcal{D})$ denote the Sobolev space of functions having generalized derivatives up to order s in the space $L^p(\mathcal{D})$. Whenever $p = 2$, we use H^s instead of $W^{s,2}$. For a Sobolev space H^1, define its stochastic counterpart, $\mathcal{H}^1$, as the product space $H^1 \otimes L^2(\Omega_q \times \Omega_\alpha)$. Denote the space of functions defined in $\partial\mathcal{D}_q$ with

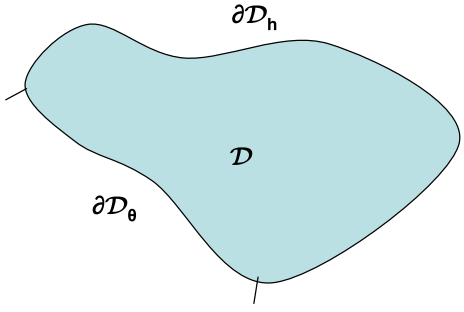

Fig. 1. Schematic of the direct problem.

finite kth norm (k any positive integer) $\|q\|^k_{\partial\mathcal{D}_q} = (\int_{\partial\mathcal{D}_q} |q(\boldsymbol{x})|^k \mathrm{d}\boldsymbol{x})^{1/k}$ as $H(\partial\mathcal{D}_q)$. Define the stochastic Banach space $L^k_{\mathcal{P}_q}(\Omega_q : H(\partial\mathcal{D}_q))$ as [28]

$$L^k_{\mathcal{P}_q}(\Omega_q : H(\partial\mathcal{D}_q)) := \left\{ q : \Omega_q \to H(\partial\mathcal{D}_q) | q \text{ measurable}, \int_{\Omega_q} \|q(\omega_q)\|^k_{\partial\mathcal{D}_q} d\mathcal{P}_q(\omega_q) < \infty \right\}. \tag{1}$$

Having defined the necessary function spaces, let us now turn to the definition of the stochastic direct problem. Consider the tensor product Hilbert space, $H_\theta \equiv H^1(\mathcal{D}, L^2(\mathcal{T})) \otimes L^2(\Omega_q \times \Omega_\alpha)$. The direct problem can be posed in this space as follows: Find a stochastic function $\theta \equiv \theta(\boldsymbol{x}, t, \omega_q, \omega_\alpha) : \mathcal{D} \times \mathcal{T} \times \Omega_q \times \Omega_\alpha \to \mathbb{R}$, such that P-a.e. (almost everywhere) in $\Omega_\alpha \times \Omega_q$, the following equation holds:

$$\frac{\partial\theta(\boldsymbol{x}, t, \omega_q, \omega_\alpha)}{\partial t} = \nabla[\alpha(\boldsymbol{x}, \omega_\alpha)\nabla\theta(\boldsymbol{x}, t, \omega_q, \omega_\alpha)], \quad \boldsymbol{x} \in \mathcal{D}, \tag{2}$$

$$\frac{\partial\theta(\boldsymbol{x}, t, \omega_q, \omega_\alpha)}{\partial \boldsymbol{n}} = q(\boldsymbol{x}, t, \omega_q), \quad \boldsymbol{x} \in \partial\mathcal{D}_h, \tag{3}$$

$$\theta(\boldsymbol{x}, t, \omega_q, \omega_\alpha) = h(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \partial\mathcal{D}_\theta, \tag{4}$$

$$\theta(\boldsymbol{x}, 0, \omega_q, \omega_\alpha) = 0, \quad \boldsymbol{x} \in \mathcal{D}. \tag{5}$$

Note that the Dirichlet conditions on $\partial\mathcal{D}_\theta$ as well as the initial conditions are deterministic (without loss of generality). For the sake of brevity, we shall denote the above set of equations as

$$\mathcal{B}(\theta : \boldsymbol{x}, t, \omega_q, \omega_\alpha) = 0. \tag{6}$$

We make the following assumptions regarding the stochastic processes q and α [7,28]:

Assumption 1. q is a bounded stochastic function.

Assumption 2. α is positive and uniformly coercive:

$$\exists \alpha_{\min}, \alpha_{\max} \in (0, \infty) \ni \mathcal{P}_\alpha(\omega_\alpha \in \Omega_\alpha : \alpha(\omega_\alpha, \boldsymbol{x}) \in [\alpha_{\min}, \alpha_{\max}] \ \forall \boldsymbol{x} \in \mathcal{D}) = 1.$$

Assumption 3. Following [28], we assume that the stochastic boundary condition q satisfies some regularity assumption, i.e. for a given $k > 2$ (defined a priori) $q \in L^k_{\mathcal{P}_q}(\Omega_q : H(\partial\mathcal{D}_q))$.

Theorem 2.1a. *There is a unique weak solution to the stochastic parabolic problem.*

Proof. With Assumptions 1 and 2 on the elliptic operator and the regularity constraints on the boundary conditions, coupled with the definition of the corresponding function spaces, this is a direct application of the existence and uniqueness results for parabolic equations given in [31, Theorem 1] and [32, Theorem 2]. □

2.1.1. Steady state problem: the stochastic elliptic operator

The steady-state behavior of the system defined by the stochastic parabolic equation defined on the physical domain $\mathcal{D}$ is the corresponding stochastic elliptic equation. The evolution equation for temperature is simplified into the following:

$$\nabla[\alpha(\boldsymbol{x}, \omega_\alpha)\nabla\theta(\boldsymbol{x}, \omega_q, \omega_\alpha)] = 0, \quad \boldsymbol{x} \in \mathcal{D}, \tag{7}$$

$$\frac{\partial\theta(\boldsymbol{x}, \omega_q, \omega_\alpha)}{\partial \boldsymbol{n}} = q(\boldsymbol{x}, \omega_q), \quad \boldsymbol{x} \in \partial\mathcal{D}_h, \tag{8}$$

$$\theta(\boldsymbol{x}, \omega_q, \omega_\alpha) = h(\boldsymbol{x}), \quad \boldsymbol{x} \in \partial\mathcal{D}_\theta. \tag{9}$$

Theorem 2.1b. *There is a unique weak solution to the stochastic elliptic problem.*

Proof. The existence and uniqueness follow from a straightforward application of the Lax–Milgram lemma [7]. □

2.2. Solution techniques for the direct problem: collocation based approximations of the stochastic spaces

The solution methodology of the above set of coupled differential equations is to first reduce the complexity of the problem by reducing the probability space into a finite-dimensional space. In some cases the random fields $q(\boldsymbol{x}, t, \omega_q)$ and $\alpha(\boldsymbol{x}, \omega_\alpha)$ can be represented/described by finite length random vectors $[\xi_q^1, \ldots, \xi_q^n] : \Omega_q \to \mathbb{R}^n$ and $[\xi_\alpha^1, \ldots, \xi_\alpha^m] : \Omega_\alpha \to \mathbb{R}^m$, respectively. In other cases the random fields can have a spatial correlation. There is rich literature on techniques to extract/fit correlations for these random fields from input experimental/numerical data. This is an area of intense ongoing research [33,34]. Using the 'finite-dimensional noise assumption' [7], the random process is decomposed into a finite set of random variables [9]. Upon decomposition and characterization of the random inputs, $q(\omega_q)$ and $\alpha(\omega_\alpha)$ into these random variables, $\xi_q^i(\omega_q)$, $i = 1, \ldots, n$, and $\xi_\alpha^i(\omega_\alpha)$, $i = 1, \ldots, m$, respectively, the solution to the system of Eq. (6) can be described by this set of random variables. The dependance of temperature can now be written as

$$\theta(\boldsymbol{x}, t, \omega_q, \omega_\alpha) = \theta(\boldsymbol{x}, t, \xi_q^1(\omega_q), \ldots, \xi_q^n(\omega_q), \xi_\alpha^1(\omega_\alpha), \ldots, \xi_\alpha^m(\omega_\alpha)) = \theta(\boldsymbol{x}, t, \boldsymbol{\xi}_q, \boldsymbol{\xi}_\alpha), \tag{10}$$

where $\boldsymbol{\xi}_q = \{\xi_q^1(\omega_q), \xi_q^2(\omega_q), \ldots, \xi_q^n(\omega_q)\}$ and $\boldsymbol{\xi}_\alpha = \{\xi_\alpha^1(\omega_\alpha), \xi_\alpha^2(\omega_\alpha), \ldots, \xi_\alpha^m(\omega_\alpha)\}$. Eq. (6) can be rewritten in the following form:

$$\mathcal{B}(\theta : \boldsymbol{x}, t, \boldsymbol{\xi}_q, \boldsymbol{\xi}_\alpha) = 0. \tag{11}$$

We denote the product space of these random variables as $\Omega_\theta \equiv \Omega_q \times \Omega_\alpha$.

Remark 2.1. It has to be emphasized that the solution of the above set of SPDEs requires an analytic (or numerical) representation of the input uncertainties (namely, the stochastic heat flux, and the thermal diffusivity). Our recent work [33] outlines a technique to construct a model for the thermal diffusivity variability. We assume that a model for the heat flux is similarly available (say, using a KL expansion) [9].

The dimensionality, d_θ of the stochastic space Ω_θ can become very large. Usually, the dimensionality, d_α of representation of the property/microstructural variability, Ω_α is $d_\alpha \sim 9$–15, while the dimensionality, d_q of representation of the heat flux is $d_q \sim 2$–5. Hence the dimensionality, $d_\theta = d_\alpha + d_q$ of the stochastic space is at least of the order of 10. Any coupled strategy would make the solution of the direct problem (let alone the inverse problem) a computationally challenging task. We therefore utilize a collocation based strategy for representing and solving the direct SPDEs [9].

In the collocation based approach, the stochastic solution to the SPDE given by Eq. (11) is converted into a finite number, n of deterministic PDEs. That is, the stochastic space Ω_θ is approximated as a multi-dimensional interpolant. This interpolant is constructed using a finite set of function evaluations [9,35]. An interpolant-based approximation is constructed by sampling the two stochastic spaces, Ω_q and Ω_α using n_q and n_α points, respectively. The collocation approach involves the solution of $n_q \times n_\alpha$ PDEs of the following form:

$$B(\theta : \boldsymbol{x}, t, \boldsymbol{\xi}_q^i, \boldsymbol{\xi}_\alpha^j) = 0, \quad i = 1, \ldots, n_q, \;\; j = 1, \ldots, n_\alpha. \tag{12}$$

The stochastic solution is then given in terms of these $n_q \times n_\alpha$ deterministic solutions, $\theta(\boldsymbol{x}, t, \boldsymbol{\xi}_q^i, \boldsymbol{\xi}_\alpha^j)$ as

$$\theta(\boldsymbol{x}, t, \boldsymbol{\xi}_q, \boldsymbol{\xi}_\alpha) = \sum_{i=1}^{n_q} \sum_{j=1}^{n_\alpha} \theta(\boldsymbol{x}, t, \boldsymbol{\xi}_q^i, \boldsymbol{\xi}_\alpha^j) L_i(\boldsymbol{\xi}_q) L_j(\boldsymbol{\xi}_\alpha), \tag{13}$$

where L_i are corresponding (sparse grid) interpolating functions [9].

2.2.1. Postprocessing formulae: computing moments

The statistics of the random solution can be obtained by

$$\langle \theta^p(\boldsymbol{x}, t) \rangle = \sum_{i=1}^{n_q} \sum_{j=1}^{n_\alpha} \left[\theta^p(\boldsymbol{x}, t, \boldsymbol{\xi}_q^i, \boldsymbol{\xi}_\alpha^j) \int_{\Omega_q} L_i(\boldsymbol{\xi}_q) \rho(\boldsymbol{\xi}_q) \mathrm{d}\boldsymbol{\xi}_q \int_{\Omega_\alpha} L_j(\boldsymbol{\xi}_\alpha) \rho(\boldsymbol{\xi}_\alpha) \mathrm{d}\boldsymbol{\xi}_\alpha \right], \tag{14}$$

where p is any positive integer ($p = 1$ gives the mean, $p = 2$ gives the second moment, etc.). The integrands on the right-hand side of the above equation are independent of the function values and depend only on the par-

ticular interpolation function, L_i. Hence these integrals can be computed a priori to enhance the computational efficiency of the algorithm. Denoting these integrals as $w_\alpha^j = \int_{\Omega_\alpha} L_j(\xi_\alpha)\rho(\xi_\alpha)\mathrm{d}\xi_\alpha$ and $w_q^i = \int_{\Omega_q} L_i(\xi_q)\rho(\xi_q)\mathrm{d}\xi_q$, the statistics can be computed as

$$\langle \theta^p(\boldsymbol{x},t)\rangle = \sum_{i=1}^{n_q}\sum_{j=1}^{n_\alpha}[\theta^p(x,t,\xi_q^i,\xi_\alpha^j)w_q^i w_\alpha^j]. \tag{15}$$

2.3. Physical interpretation of the stochastic solution: a prelude to the inverse problem

The previous subsection discusses the numerical methodology to compute the stochastic solution $\theta(\cdot,\cdot,\xi_q,\xi_\alpha) \in \Omega_q \times \Omega_\alpha$. *Given stochastic representations of the input uncertainties* $q = \mathcal{G}_1(\xi_q)$ and $\alpha = \mathcal{G}_2(\xi_\alpha)$, the sparse grid collocation allows one to construct a stochastic representation of the dependant variable $\theta = \mathcal{G}(\xi_\theta)$. This is pictorially represented in Fig. 2 for a simple case where α is a constant and the heat flux can be represented in a two-dimensional stochastic space. The stochastic solution procedure constructs the resultant stochastic temperature. This stochastic temperature is consequently defined in a two-dimensional stochastic space.

We emphasize that the stochastic representation, $\theta = \mathcal{G}(\xi_\theta)$, of any random process is a purely abstract scheme. From a physically realizable/observable perspective, only *operations* on the stochastic representation, $\theta = \mathcal{G}(\xi_\theta)$, make sense (for instance, realizations of events, moments or the probability distribution function). It is clear that these operations are *reduced* (or averaged) representations of the abstract random process. This observability/measurability argument strongly points to the necessity of posing the inverse stochastic problem in a way that the measured or desired dependant variables are given in the form of moments or PDF.

In this work, we argue that the design variables should also be computed in the form of moments or PDF. The reason for this arises from the analysis of the existence and uniqueness of the kth order moment of the solution of Eq. (6), $\langle \theta^k(\boldsymbol{x},t)\rangle = \int_{\Omega_q}\int_{\Omega_\alpha}\theta^k(\boldsymbol{x})\mathrm{d}\mathcal{P}_\alpha(\omega_\alpha)\mathrm{d}\mathcal{P}_q(\omega_q)$, given the kth moment of the input stochastic heat flux q. In [28], an investigation is provided of the existence/uniqueness of the kth moment of the dependant variable in a stochastic elliptic equation (Eq. (7)) in the context of stochastic loading and boundary conditions. An elegant description/construction of an equation of the kth moment of the dependant variable is given based on tensor product of appropriate function spaces. It is shown that given the kth moment of the input stochastic variable, a unique kth moment for the corresponding dependant variable exists. It is simple to extend these developments to the present case where the coefficients in the elliptic operator are also random (in the present work, Assumption 2 is equivalent to Eq. (7) in [28]). We state the corresponding theoretical results without going into the details of the proofs. The interested reader is referred to [28] for an exhaustive discussion of the same.

Theorem 2.2. *If q satisfies Assumption 3, the kth order moment of θ (that solves Eq. (7)) exists.*

Proof. Follows from Theorem 3.1 in [28]. □

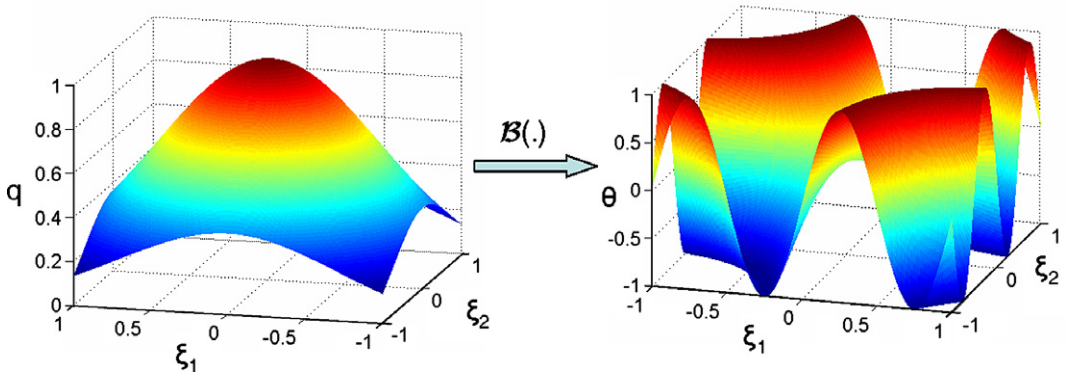

Fig. 2. The stochastic direct problem, $\mathcal{B}$, converting a random process q to a random process θ.

Theorem 2.3. *Given the kth order moments of the stochastic function q,* $\langle q^k(\boldsymbol{x},t)\rangle = \int_{\Omega_q} q^k(\boldsymbol{x},t,\omega_q)\mathrm{d}\mathcal{P}_q(\omega_q)$, *a unique solution of* $\langle\theta^k(\boldsymbol{x},t)\langle = \int_{\Omega_q}\int_{\Omega_\alpha}\theta^k(\boldsymbol{x},t,\omega_q,\omega_\alpha)\mathrm{d}\mathcal{P}_\alpha(\omega_\alpha)\mathrm{d}\mathcal{P}_q(\omega_q)$ *exists, where* θ *solves Eq. (7).*

Proof. Follows from Theorem 3.2 in [28] (which is an application of the Lax–Milgram theorem to the bilinear operator that describes the kth order moment equation). □

Remark 2.2. It has to be noted that the uniqueness/existence proofs stated above are for the stochastic elliptic problem. The authors contend that it is possible to apply the theory of [28] coupled with the uniqueness/existence proofs of the stochastic parabolic equation (Theorem 2.1a) [32] to construct uniqueness/existence proofs for the kth moments of the dependant variable in the stochastic parabolic equation. We have not attempted to do so as this is beyond the scope of the current article. Nevertheless, our numerical experiments in Section 4.1 clearly seem to validate this hypothesis.

Remark 2.3. Given a complete description of the input stochastic quantities (i.e. a functional form for q (say) $q = \mathcal{G}(\boldsymbol{\xi}_q)$, $\boldsymbol{\xi}\in\Gamma\subset\mathbb{R}^d$), using Theorem 1, one can construct a complete description of the dependant stochastic quantities. On the other hand, given only finite/limited information in terms of (say) the kth moment of the input stochastic quantities, Theorems 2.2 and 2.3 show that a unique kth moment solution exists. Our computational framework for the solution of the direct problem requires a complete description of the input stochastic quantities. Given k moments of the input stochastic quantities, it is possible to arbitrarily construct a stochastic process that *exactly* satisfies these moments. Now, using this input stochastic process, we can solve for the dependant stochastic variable in Eq. (6), with Theorems 2 and 3 guaranteeing that k moments of the dependant variable exist and are uniquely determined by the k moments of the input stochastic process.

It has to be noted here that we only have guarantees of uniqueness and existence of the moments/PDF of the dependant variable given the moments/PDF of the input stochastic variable. No such guarantee can be made on the existence (let alone uniqueness) of the moments of the input stochastic process given some data (in terms of moments/PDF) of the dependant variable. This is the typical ill-posedness phenomena seen in most inverse problems [19]. However, this problem can be partially resolved by expanding the solution space and posing the problem in a least-squares (Tikhonov) sense [20].

We now turn to the definition and solution strategy for the stochastic inverse problem.

3. The stochastic inverse problem

3.1. Physical motivation of the stochastic inverse problem: design in the presence of uncertainties

The motivation of the inverse problem is as follows: we are interested in designing a micro-scale heat sink that maintains a specific temperature profile at one end. The aim is to design the optimal heat flux on the other end such that this thermal profile is maintained. The thermal evolution implicitly depends on the material properties (here, the thermal diffusivity and heat capacity) of the device. The problem is complicated by the fact that the exact property variation of the device is unknown. That is, the complete microstructural description of the device is unknown (limitations in experimental characterization, time and cost restrictions preclude a complete characterization of the device). Only some statistical correlations about the microstructure are available. This necessitates a stochastic analysis assuming the thermal property to be a stochastic field. The problem definition is then as follows (see Fig. 3): Design the optimal heat flux on one end of the device such that a specified thermal profile is maintained on the other end (that is assumed to be insulated), in the presence of microstructural variability. The physical phenomena that directs the thermal evolution is simple diffusion. We will return to the solution of this problem in Section 4.2.

3.2. Definition of the inverse/design problem

The inverse/design problem is defined along the lines of the direct problem defined in Section 2. Without loss of generality, we assume that the heat flux $q(\boldsymbol{x},t,\boldsymbol{\xi}_q)$ is unknown in the boundary $\partial\mathcal{D}_h$. A set, $Y(\boldsymbol{x}_i,t)$, of

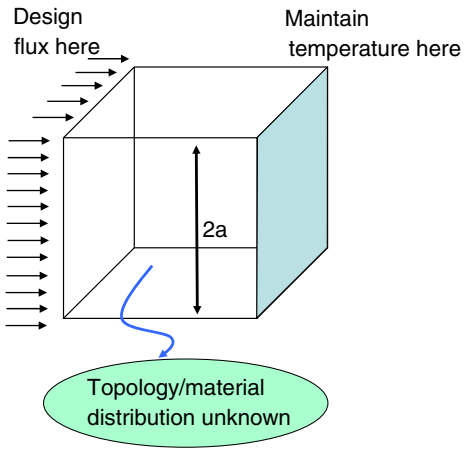

Fig. 3. Schematic of the design problem: The thermal diffusivity of the device is a random field. The variability of the heat flux at the left wall is sought leading to a desired temperature distribution on the right wall. The right wall is assumed to be insulated.

observed/desired temperature readings at a finite number of locations, $i = 1, \ldots, s$, in the domain (and/or boundary) are provided. These temperature measurements can be given in two forms:

- *Case I*: The first p moments of the temperature measurements at these sensor locations are given.
- *Case II*: The complete probability distribution functions of the temperature variation at these sensor locations are given.

Given this information, the problem is to construct the optimal stochastic heat flux, $q(\boldsymbol{x}, t), \boldsymbol{x} \in \partial\mathcal{D}_h$, such that the observed/desired temperature measurements (see Fig. 4) at the s sensor locations are maintained. The design problem can be stated for each of the two cases above as follows:

Case I: Find moments of the stochastic heat flux, $\langle q^k(\boldsymbol{x}, t, \boldsymbol{\xi}_q)\rangle$, $\boldsymbol{x} \in \partial\mathcal{D}_h$, $t \in \mathcal{T}$, $k = 1, \ldots, p$, such that the stochastic system defined by Eqs. (2)–(5) (with q in Eq. (3) taken as $q(\boldsymbol{x}, t, \boldsymbol{\xi}_q)$), results in a temperature profile that matches p moments of the temperature measurements, $\langle Y^k(\boldsymbol{x}_i, t)\rangle$, $k = 1, \ldots, p$ at s sensor locations $(\boldsymbol{x}_i, i = 1, \ldots, s)$.
Case II: Find the probability distribution function (PDF) of the stochastic heat flux, $Pr[q(\boldsymbol{x}, t, \boldsymbol{\xi}_q)]$, $\boldsymbol{x} \in \partial\mathcal{D}_h$, $t \in \mathcal{T}$, such that the stochastic system defined by Eqs. (2)–(5) (with q in Eq. (3) taken as $q(\boldsymbol{x}, t, \boldsymbol{\xi}_q)$), results in a temperature profile that matches the PDF of the temperature measurements, $Pr[Y^k(\boldsymbol{x}_i, t)]$ at s sensor locations $(\boldsymbol{x}_i, i = 1, \ldots, s)$.

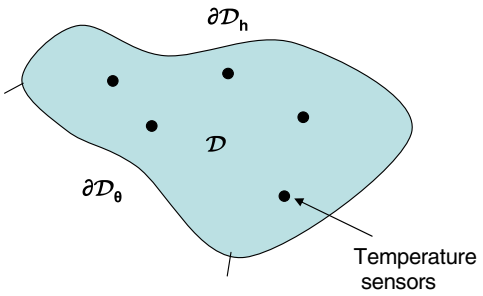

Fig. 4. Schematic of the inverse problem.

The motivation for posing the inverse problems in terms of construction of moments (CASE I) and the PDF (CASE II) of the stochastic heat flux is based on the arguments on physical observability/measurability made in Section 2.3.

Remark 3.1. In the context of design applications such as the one introduced in Fig. 3, the given moments or PDFs should be interpreted as 'desired' variability or performance robustness. In the context of inverse problems driven by data, this variability maybe induced either by sensor noise (in the case of a single experiment) or because of the variability of the random topology (repeated experimentation with random realizations of the medium will lead to variability in measurements even without measurement noise). All of these problems can be stated in the form of the problems given above.

3.3. Representing the stochastic design variable (heat flux): computational aspects

Following the usual procedure used in stochastic analysis [1], the unknown random process, q, can be represented as a function of n uncorrelated random variables, $\boldsymbol{\xi}_q = \{\xi_q^1, \ldots, \xi_q^n\}$. Following the rational of using a collocation based strategy for representing random processes, we can approximate the unknown stochastic heat flux as a polynomial interpolant using a finite number, n_q, of realizations of the flux over the stochastic space. Then, the stochastic flux at any point $\boldsymbol{\xi}_q$ is evaluated as

$$q(\boldsymbol{x}, t, \boldsymbol{\xi}_q) = \sum_{i=1}^{n_q} q(\boldsymbol{x}, t, \boldsymbol{\xi}_q^i) L_i(\boldsymbol{\xi}_q), \quad i = 1, \ldots, n_q, \tag{16}$$

where the $\boldsymbol{\xi}_q^i$ are collocation points in the stochastic space where the flux is sampled [9]. With this assumption, we have transferred the problem of computing a stochastic function, $q(\boldsymbol{x}, t, \boldsymbol{\xi}_q)$ (defined in $\partial\mathcal{D}_h \times \Omega_q$), into a problem of computing a finite set of deterministic functions, $q(\boldsymbol{x}, t, \boldsymbol{\xi}_q^i), i = 1, \ldots, n_q$ (defined in $\partial\mathcal{D}_h$).

Remark 3.2. Without loss of generality, we can assume that $\boldsymbol{\xi}_q = \{\xi_q^1, \ldots, \xi_q^n\}$ are uncorrelated uniform random variables defined in the unit hypercube $[0,1]^n$. In Section 3.7, we provide numerical illustrations of the invariance of the random process, q, to the choice of the random variables used to represent it.

Remark 3.3. In the first design case, we are interested in estimating/designing only the moments of the stochastic function, $\langle q^k \rangle = \int_{\Omega_q} q^k(\boldsymbol{\xi}_q)\mathcal{P}_q(\boldsymbol{\xi}_q)\mathrm{d}\boldsymbol{\xi}_q$ and in the second case we are interested in estimating the probability distribution function of this stochastic function $Pr(q \leqslant q^*)$. These can be viewed as essentially postprocessing operations once the stochastic heat flux representation in Eq. (16) is constructed. The design problem can now be stated in terms of estimating the heat flux at these finite number of collocation points.

Case I: Compute $q(\boldsymbol{x}, t, \boldsymbol{\xi}_q^i), \boldsymbol{x} \in \partial\mathcal{D}_h$, $i = 1, \ldots, n_q$, such that the stochastic system defined by Eqs. (2)–(5) (with q in Eq. (3) computed from Eq. (16)), results in a temperature profile that matches p moments of the temperature measurements, $\langle Y^k(\boldsymbol{x}_i, t)\rangle, k = 1, \ldots, p$ at s sensor locations $(\boldsymbol{x}_i, i = 1, \ldots, s)$.
Case II: Compute $q(\boldsymbol{x}, t, \boldsymbol{\xi}_q^i), \boldsymbol{x} \in \partial\mathcal{D}_h$, $i = 1, \ldots, n_q$, such that the stochastic system defined by Eqs. (2)–(5) (with q in Eq. (3) computed from Eq. (16)), results in a temperature profile that matches the PDF of the temperature measurements, $Pr[Y^k(\boldsymbol{x}_i, t)]$ at the s sensor locations $(\boldsymbol{x}_i, i = 1, \ldots, s)$.

Eq. (2) together with the known temperature boundary conditions on $\partial\mathcal{D}_\theta$ (Eq. (4)), initial conditions (Eq. (5)) and the temperature conditions (at the sensor locations) defines an *ill-posed* problem that can be solved to calculate the unknown stochastic heat flux q in $\partial\mathcal{D}_h$. We assume that a solution to the inverse problem exists in the sense of Tikhonov. We are looking for a solution $\{q_i^*\}_{i=1}^{n_q}$ such that

$$\mathcal{CF}[\{q_i^*\}_{i=1}^{n_q}] \leqslant \mathcal{CF}[\{q_i\}_{i=1}^{n_q}] \quad \forall q(\boldsymbol{x}, t, \boldsymbol{\xi}_q) = \sum_{i=1}^{n_q} q_i L_i(\boldsymbol{\xi}_q). \tag{17}$$

Here, $\mathcal{CF}[\{q_i\}_{i=1}^{n_q}]$ is a cost functional that quantifies how well the designed heat flux performs.

Remark 3.4. For a given $\{q_i\}_{i=1}^{n_q}$, the cost functional is computed using the temperature, $\theta(\boldsymbol{x}, t, \boldsymbol{\xi}_q, \boldsymbol{\xi}_\alpha : \{q_i\}_{i=1}^{n_q})$. This temperature is obtained from the solution of the direct problem in Eqs. (2)–(5) with q as a (known)

stochastic function, that is, using $\{q_i\}_{i=1}^{n_q}$ as the guessed heat flux on $\partial\mathcal{D}_h$, the known initial conditions and the known temperature conditions on $\partial\mathcal{D}_\theta$, solve Eq. (2) for the stochastic temperature. Note that for the sake of clarity, we write $\theta(\boldsymbol{x}, t, \boldsymbol{\xi}_q, \boldsymbol{\xi}_\alpha : \{q_i\}_{i=1}^{n_q})$ as $\theta(\boldsymbol{x}, t, \boldsymbol{\xi}_q, \boldsymbol{\xi}_\alpha)$.

This cost functional is usually quadratic in nature to utilize the extensive quadratic optimization tools/algorithms available. This definition of the inverse problem converts the above design problem into an optimization problem.

Remark 3.5. With the collocation based representation of the unknown stochastic heat flux (Eq. (16)), the stochastic inverse problem is transformed into a problem of determining a set of deterministic quantities. By specifying a quadratic cost functional, this *deterministic inverse problem* is posed as a *deterministic optimization problem*.

3.4. Definition of the cost functional

The design variables for both cases discussed above are now the heat flux at a finite number of collocation points $\{q_i\}, i = 1, \ldots, n_q$. The cost functional quantifies how well the designed heat flux performs.

Remark 3.6. Note that for the sake of clarity, we write $\theta(\boldsymbol{x}_j, t, \boldsymbol{\xi}_q^r, \boldsymbol{\xi}_\alpha^e : \{q_i\}_{i=1}^{n_q})$ as $\theta(\boldsymbol{x}_j, t, \boldsymbol{\xi}_q^r, \boldsymbol{\xi}_\alpha^e)$.

Case I: In the first case, the cost functional must measure how well the stochastic heat flux satisfies the moments of temperature at the sensor points. We define the cost functional in this case as

$$\begin{aligned}\mathcal{CF}[\{q_i\}_{i=1}^{n_q}] = \frac{1}{2}\sum_{j=1}^{s}\int_{t=0}^{t_{\max}} [\beta_1(\langle\theta(\boldsymbol{x}_j,t)\rangle - \langle Y(\boldsymbol{x}_j,t)\rangle)^2 + \beta_2(\langle\theta^2(\boldsymbol{x}_j,t)\rangle - \langle Y^2(\boldsymbol{x}_j,t)\rangle)^2 + \cdots \\ + \beta_p(\langle\theta^p(\boldsymbol{x}_j,t)\rangle - \langle Y^p(\boldsymbol{x}_j,t)\rangle)^2]\mathrm{d}t,\end{aligned} \tag{18}$$

where $\langle\theta^k(\boldsymbol{x}_j,t)\rangle$ is the kth moment of the temperature at the jth sensor location at time t, and the β_i are (user defined) positive scalars that weigh the different moments differently. Here, θ is the solution of the direct stochastic temperature evolution equation (see Remark 3.4) and Y is the desired temperature. Using Eq. (15) to represent the various moments in terms of the weights w, the cost functional can be written as

$$\begin{aligned}\mathcal{CF}[\{q_i\}_{i=1}^{n_q}] = \frac{1}{2}\sum_{j=1}^{s}\int_{t=0}^{t_{\max}} \Bigg[\beta_1\Bigg(\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \theta(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e) - \langle Y(\boldsymbol{x}_j,t)\rangle\Bigg)^2 \\ + \beta_2\Bigg(\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \theta^2(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e) - \langle Y^2(\boldsymbol{x}_j,t)\rangle\Bigg)^2 \\ + \cdots + \beta_p\Bigg(\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \theta^p(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e) - \langle Y^p(\boldsymbol{x}_j,t)\rangle\Bigg)^2\Bigg]\mathrm{d}t.\end{aligned} \tag{19}$$

This can be compactly written as

Cost functional – Case I:

$$\mathcal{CF}[\{q_i\}_{i=1}^{n_q}] = \sum_{\gamma=1}^{p}\frac{1}{2}\int_{t=0}^{t_{\max}}\Bigg[\sum_{j=1}^{s}\Bigg[\beta_\gamma\Bigg(\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \theta^\gamma(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e) - \langle Y^\gamma(\boldsymbol{x}_j,t)\rangle\Bigg)^2\Bigg]\mathrm{d}t, \tag{20}$$

where γ loops over the p moments of the stochastic temperature.

Case II: In this case, the cost functional must measure how well the stochastic heat flux results in the probability distribution function of temperature at the sensor points being satisfied.

Since the probability distribution function, the cumulative distribution function and inverse cumulative distribution function represent the same distribution in different ways (see Fig. 5), for the sake of computational effectiveness, the cost functional is defined in terms of the inverse cumulative distribution

function. This choice is also motivated by the fact that the inverse CDF has a fixed, known support ($[0,1]$) and this greatly simplifies defining the cost functional. Denote $\Upsilon^{-1}[f,u]$ as the inverse cumulative distribution of the random variable f. u is the spanning variable, $u \in [0,1]$ (Fig. 5 right). The inverse CDF is utilized because it is computationally easy to represent it using the collocation based polynomial interpolation representation of the temperature (Eq. (13)). The cost functional is defined as

Cost functional – Case II:

$$\mathcal{CF}[\{q_i\}_{i=1}^{n_q}] = \frac{1}{2}\sum_{j=1}^{s}\int_{t=0}^{t_{\max}}\int_0^1 (\Upsilon^{-1}[\theta(\boldsymbol{x}_j,t),u] - \Upsilon^{-1}[Y(\boldsymbol{x}_j,t),u])^2 \mathrm{d}u\,\mathrm{d}t. \tag{21}$$

Consider the stochastic temperature at a sensor location $\boldsymbol{x}_s$ at some time t. This can be represented in terms of the computed collocated temperatures as

$$\theta(\boldsymbol{x}_s,t,\boldsymbol{\xi}_q,\boldsymbol{\xi}_\alpha) = \sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha}\theta(\boldsymbol{x}_s,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)L_r(\boldsymbol{\xi}_q)L_e(\boldsymbol{\xi}_\alpha), \tag{22}$$

where $\boldsymbol{\xi}_q = \{\xi_q^1,\ldots,\xi_q^n\}$ and $\boldsymbol{\xi}_\alpha = \{\xi_\alpha^1,\ldots,\xi_\alpha^m\}$. We have assumed that $\xi_q^i, i = 1,\ldots,n$ are uniform random variables in $[0,1]$. Υ^{-1} is a function mapping $[0,1]$ to the range of $\theta(\boldsymbol{x}_s,t)$. It follows that Υ^{-1} can be expressed as the marginal of the temperature over one of the stochastic dimensions:

$$\begin{aligned}
\Upsilon^{-1}[\theta(\boldsymbol{x}_s,t),\xi_1] &= \int_{\xi_q^2=0}^1\cdots\int_{\xi_q^n=0}^1\int_{\Omega_\alpha}\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha}\theta(\boldsymbol{x}_s,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)L_r(\xi_q^1,\xi_q^2,\ldots,\xi_q^n)L_e(\boldsymbol{\xi}_\alpha)\mathrm{d}\xi_q^2\cdots\mathrm{d}\xi_q^n\,\mathrm{d}\boldsymbol{\xi}_\alpha \\
&= \sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha}\theta(\boldsymbol{x}_s,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)\int_{\xi_q^2=0}^1\cdots\int_{\xi_q^n=0}^1\int_{\Omega_\alpha}L_r(\xi_q^1,\xi_q^2,\ldots,\xi_q^n)L_e(\boldsymbol{\xi}_\alpha)\mathrm{d}\xi_q^2\cdots\mathrm{d}\xi_q^n\,\mathrm{d}\boldsymbol{\xi}_\alpha \\
&= \sum_{r=1}^{n_q}\left[\sum_{e=1}^{n_\alpha}w_\alpha^e\theta(\boldsymbol{x}_s,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)\right]\int_{\xi_q^2=0}^1\cdots\int_{\xi_q^n=0}^1 L_r(\xi_q^1,\xi_q^2,\ldots,\xi_q^n)\mathrm{d}\xi_q^2\cdots\mathrm{d}\xi_q^n \\
&= \sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha}w_\alpha^e\theta(\boldsymbol{x}_s,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)f_r(\xi_q^1).
\end{aligned} \tag{23}$$

For the sake of brevity, we will denote this as

$$\Upsilon^{-1}[\theta(\boldsymbol{x}_s,t),u] = \sum_{r=1}^{n_q}\left\{\sum_{e=1}^{n_\alpha}w_\alpha^e\theta(\boldsymbol{x}_s,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)\right\}f_r(u). \tag{24}$$

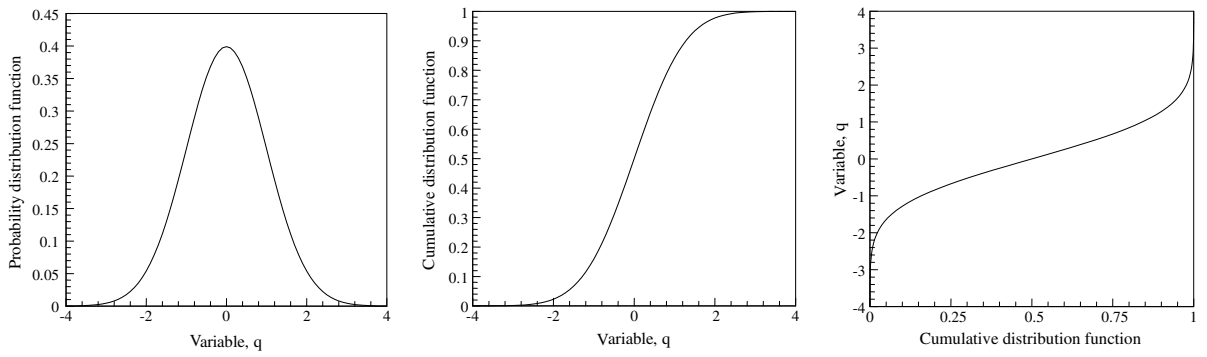

Fig. 5. The bijective relation between a PDF and its inverse CDF. The cumulative integral of the PDF gives the CDF. A simple rotation of the CDF (interchanging the axes) gives the inverse CDF.

3.5. Optimization strategy: gradient of the cost functional and the stochastic sensitivity equations

A gradient based optimization strategy is used to design the optimal heat flux $\{q_i\}$. The first step is to compute the gradient of the cost functional with respect to the design variables. The directional derivative, $\mathrm{D}_{\Delta q_v}\mathcal{CF}[\{q_i\}_{i=1}^{n_q}]$ of the cost functionals in Eqs. (20) and (21) can be computed as follows:

Gradient of the cost functional – Case I:

$$\mathrm{D}_{\Delta q_v}\mathcal{CF}[\{q_i\}_{i=1}^{n_q}] = \sum_{\gamma=1}^{p}\sum_{j=1}^{s}\int_{t=0}^{t_{\max}}\left[\beta_\gamma\left(\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \theta^\gamma(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e) - \langle Y^\gamma(\boldsymbol{x}_j,t)\rangle\right)\right.$$
$$\left.\times\,\mathrm{D}_{\Delta q_v}\left[\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha}\theta^\gamma(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e) w_q^r w_\alpha^e\right]\right]\mathrm{d}t. \tag{25}$$

Gradient of the cost functional – Case II:

$$\mathrm{D}_{\Delta q_v}\mathcal{CF}[\{q_i\}_{i=1}^{n_q}] = \sum_{j=1}^{s}\int_{t=0}^{t_{\max}}\int_0^1 (\Upsilon^{-1}[\theta(\boldsymbol{x}_j,t),u] - \Upsilon^{-1}[Y(\boldsymbol{x}_j,t),u])$$
$$\times\,\mathrm{D}_{\Delta q_v}\left(\sum_{r=1}^{n_q}\left\{\sum_{e=1}^{n_\alpha} w_\alpha^e\theta(\boldsymbol{x}_s,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)\right\} f_r(u)\right)\mathrm{d}u\,\mathrm{d}t. \tag{26}$$

The directional derivative of the cost functional with respect to a design variable q_v depends on evaluation of the terms $\mathrm{D}_{\Delta q_v}(\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha}\theta^\gamma w_q^r w_\alpha^e)$. This can be simplified as follows:

$$\mathrm{D}_{\Delta q_v}\left(\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha}\theta^\gamma(\boldsymbol{x},t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e) w_q^r w_\alpha^e\right) = \sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \mathrm{D}_{\Delta q_v}\theta^\gamma(\boldsymbol{x},t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)$$
$$= \sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \gamma\theta^{\gamma-1}(\boldsymbol{x},t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)\mathrm{D}_{\Delta q_v}\theta(\boldsymbol{x},t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e). \tag{27}$$

This requires the calculation of $\mathrm{D}_{\Delta q_v}\theta(\boldsymbol{x},t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)$. Define the directional derivative of the temperature with respect to each design variable as $\Theta(\boldsymbol{x},t,\boldsymbol{\xi}_q,\boldsymbol{\xi}_\alpha:\{q_i\}_{i=1}^{n_q},\Delta q_v) \equiv \mathrm{D}_{\Delta q_v}\theta(\boldsymbol{x},t,\boldsymbol{\xi}_q,\boldsymbol{\xi}_\alpha:\{q_i\}_{i=1}^{n_q})$. This defines the sensitivity temperature field as the linear in Δq_v part of $\theta(\boldsymbol{x},t,\boldsymbol{\xi}_q,\boldsymbol{\xi}_\alpha:\{q_i\}_{i=1,i\neq v}^{n_q},q_v+\Delta q_v)$, where Δq_v is a perturbation to one of the design variables:

$$\theta(\boldsymbol{x},t,\boldsymbol{\xi}_q,\boldsymbol{\xi}_\alpha:\{q_i\}_{i=1,i\neq v}^{n_q},q_v+\Delta q_v) = \theta(\boldsymbol{x},t,\boldsymbol{\xi}_q,\boldsymbol{\xi}_\alpha:\{q_i\}_{i=1}^{n_q}) + \Theta(\boldsymbol{x},t,\boldsymbol{\xi}_q,\boldsymbol{\xi}_\alpha:\{q_i\}_{i=1}^{n_q},\Delta q_v) + \text{h.o.t.} \tag{28}$$

The stochastic temperature sensitivity equations are simply obtained by taking the directional derivative of the equations that define the parametric direct problem used to compute θ for each heat flux q. These directional derivatives are taken with respect to n_q collocation points used to represent q (i.e. w.r.t. the design variables).

Stochastic sensitivity equations:

$$\frac{\partial\Theta(\boldsymbol{x},t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e:\{q_k\}_{k=1}^{n_q},\Delta q_v)}{\partial t} = \nabla[\alpha(\boldsymbol{x},\boldsymbol{\xi}_\alpha^e)\nabla\Theta(\boldsymbol{x},t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e:\{q_k\}_{k=1}^{n_q},\Delta q_v)], \tag{29}$$

$$\frac{\partial\Theta(\boldsymbol{x},t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e:\{q_k\}_{k=1}^{n_q},\Delta q_v)}{\partial\boldsymbol{n}} = \Delta q_v\delta_{v,r}, \quad \boldsymbol{x}\in\partial\mathcal{D}_h, \tag{30}$$

$$\Theta(\boldsymbol{x},t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e:\{q_k\}_{k=1}^{n_q},\Delta q_v) = 0, \quad \boldsymbol{x}\in\partial\mathcal{D}_\theta, \tag{31}$$

$$\Theta(\boldsymbol{x},0,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e:\{q_k\}_{k=1}^{n_q},\Delta q_v) = 0, \quad \boldsymbol{x}\in\mathcal{D}. \tag{32}$$

Notice the elegant decoupling that occurs in Eq. (30) above. The sensitivity of the stochastic temperature computed at the ξ_q^i collocation point depends only on perturbations to the q_ith design variable. Also note that the sensitivity equations depend only on the perturbation to the heat flux, Δq_v and not on the heat flux, q_v. This is due to the fact that the heat flux enters in Eq. (6) as a linear term. Furthermore, no other variable/parameter depends nonlinearly on the heat flux.

Returning back to Eq. (27), this can now be rewritten as

$$\begin{aligned}
\mathrm{D}_{\Delta q_v}\left(\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha}\theta^\gamma(\boldsymbol{x},t,\xi_q^r,\xi_\alpha^e)w_q^r w_\alpha^e\right) &= \sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \mathrm{D}_{\Delta q_v}\theta^\gamma(\boldsymbol{x},t,\xi_q^r,\xi_\alpha^e)\\
&= \sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \gamma\theta^{\gamma-1}(\boldsymbol{x},t,\xi_q^r,\xi_\alpha^e)\mathrm{D}_{\Delta q_v}\theta(\boldsymbol{x},t,\xi_q^r,\xi_\alpha^e)\\
&= \sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \gamma\theta^{\gamma-1}(\boldsymbol{x},t,\xi_q^r,\xi_\alpha^e)\Theta(\boldsymbol{x},t,\xi_q^r,\xi_\alpha^e:\{q_k\}_{k=1}^{n_q},\Delta q_v)\delta_{v,r}\\
&= \sum_{e=1}^{n_\alpha} w_q^v w_\alpha^e \gamma\theta^{\gamma-1}(\boldsymbol{x},t,\xi_q^v,\xi_\alpha^e)\Theta(\boldsymbol{x},t,\xi_q^v,\xi_\alpha^e:\{q_k\}_{k=1}^{n_q},\Delta q_v). \qquad (33)
\end{aligned}$$

Thus, the calculation of the gradient of the cost functional requires the computation of the sensitivity of the stochastic temperature to perturbations in these design variables. The complete stochastic temperature sensitivity can be constructed as

$$\Theta(\boldsymbol{x},t,\xi_q,\xi_\alpha:\{q_k\}_{k=1}^{n_q},\Delta q_v)=\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha}\Theta(\boldsymbol{x},t,\xi_q^r,\xi_\alpha^e:\{q_k\}_{k=1}^{n_q},\Delta q_v)L_r(\xi_q)L_e(\xi_\alpha). \qquad (34)$$

Denote the gradient of the cost functional with respect to the design variables $\{q_i\}_{i=1}^{n_q}$ as $\boldsymbol{d}=\{d_i\}^T$. The gradient of the cost functional can be written in terms of the directional derivative (the directional derivative is just the gradient computed in a specific direction). Since we have a scheme to compute the directional derivative by solving the continuum stochastic sensitivity equations, the gradient of the cost function is then simply given as $\boldsymbol{d}=\{d_v\}^T=\{\mathrm{D}_{\Delta q_v}\mathcal{CF}[\{q_i\}_{i=1}^{n_q}]/\Delta q_v\}^T$.

3.6. Optimization algorithm

In the current work, the gradient of the cost functional is evaluated and used in a steepest descent minimization framework to estimate the optimal stochastic heat flux parameters. This procedure is schematically illustrated in Fig. 6.

Remark 3.7. With known $\{q_i\}_{i=1}^{n_q}$, computing the direct temperature is a fully-decoupled procedure involving the solution of $n_q \times n_\alpha$ deterministic problems. The evaluation of the cost functional involves either computing moments of the temperature or the inverse CDF of the temperature. This step involves collecting data (from each decoupled deterministic problem) and computing the necessary quantities ($\langle\theta^p\rangle$ or $\Upsilon^{-1}[\theta,u]$). The solution of the stochastic sensitivity equations is again a fully-decoupled operation. This presents the possibility of solving this problem in an embarrassingly parallel format where the only places where any communication is required is in the construction of the cost functional and the construction of the gradient with respect to the n_q collocation points.

Step size in the steepest descent algorithm: After computing the gradient of the cost functional, $\boldsymbol{d}$, the updated values of the heat flux q_i^{iter+1} are given as $q_i^{iter+1}=q_i^{iter}-vd_i$, where v is the step size of descent along the negative of the gradient. The step size for each of the two cases is given as

Step size, Case I:

$$v=\frac{\boldsymbol{d}^T\boldsymbol{d}}{\sum_{\gamma=1}^{p}\beta_\gamma\mathbf{R}_\gamma}, \qquad (35)$$

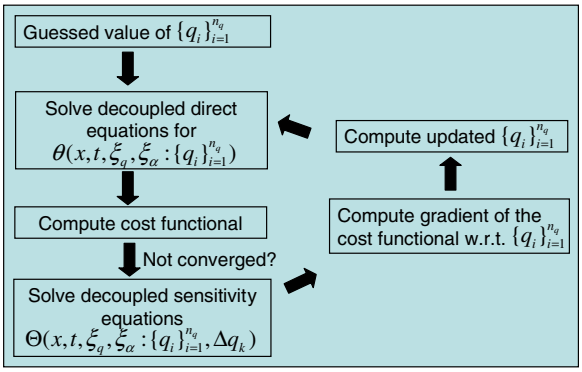

Fig. 6. Schematic of the inverse problem.

where

$$\mathbf{R}_\gamma = \sum_{j=1}^{s}\int_{t=0}^{t_{\max}}\left\{\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \gamma\theta^{\gamma-1}(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)\frac{\Theta(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e:\{q_k\}_{k=1}^{n_q},\Delta q_r)}{\Delta q_r}d_r\right\}^2 \mathrm{d}t$$
$$+\sum_{j=1}^{s}\int_{t=0}^{t_{\max}}\left(\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \theta^{\gamma}(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e) - \langle Y^\gamma(\boldsymbol{x}_j,t)\rangle\right)$$
$$\times\left\{\sum_{r=1}^{n_q}\sum_{e=1}^{n_\alpha} w_q^r w_\alpha^e \gamma(\gamma-1)\theta^{\gamma-2}(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e)\left[\frac{\Theta(\boldsymbol{x}_j,t,\boldsymbol{\xi}_q^r,\boldsymbol{\xi}_\alpha^e:\{q_k\}_{k=1}^{n_q},\Delta q_r)}{\Delta q_r}d_r\right]^2\right\}\mathrm{d}t. \tag{36}$$

Step size, Case II:

$$v = \frac{\boldsymbol{d}^T\boldsymbol{d}}{\boldsymbol{d}^T\mathbf{R}\boldsymbol{d}}, \tag{37}$$

where

$$\mathbf{R}_{ij} = \sum_{r=1}^{s}\int_{t=0}^{t_{\max}}\left\{\sum_{k=1}^{n_\alpha} w_\alpha^k \Theta(\boldsymbol{x}_r,t,\boldsymbol{\xi}_q^i,\boldsymbol{\xi}_\alpha^k:\{q_k\}_{k=1}^{n_q},\Delta q_i)\right\}\times\left\{\sum_{k=1}^{n_\alpha} w_\alpha^k \Theta(\boldsymbol{x}_r,t,\boldsymbol{\xi}_q^j,\boldsymbol{\xi}_\alpha^k:\{q_k\}_{k=1}^{n_q},\Delta q_j)\right\}\mathrm{d}t. \tag{38}$$

The complete optimization algorithm is given as

Optimization algorithm:

- *Step 1*: Select initial guess values for the flux parameters $\{q_i\}_{i=1}^{n_q}$. Set the iteration counter $iter = 0$.
- *Step 2*: Solve the direct problem (defined in Eq. (12)) using the guessed parameters.
- *Step 3*: Compute the cost functional $\mathcal{CF}[\{q_i\}_{i=1}^{n_q}]$ (using Eq. (20) or Eq. (21)). If $\mathcal{CF}[\{q_i\}_{i=1}^{n_q}] < tolerance$ STOP, else GOTO Step 4.
- *Step 4*: Compute the gradient of the cost functional $\nabla\mathcal{CF}[\{q_i\}_{i=1}^{n_q}]$ (using Eq. (25) or Eq. (26)).
- *Step 5*: Compute the step size, v, for use in the steepest descent method (using Eq. (35) or Eq. (37)).
- *Step 6*: Compute the updated flux parameters $q_i^{iter+1} = q_i^{iter} - v\times\nabla_i\mathcal{CF}[\{q_i\}_{i=1}^{n_q}]$. Update the iteration counter by 1. GOTO Step 2.

3.7. Numerical illustration: 1D inverse heat conduction problem

The methodology discussed above is numerically illustrated here using a series of one-dimensional stochastic inverse heat conduction problems (see Fig. 7). A one-dimensional rod of length $L = 1.0$ has a temperature

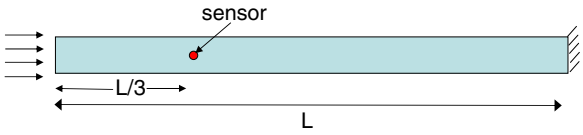

Fig. 7. Schematic of the 1D inverse heat conduction problem.

sensor placed at a location one-third the length of the rod. The right end is maintained at $T = 0$. The thermal diffusivity, α, of the material is random and follows a chi-square distribution $\alpha = 1 + 0.1(\xi^2 - 1)$, where $\xi \sim N(0,1)$. The problem is to reconstruct the steady-state stochastic heat flux on the left end that results in the measured temperature. For the sake of clarity and presentation, we consider the non-dimensionalized form of the equations (unless specified otherwise). Two cases are examined: (I) when only moments of the temperature measured at the sensor are provided and (II) when the PDF of the temperature measured at the sensor is provided.

Case I: Moments of temperature provided. The first 15 moments of the temperature at the sensor location are provided. This temperature is computed by first solving a direct problem with a known stochastic heat flux. The known stochastic heat flux (denoted as 'experimental' in the figures and tables) follows a Beta $(p = 1.5, q = 2)$ distribution:

$$\text{Beta distribution}: \quad \xi \in (0,1), \quad \rho(\xi) = \frac{\xi^{p-1}(1-\xi)^{q-1}}{\int_0^1 t^{p-1}(1-t)^{q-1}\mathrm{d}t}.$$

The direct problem to obtain the 'experimental' sensor location temperature is solved using $n_q = 1025$ collocation points to represent the stochastic heat flux, $n_\alpha = 65$ collocation points to represent the stochastic diffusivity and using 80 linear finite elements to discretize the spatial domain. It has to be emphasized that very high levels of interpolation representation of the stochastic spaces as well as the spatial domain (compared to the discretization used in the inverse problem) are considered to prevent the possibility of an inverse crime being committed.

Convergence with respect to increasing depth of interpolation: A representation based on a uniform random variable is chosen to construct the optimal stochastic heat flux i.e. $q(\xi) = \sum_{i=1}^{n_q} q(\xi_i) L_i(\xi)$, where $\xi \in [0,1]$. The optimization problem was run using different depths $(2,4,6)$ of interpolation (corresponding to increasing values of $n_q = 5, 17, 65$). Fig. 8 plots the variation in the cost functional for each of the optimization problems. Fig. 8 allows us to draw several interesting conclusions about the behavior of the optimization framework. Notice that the cost functional decreases by two orders of magnitude very quickly (within 15 iterations) in all three optimization cases (levels 2, 4 and 6). Following this initial rapid decrease, each optimization problem has a different rate of decrease of the cost functional. It appears that the optimization problem with the lowest number of design variables (levels 2 and 5 design variables) has the fastest decrease in the cost functional followed by levels 4 and 6 optimization problems.

The behavior observed is explained by looking at the form of the cost functional in Eq. (20). We have chosen all the β_i to be 1.0 in all the numerical examples. It has to be noted that this particular choice of β_i results in equal weights to all the moments. In contrast, if one knows a priori that the higher moments are much smaller than the lower moments, appropriately weighted moments ($\beta_i \ll \beta_j$ for $i < j$) can be used to ensure that the higher moments are captured accurately. Level 2 optimization problem quickly captures the first few moments and the cost functional rapidly approaches an asymptotic value. There are too few variables to completely capture the required behavior. On the other hand, due to the larger number of design variables (for the level 4 and level 6 optimization problems), the design space is explored at a slower rate, leading to the slower rate of convergence.

One way of verifying this hypothesis is to let the optimization process run for a large number of iterations and analyze the behavior of the cost functional. The limited number of design variables in level 2 case should cause the cost functional of this case to remain almost stagnant, while the cost functional of the other two cases should decrease. Since, each iteration of this one-dimensional problem takes less than 30 s to complete on a serial machine, we run all three optimization problems for an additional 1000 iterations. Fig. 9 plots the

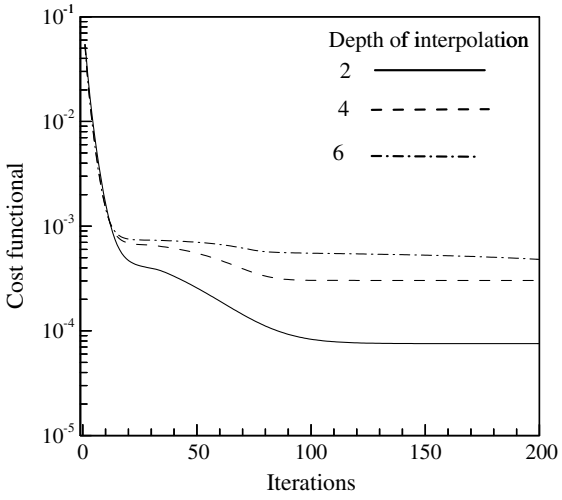

Fig. 8. Variation of the cost functional for different depths of interpolation.

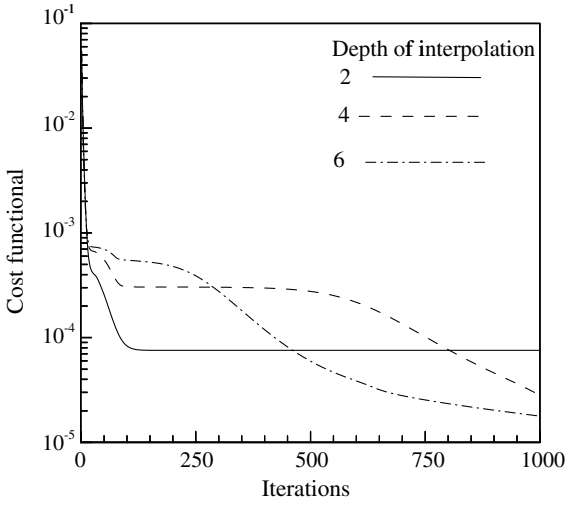

Fig. 9. Variation of the cost functional for different depths of interpolation.

corresponding reduction in the cost functional for these problems, which follows the expected behavior. Another observation that one can make looking at Figs. 8 and 9 is the 'arm-chair' like behavior of the cost functional. In all three cases, the cost functional has regions of rapid decrease, followed regions of almost asymptotic behavior, followed again by regions of rapid decrease. This pattern seems to reappear as the number of iterations are increased. Interestingly, the region of asymptotic behavior (i.e. the number of iterations where the decrease in the cost functional is very small) becomes smaller as the level of interpolation increases.

Fig. 10 shows a comparison of the first 15 moments of the optimal heat flux computed from each of the optimization problems with respect to the 'experimental' heat flux. Notice that the first few ($\sim$7) moments are accurately captured even by a level 2 interpolation scheme. The results match the 'experimental' moments very accurately as the depth of interpolation is increased, with no difference between the actual value and those computed using a depth 8 interpolation.

Remark 3.8. From the above numerical experiments, the following conclusions can be drawn: (1) The optimization problem with coarser representation (i.e. smaller level of interpolation) of the design variable

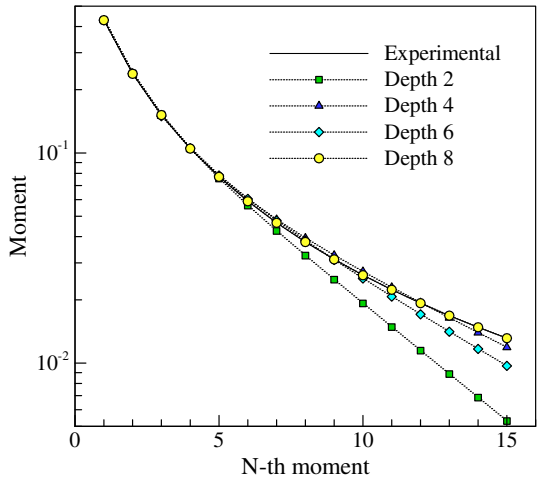

Fig. 10. Comparison of the optimal computed heat flux with the actual heat flux.

converges faster initially, but becomes very slowly converging later and (2) the optimization problem with coarser representation is numerically faster to solve. These two observations naturally lead to the possibly of formulating a *hierarchical stochastic optimization* problem, where a coarser problem is solved to quickly compute a coarse solution that is in turn used as an initial guess to solve a finer problem. This idea (similar to the accelerated convergence using multi-grid methods) seems to offer great promise in terms of rapid solutions to inverse problems through solution of a hierarchy of coarser optimization problems. We make no attempt to investigate this novel concept in the present paper. Nevertheless, we have applied a rudimentary version of this concept to solve the above problem.

A hierarchical stochastic optimization framework: Starting from a level 2 representation of the unknown heat flux, the optimization problem is solved until the cost functional reduces to a value below an a priori defined cutoff ($\sim 2 \times 10^{-4}$). The stochastic solution at this stage was used as an initial guess to start an optimization with a level 4 representation. This was solved until the cost functional reduces to a value below an a priori defined cutoff ($\sim 2 \times 10^{-5}$). The optimization was stopped when the cost functional reached a value of 4×10^{-6}. Fig. 11 plots the reduction in the cost functional using this hierarchical multi-level optimization method as well as the conventional single level (level 6) optimization method. The jumps in the cost functional

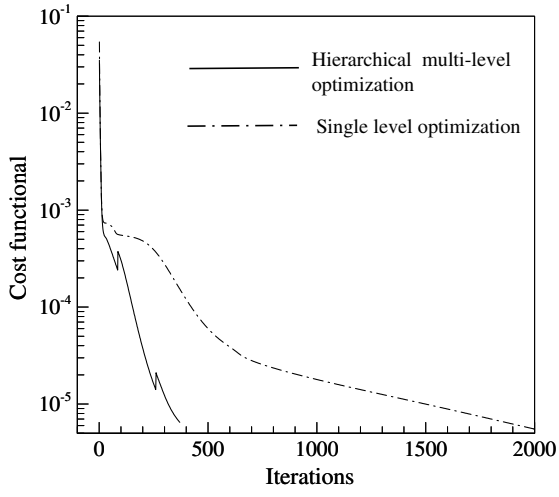

Fig. 11. Comparison of the cost functional of the single level and the hierarchical multi-level optimization algorithm.

represent the points where a finer approximation of the stochastic space is utilized. We have made no attempt to optimize the location where this transition to the next higher-level occurs. The figure shows that the hierarchical multi-level optimization strategy is indeed a promising approach to solving stochastic inverse problems (with almost an order in magnitude decrease in computational effort).

Invariance with the choice of representation: An analysis of the effect of choosing different representations to construct the stochastic heat flux is performed next. Four different representations of q based on different supports and associated PDFs are chosen as follows:

Beta distribution:

$$q(\xi) = \sum_{i=1}^{n_q} q(\xi_i) L_i(\xi), \quad \xi \in (0,1), \quad \rho(\xi) = \frac{\xi^{p-1}(1-\xi)^{q-1}}{\int_0^1 t^{p-1}(1-t)^{q-1}\,\mathrm{d}t}.$$

Gamma distribution:

$$q(\xi) = \sum_{i=1}^{n_q} q(\xi_i) L_i(\xi), \quad \xi \in (0,\infty), \quad \rho(\xi) = \frac{\xi^{\gamma-1}\,\mathrm{e}^{-\xi}}{\int_0^\infty t^{\gamma-1}\,\mathrm{e}^{-t}\,\mathrm{d}t}.$$

Normal distribution:

$$q(\xi) = \sum_{i=1}^{n_q} q(\xi_i) L_i(\xi), \quad \xi \in (-\infty,\infty), \quad \rho(\xi) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\xi^2}{2}}.$$

Uniform distribution:

$$q(\xi) = \sum_{i=1}^{n_q} q(\xi_i) L_i(\xi), \quad \xi \in [0,1], \quad \rho(\xi) = 1.$$

The four optimization problems were run with $n_q = 257$ collocation points to represent the domain of ξ. Fig. 12 plots the reduction in the cost functional as a function of the optimization iteration counter for each of the four cases stated above. Table 1 tabulates the results of the optimization problems. The first column consists of the first 10 moments of the 'experimental' stochastic heat flux. The next four columns consist of the first 10 moments of the reconstructed stochastic heat flux based on the four different representations of the random process. Notice that there is negligible difference between the reconstructed values of the optimal stochastic heat flux and the 'experimental' values. The cost functional in all the four cases is less than 10^{-6}. This corresponds to a reduction in the cost functional by five orders of magnitude. Notice that there is some

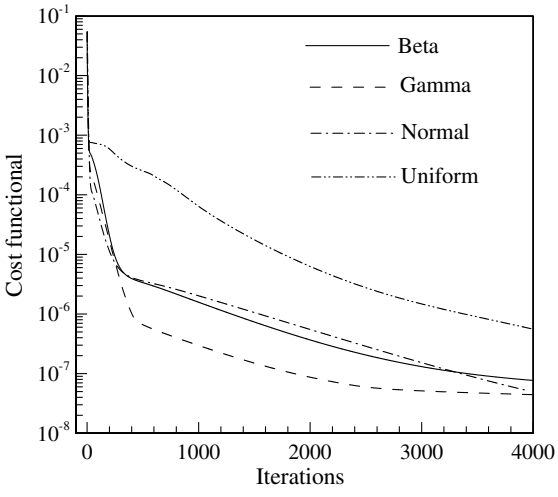

Fig. 12. Variation of the cost functional for optimizations based on different representations of the stochastic heat flux.

Table 1
Comparison of the first 10 moments

*N*th moment	Experimental	Beta	Gamma	Normal	Uniform
1	0.428581	0.428556	0.428648	0.428452	0.428561
2	0.238119	0.238023	0.237794	0.238350	0.238824
3	0.151551	0.151934	0.151804	0.151774	0.149193
4	0.104942	0.105334	0.105383	0.105039	0.105726
5	0.076981	0.076984	0.077244	0.076812	0.077641
6	0.058892	0.058310	0.058791	0.058348	0.060955
7	0.046519	0.045299	0.045992	0.045543	0.048112
8	0.037683	0.035853	0.036737	0.036266	0.039379
9	0.031155	0.028783	0.029827	0.029317	0.032157
10	0.026195	0.023366	0.024536	0.023977	0.026815

similarity in how well each of the representations captures the moments. For instance, all the representations capture the first two moments to 4 decimal digits of accuracy. Similarly, all four schemes compute the ninth and tenth moments with 2 decimal digits of accuracy. Since the aim of this exercise was to show the invariance of resulting solution on the representation of the unknown variable, further investigations on utilizing the hierarchical multi-level optimization method are not performed.

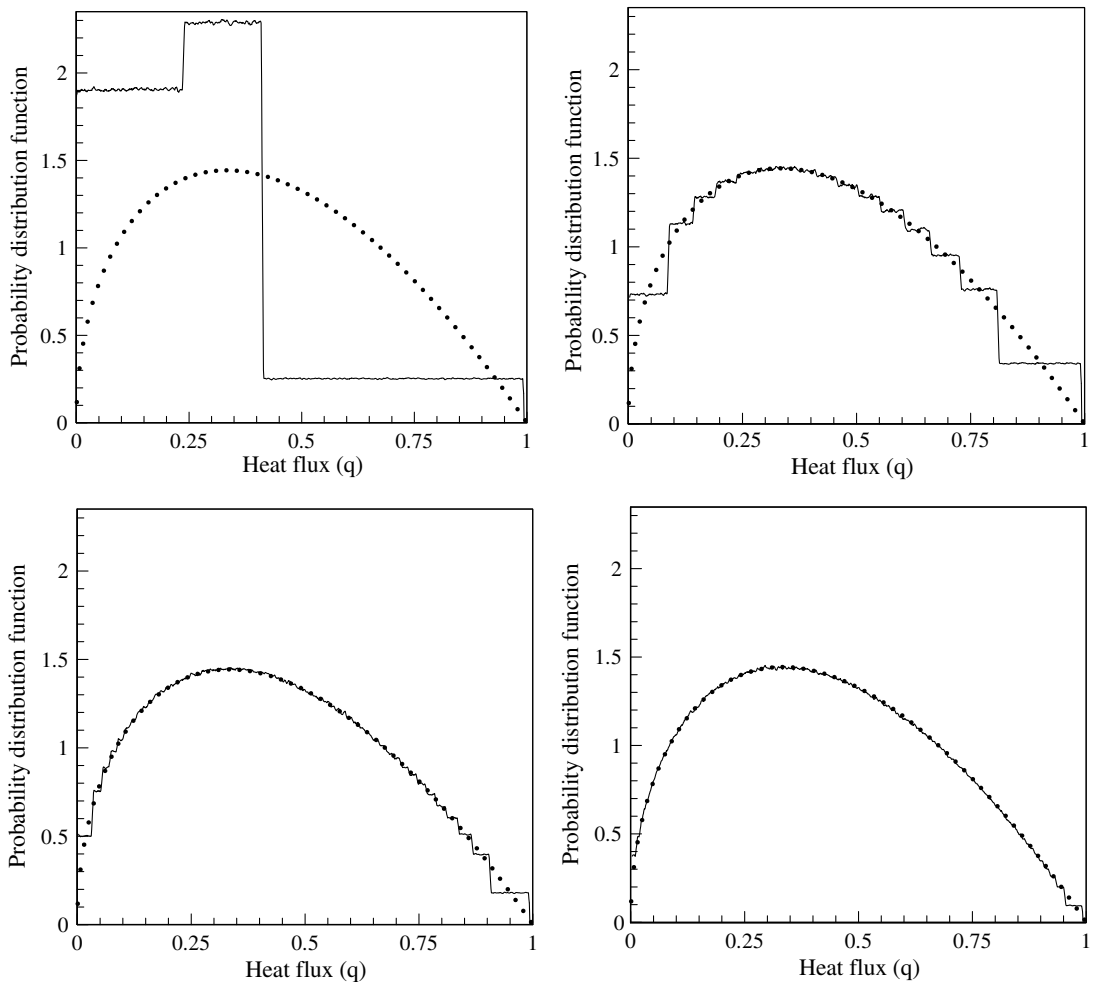

Fig. 13. Optimal PDF of stochastic heat flux computed with various depths of interpolation ($q \sim \mathrm{Beta}(1.5, 2)$). Top row: Depth of interpolation 2 and 4. Bottom row, depth of interpolation 6 and 8. The filled circles represent the 'experimental' data and the solid lines represent the reconstructed solution.

Case II: PDF of temperature provided.

In this numerical investigation, the PDF of the temperature at the sensor location is provided. This temperature is computed by first solving a direct problem with a known stochastic heat flux. Two cases are considered here: (a) the known stochastic heat flux has finite support (Beta distribution); (b) the known stochastic heat flux has unbounded support (Normal distribution). The known stochastic heat flux follows a Beta $(p = 1.5, q = 2)$ distribution (for case (a)) and Normal distribution $(\mu = 1.0, \sigma = 0.1)$ (for case (b)). The direct problem (to define the 'experimental' sensor location temperature) is solved using $n_q = 1025$ collocation points to represent the stochastic heat flux, $n_\alpha = 65$ collocation points to represent the stochastic diffusivity and using 80 element to discretize the spatial domain.

Convergence with respect to increasing depth of interpolation: A representation based on a uniform random variable is chosen to construct the optimal stochastic heat flux i.e. $q(\xi) = \sum_{i=1}^{n_q} q(\xi_i) L_i(\xi)$, where $\xi \in [0, 1]$. The optimization problem was run using different depths $(2, 4, 6, 8)$ of interpolation (corresponding to increasing values of $n_q = 5, 17, 65, 257$).

Fig. 13 plots the computed optimal PDF of the stochastic heat flux for four optimization processes using different depth of interpolation for case (a), where the 'experimental' q has finite support. Similarly, Fig. 14 compares the reconstructed solution with the actual 'experimental' result for case (b), where the 'experimental'

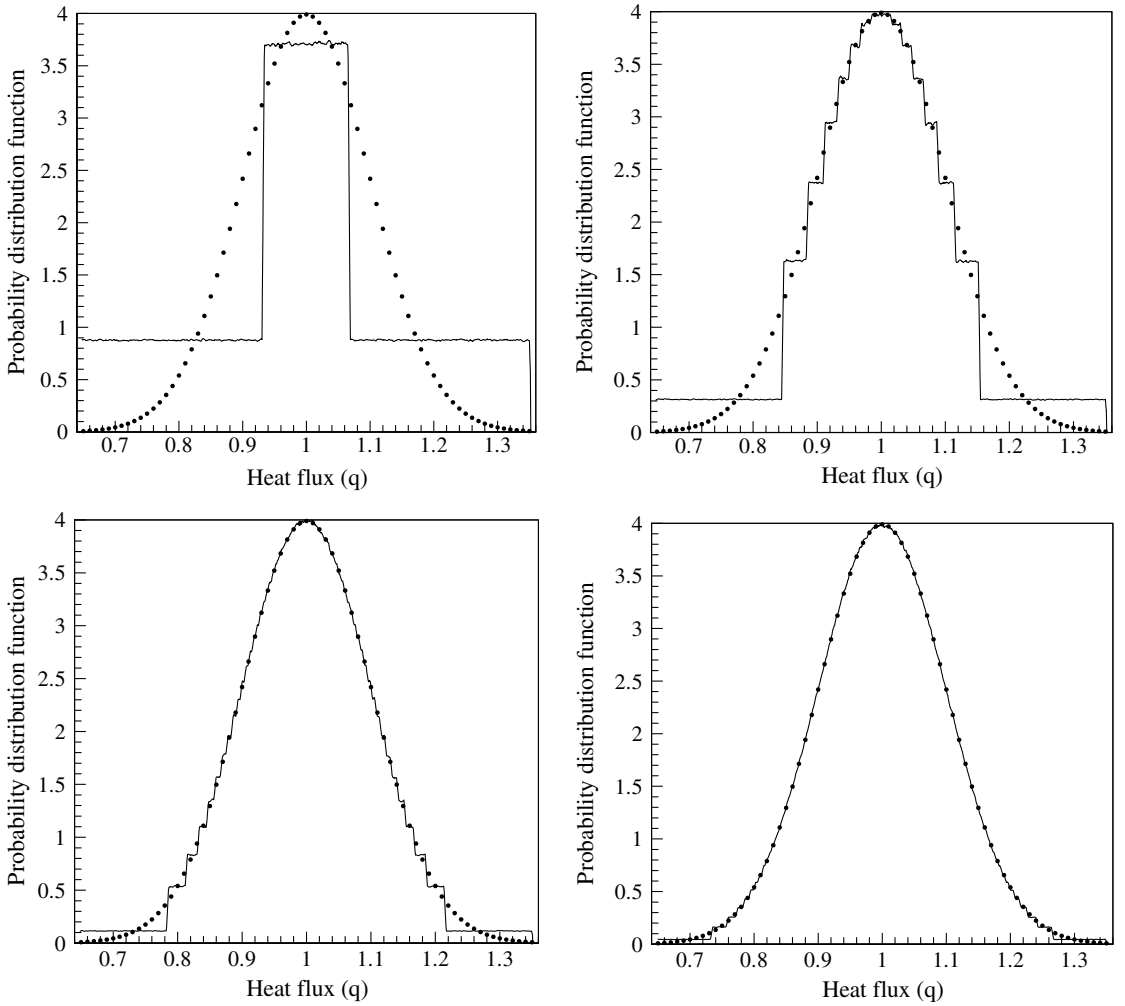

Fig. 14. Optimal PDF of stochastic heat flux computed with various depths of interpolation $(q \sim N(1.0, 0.1))$. Top row: Depth of interpolation 2 and 4. Bottom row, depth of interpolation 6 and 8. The filled circles represent the 'experimental' data and the solid lines represent the reconstructed solution.

q follows a normal distribution (unbounded support). As the depth of interpolation is increased, the PDF in both cases is captured very well. Even a low depth of interpolation of 4 ($n_q = 17$) results in a relatively accurate match of the main body of the PDF. As the interpolation depth is increased, the tails of the PDF in Fig. 14 are computed more accurately.

Invariance with the choice of representation: An analysis of the effect of choosing different representations to construct the stochastic heat flux is performed next (for case (a)). Four different representations of q based on different supports and associated PDFs are chosen as discussed earlier.

The four optimization problems were run with $n_q = 257$ collocation points to represent the domain of $\boldsymbol{\xi}$. Fig. 15 plots the PDF of the optimal stochastic heat flux computed from each of the four optimization problems. Notice the negligible difference between the computed PDFs.

Additional illustrations: Reconstructing PDFs of higher-dimensional random processes. In the previous examples, the input stochastic process that drove the direct problem to obtain the 'experimental' statistics was assumed to be a one-dimensional function (i.e. a univariate random variable). In the following sub-case, this assumption is relaxed and the ability of the optimization procedure to reconstruct the PDF of the stochastic heat flux irrespective of the dimensionality of the space in which the flux resides is showcased. As before, the PDF of the temperature at the sensor location is provided. This temperature is computed by first solving a direct problem with a known stochastic heat flux. The known stochastic heat flux is defined as follows:

$$q(\xi_1, \xi_2) = \sum_{i=1}^{n} a_i \xi_i, \quad \xi_i \in [0, 1]. \tag{39}$$

The ξ_i are independent uniform random variables. n is taken to be 2 and $(a_1, a_2) = (3.5, 2.5)$. The direct problem (to obtain the 'experimental' sensor location temperature) is solved using $n_q = 1537$ collocation points to represent the (2 stochastic dimensional) stochastic heat flux, $n_\alpha = 65$ collocation points to represent the stochastic diffusivity and using 80 elements to discretize the spatial domain.

The solution of the inverse problem to reconstruct the PDF of the stochastic heat flux is based on a representation using one uniform random variable i.e. $q(\boldsymbol{\xi}) = \sum_{i=1}^{n_q} q(\boldsymbol{\xi}_i) L_i(\boldsymbol{\xi})$, where $\boldsymbol{\xi} \in [0, 1]$. The optimization problem was run using different depths (2, 4, 6, 8) of interpolation (corresponding to increasing values of $n_q = 5, 17, 65, 257$). Fig. 16 plots the computed optimal PDF of the stochastic heat flux for the four different optimizations using different depths of interpolation. There is negligible difference between the PDFs generated using a depth of interpolation 6 and 8. For a depth of interpolation 8, the reconstructed PDF matches extremely well with the 'experimental' PDF.

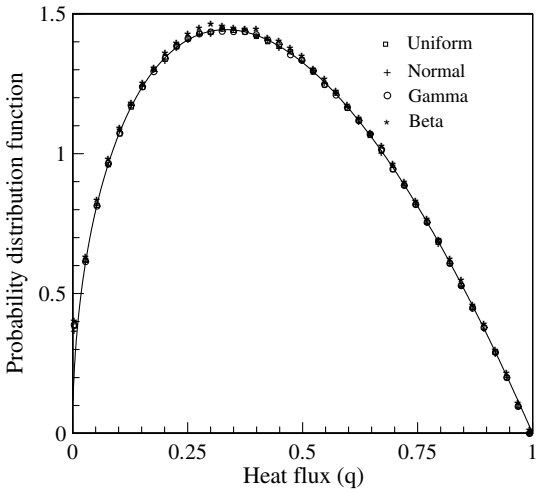

Fig. 15. Comparison of the performance of four different representations of the stochastic heat flux. The solid line is the 'experimental' PDF.

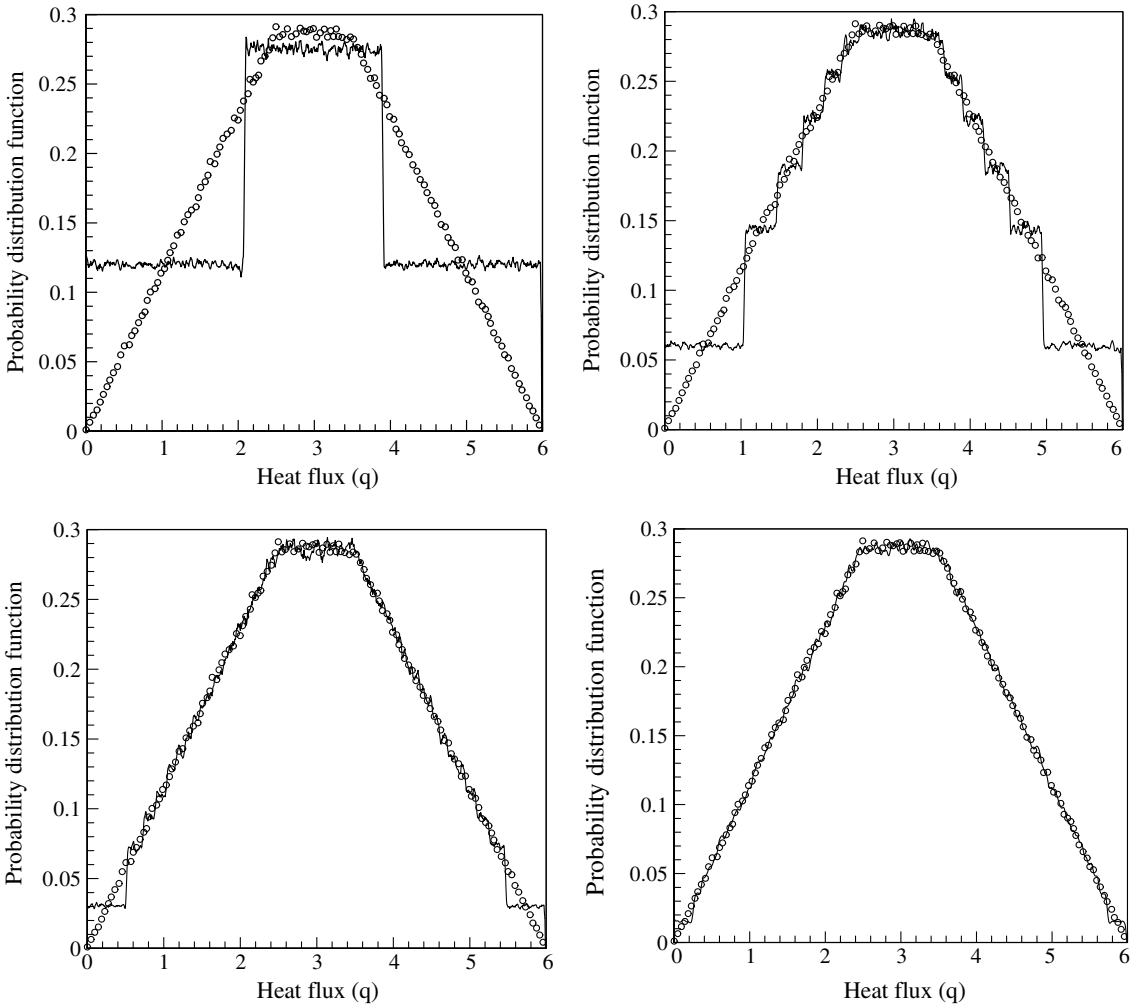

Fig. 16. Capturing random processes defined in higher dimensions: Convergence of the optimization procedure with increasing depth of interpolation. Comparison of the 'experimental' PDF (circles) with the reconstructed PDF (solid line): Top row, depth of interpolation 2 and 4. Bottom row, depth of interpolation 6 and 8.

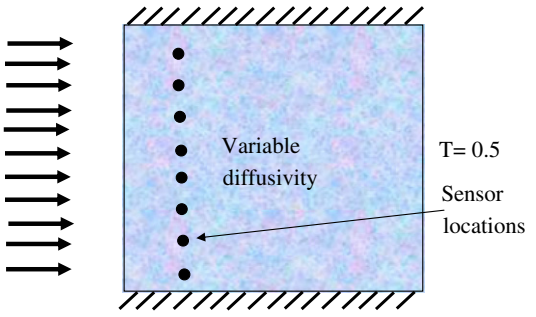

Fig. 17. Schematic of the 2D problem: The thermal diffusivity in the domain is defined by a known correlation kernel. The temperature is given at the sensor location shown by the dark circles. The problem is to compute the unknown stochastic heat flux on the left boundary.

4. Numerical examples

The following section contains two numerical examples. In the first example, a two-dimensional stochastic transient inverse problem is attempted while in the second example a three-dimensional steady-state stochastic design problem is considered.

4.1. 2D stochastic inverse problem

In the following numerical problem, we showcase the scalability of the proposed framework in solving large-scale stochastic inverse problems. The region of interest is a two-dimensional domain $\mathcal{D} \equiv [-0.5, 0.5] \times [-0.5, 0.5]$. The right boundary of the domain is maintained at a steady (deterministic) value of $\theta = 0.5$. At $t = 0$, the temperature in the interior of the domain is $\theta = 0$. The thermal diffusivity of the domain is uncertain and is assumed to be heterogeneously distributed over the domain (Fig. 17). The thermal diffusivity has a mean value of $\alpha = 10$. It is further assumed that the spatial variation in the thermal diffusivity follows an exponential correlation (i.e. $C(r_1, r_2) = \exp(-|r_1 - r_2|/b)$). The correlation length is set at $b = 10$. There are $s = 41$ equally spaced sensors inside the domain at a distance $d = 0.1$ from the left boundary. Each sensor collects data over the time interval $[0, 0.5]$. This data is given in terms of PDF of the temperature at each of these sensor location at 50 equally spaced time intervals in time range $[0, 0.5]$. The inverse problem can now be posed as the following:

> Identify the PDF of the heat flux on the left boundary in the time range $[0, 0.5]$ such that the experimental measurements are reconstructed.

Following the standard techniques in representing correlated spatial stochastic processes, the stochastic thermal diffusivity is expanded using a Karhunen–Loève expansion as follows:

$$\alpha(x, y, \omega_\alpha) = \alpha_{\text{mean}}(x, y) + \sum_{i=1}^{m} \sqrt{\lambda_i} f_i(x, y) \xi_i, \tag{40}$$

where λ_i and f_i are the eigenvalues and eigenvectors of the correlation kernel of the thermal diffusivity variation. Fig. 18 plots the first few eigenvalues of the correlation matrix. The first three eigenvalues represent about 96% of the variation. Correspondingly, the thermal diffusivity is represented using three random variables $\xi_i, i = 1, 2, 3$. Fig. 19 shows the eigenmodes corresponding to these random variables.

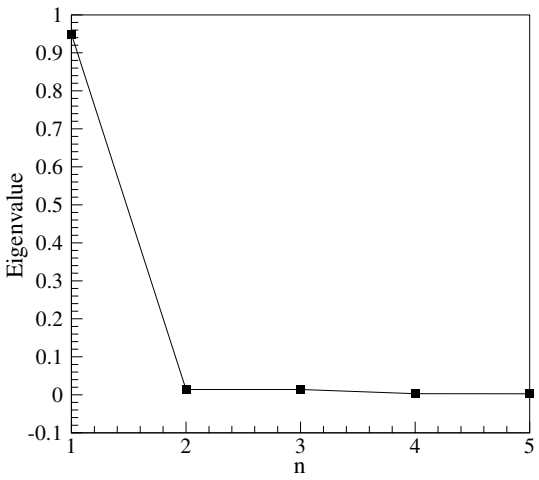

Fig. 18. The eigenvalues of the correlation kernel that represents the thermal diffusivity variation.

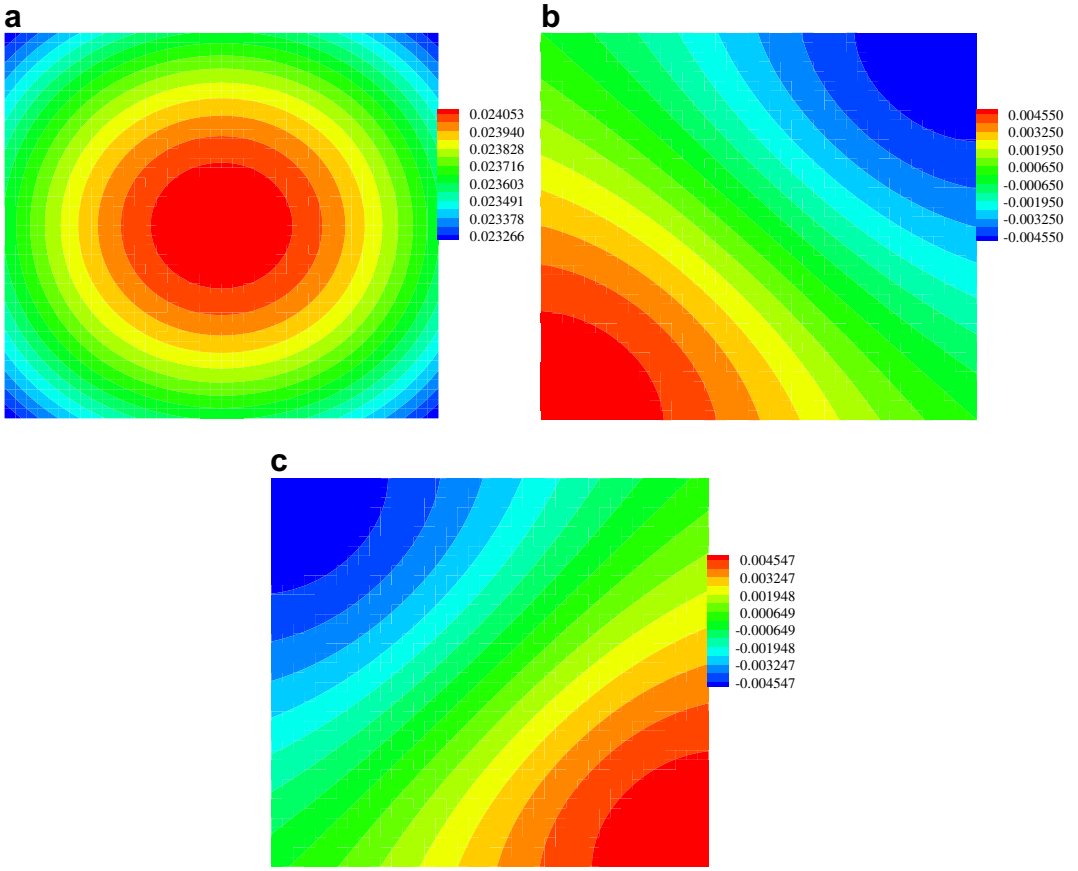

Fig. 19. The first three modes of the Karhunen–Loève expansion of the thermal conductivity. The first three modes represent ~95% of the variation of α.

4.1.1. Computing the 'experimental' PDF

A direct problem is solved using an assumed stochastic variation for the heat flux applied on the left boundary. The PDFs of the temperature variation at the sensor locations are computed. This serves as the 'experimental' data that is used to run the inverse problem. The direct problem is run using a correlated thermal flux. The spatial variation of the heat flux is assumed to have an exponential correlation, i.e. $C(y_1, y_2) = \exp(-|y_1 - y_2|/b_1)$, with $b_1 = 0.5$. This stochastic heat flux is expanded using a Karhunen–Loève expansion. The eigenvalues of this 'experimental' heat flux are plotted in Fig. 20. The first three eigenvalues represent about ~94% of the variation. The applied heat flux is represented using three random variables.

The heat flux applied has a mean value of 20 and is given a time-dependent damping term $\exp(-\beta t)$ with $\beta = 2.0$. The input heat flux is taken to be of the form:

$$q_{\exp}(y, t, \omega_q) = \mathrm{e}^{-\beta t}\left[20.0 + 5.0\sum_{i=1}^{3}\sqrt{\lambda_i}g_i(y)\xi_i\right]. \tag{41}$$

The direct problem to obtain the 'experimental' statistics is solved using a 80×80 quad element discretization of the physical domain using a time step of $\Delta t = 0.0025$. There are two sources of uncertainty in this problem: the uncertainty in the thermal diffusivity (represented in 3D stochastic space) and the uncertainty in the heat flux applied (represented in 3D stochastic space). The temperature consequently resides in the product space of these variations – a six-dimensional stochastic space. A sparse grid collocation strategy is used to compute the temperature variability at the sensor locations. The thermal diffusivity variation is represented

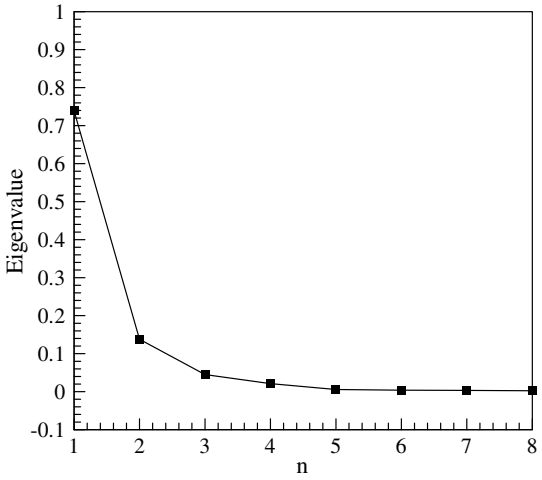

Fig. 20. The eigenvalues of the correlation kernel that represents the 'experimental' heat flux.

using a level 6 depth of interpolation (1073 points in 3D space). Similarly, 1073 points are used to represent the heat flux variation. The stochastic direct problem involves the solution of $1073 \times 1073 = 1.151 \times 10^6$ direct deterministic problems.

The PDF of the applied 'experimental' heat flux at different times at two locations on the boundary ($(x, y) = (-0.5, 0.0)$ and $(x, y) = (-0.5, -0.475)$) is given in Fig. 21. Notice that the initially diffuse PDF peaks and shifts towards a value of zero (i.e. the peak moves left) with increasing time because of the damping effect. This is clearly seen in Fig. 22 that plots the time variation of the mean flux.

The PDFs of the resultant 'experimental' temperature at two locations in the domain ($(x, y) = (-0.4, 0.0)$ and $(x, y) = (-0.4, -0.475)$) are given in Fig. 23. Notice that the initially diffuse PDF (at $t = 0.01$) slowly peaks and shifts towards zero with increasing time because of the damping effect. Interestingly, the effect of the uncertain thermal diffusivity is also seen in the bimodal structure of the resultant PDF.

4.1.2. The optimization problem: computational details and results

A 40×40 quad element discretization of the domain is utilized to solve the stochastic inverse problem. The time domain is discretized into $n_t = 50$ equal time steps. The inverse problem is to estimate the time evolution

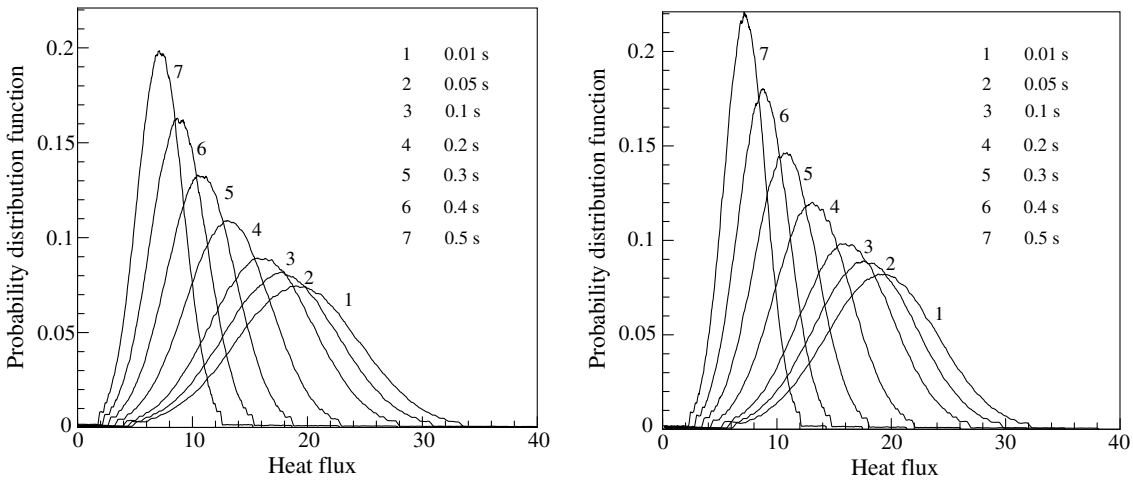

Fig. 21. The 'experimental' PDF of the heat flux at two boundary locations $y = 0.0$ and $y = -0.475$ at different times.

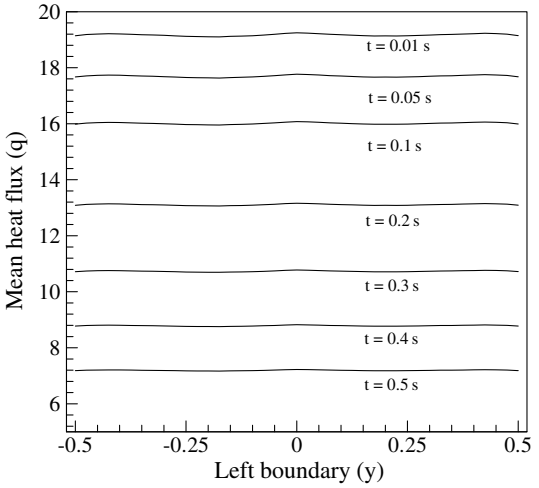

Fig. 22. The time variation of the mean 'experimental' heat flux applied.

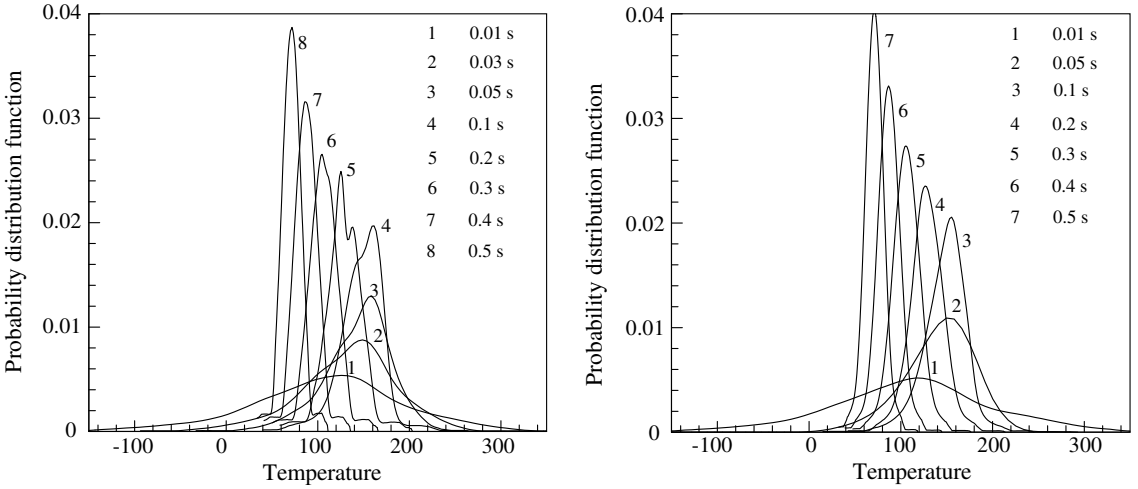

Fig. 23. The 'experimental' PDF of temperature at two sensor locations $y = 0.0$ and $y = -0.475$ at different times.

of the thermal flux at each nodal point on the left boundary. The total number of nodal points on the left vertical boundary is $n_y = 41$. The heat flux at each of these nodal points is assumed to be an independent random variable. Thus, the variation in the heat flux at each nodal point at each time instant has to be estimated. The total number of random variables that have to be estimated is equal to $n_y \times n_t = 2050$. The thermal diffusivity variation is represented using a level 5 depth of interpolation. This corresponds to using $n_\alpha = 441$ realization of the thermal diffusivity field. Following the collocation based representation rational used in the current work, each of the random flux is represented as

$$q(x,t,\xi) = \sum_{i=1}^{n_q} q(x,t,\xi_i) L_i(\xi). \tag{42}$$

Without loss of generality, we assume that the heat flux can be represented using one uniform random variable. That is, $\xi = U[0,1]$. A level 6 depth of interpolation is used to represent each random variable. This corresponds to $n_q = 65$. The total number of design variables is consequently $n_q \times n_y \times n_t = 133,250$.

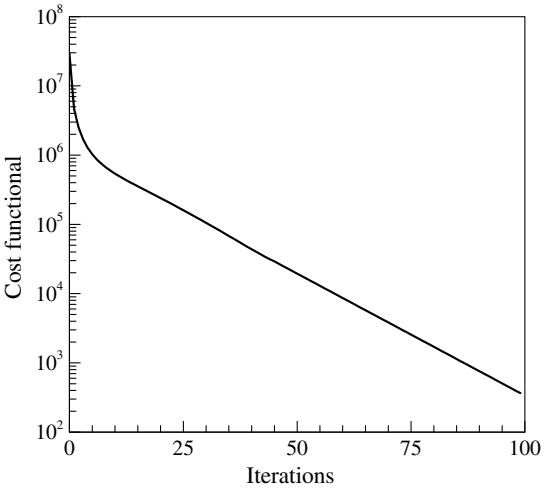

Fig. 24. Reduction in the cost functional.

The optimization problem requires the estimation of the sensitivity of the temperature at the sensor locations to perturbations to each of the design variables. A set of decoupled deterministic sensitivity problems were run to construct the stochastic temperature sensitivity. The number of such deterministic sensitivity problems run is $n_\alpha \times n_y \times n_t = 0.9 \times 10^6$. Each iteration of the optimization problem requires the solution of the stochastic forward problem. The stochastic forward problem is solved as a set of decoupled direct deterministic problems. The total number of such direct deterministic problems was $N_{\text{run}} = n_\alpha \times n_q = 28,665$. Each deterministic problem requires the solution on a 40×40 quad grid for n_t time steps. The number of degrees of freedom (DOF) in each deterministic problem is $N_{\text{det}} = (41)^2 \times n_t = 84,050$. The total number of DOF in *each* direct stochastic solve of the optimization algorithm is $N_{\text{stochastic}} = N_{\text{det}} \times N_{\text{run}} = 2.41 \times 10^9$. Thus, more than a *billion* DOF are solved at each iteration of the stochastic optimization problem.

The reduction in the cost functional with the number of iterations of the stochastic optimization is shown in Fig. 24. After the initial large rate of reduction in the cost functional, the rate of decrease of the cost functional is more or less constant over the course of about 100 iterations. The optimization problem was run using our in-house Linux super computing cluster. Forty nodes, corresponding to 160 processors were utilized for the current problem. Each optimization iteration of the problem took 74 min to complete.

Fig. 25 plots the time evolution of the PDF of the heat flux at two different locations on the boundary. Comparing this with Fig. 23 reveals that the stochastic heat fluxes are reconstructed very well though there is some pixellation near the tails of the PDF. This could be due to the lower level of resolution of the ends of the support of the heat flux considering the choice of the depth of interpolation of the random variable ($n_q = 65$).

To see if this pixellation disappears when a finer representation of the stochastic flux is used, another optimization was run using a depth 8 interpolation of the stochastic heat flux. This corresponds to $n_q = 257$ design variables to represent each unknown. We utilize the hierarchical stochastic optimization method (see Remark 3.8) to solve this larger optimization problem. That is, the solution of the previous optimization problem is used as the initial guess of this finer stochastic optimization problem. Fig. 26 plots the time evolution of the PDF of the heat flux at two different locations on the boundary. The pixellation near the tails of the PDF in Fig. 25 is smoothed out by using a higher-depth of interpolation of the stochastic space.

The difference between the exact and the reconstructed stochastic heat flux is defined in terms of the error, $e = \sum_{i=1}^{n_y} (\Upsilon^{-1}[q_{\text{ex}}(\boldsymbol{x}_i), u] - \Upsilon^{-1}[q_{\text{rc}}(\boldsymbol{x}_i), u])^2$, where $\Upsilon^{-1}(\cdot, u)$ is the corresponding inverse CDF of the heat flux at each nodal point on the left boundary and q_{ex} denotes the actual heat flux while q_{rc} denotes the reconstructed solution. Fig. 27 plots the error for the two optimization problems.

4.2. 3D design problem: design in the presence of topological uncertainty

We briefly state the problem of interest for the sake of completeness. We are given a micro-scale heat sink device, made up of a two-phase material of size 20 μm × 20 μm × 20 μm. The exact microstructure of this device is unknown. Some statistical descriptors of the microstructure can be extracted from the given 2D experimental image of the microstructure. A specific temperature profile has to be maintained at one end of the device. This desired temperature is specified as follows: The mean temperature should follow a specified profile, i.e. $T = -0.45 + \frac{z-a}{a}0.05$, where a is the half width of the device, $a = 10$ μm (see Fig. 28). The standard deviation of the temperature variation must be limited to 0.05 everywhere. Hence, assume that the required probability distribution of the temperature is given by $N(\mu, \sigma)$, where $\mu = -0.45 + \frac{z-a}{a}0.05$ and $\sigma = 0.05$. The design problem can be stated as follows: *Design the stochastic heat flux on the left wall, such that in the presence of topological uncertainties, the above specified thermal profile is maintained.* Insulated boundary conditions are enforced on the right boundary.

For clarity of presentation, we divide the solution into multiple sections. The first step is to represent the topological uncertainty that the limited information about the microstructure causes. We have recently developed a data-driven model reduction framework to construct viable, realistic low-dimensional stochastic models of microstructure variability given some limited information [33]. We do not intend to go into the details of the nonlinear model reduction technique though we briefly state the overall methodology for the sake of completeness. The interested reader is referred to [33] for a comprehensive treatment of the same.

4.2.1. Representing the topological uncertainty

The limited information about the microstructure results in topological uncertainty which translates to uncertainties in the thermal properties of the device. We first extract topological statistics from the experimental image provided. These statistics are then utilized to reconstruct a large set of 3D microstructures $\{\boldsymbol{x}_i\}, i = 1, \ldots, N$. The next step is then to construct the low-dimensional representation of the class of microstructures utilizing the samples $\{\boldsymbol{x}_i\}, i = 1, \ldots, N$.

Data extraction and sample set construction: We assume that we are given one experimental image of the underlying microstructure of the device. The image, shown in Fig. 29, is of a tungsten–silver composite [36]. This composite was produced by infiltrating a porous tungsten solid with molten silver. The first step is to extract the necessary statistical information from the experimental image. The image is cropped, deblurred and discretized. The volume fraction of silver is $p = 0.2$. The experimental two-point correlation $S_2(r)$ is extracted from the image. The normalized two-point correlation $(g(r) = \frac{S_2(r)-p^2}{p-p^2})$, is shown in Fig. 30.

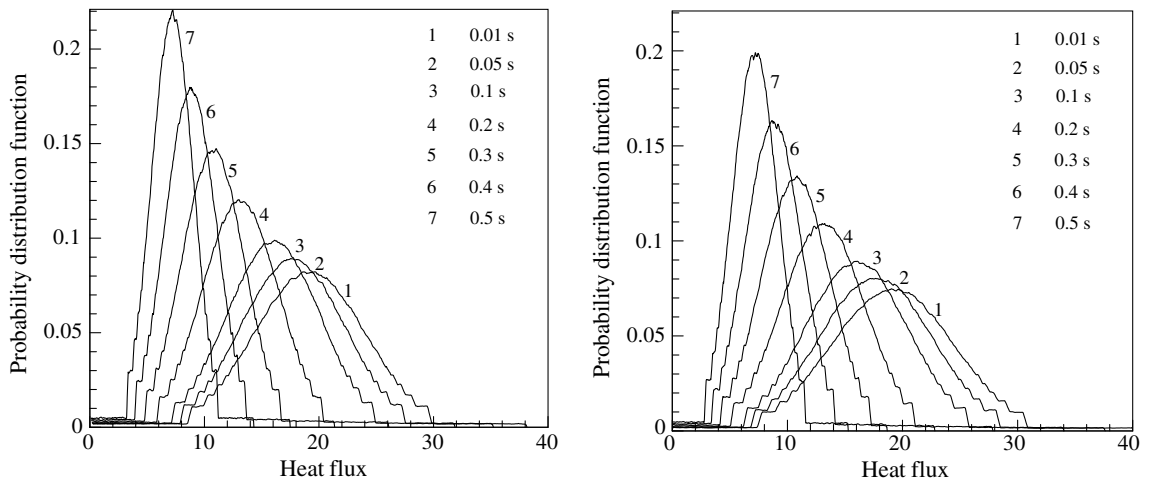

Fig. 25. The reconstructed PDF of heat flux at two boundary locations $y = 0.0$ and $y = -0.475$ (with depth of interpolation 6).

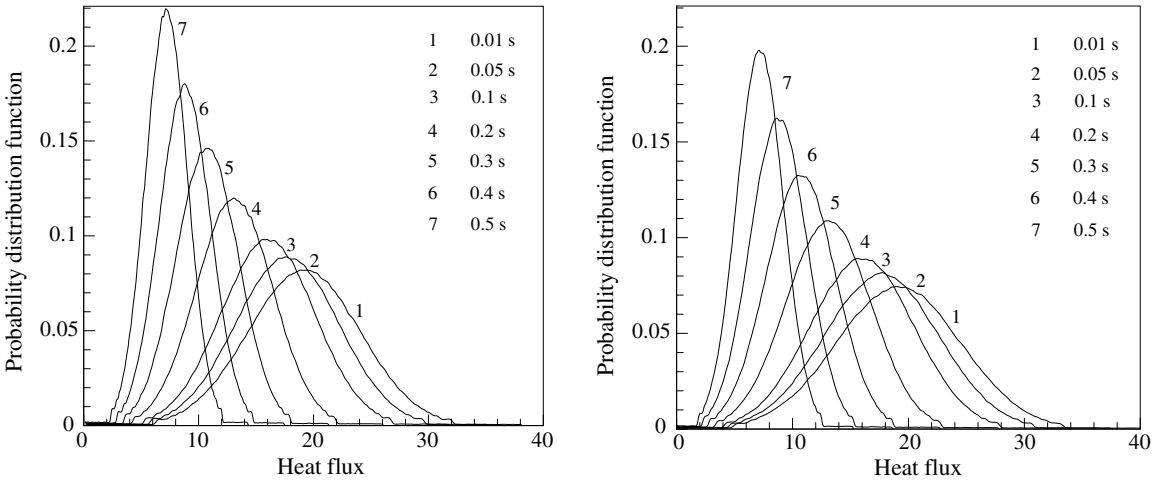

Fig. 26. The reconstructed PDF of heat flux at two boundary locations $y = 0.0$ and $y = -0.475$ at different times (with depth of interpolation 8).

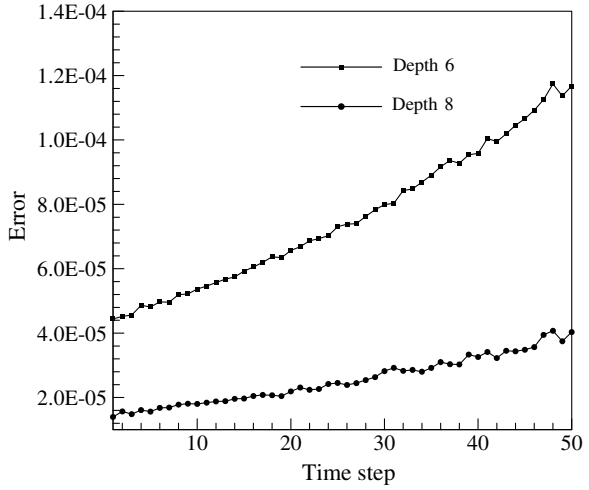

Fig. 27. Error between the actual solution and the reconstructed solution.

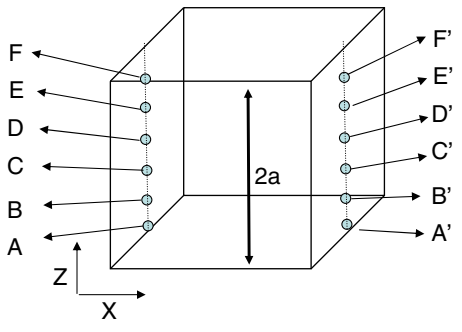

Fig. 28. Schematic of the 3D problem.

Fig. 29. Experimental image of a two-phase composite (from [36]).

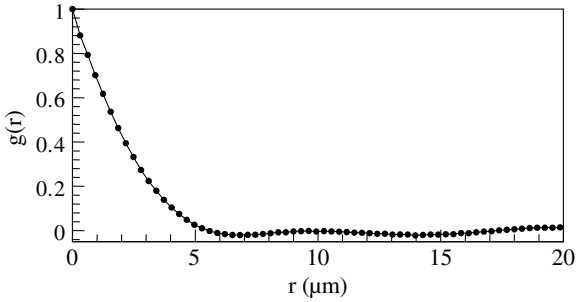

Fig. 30. The two-point correlation function.

The next step is to utilize these extracted statistical relations (volume fraction and two-point correlation) to reconstruct a class of 3D microstructures. We utilize a statistics based reconstruction procedure based on Gaussian random fields (GRF). Using this procedure, realizations of 3D microstructure were computed. Each microstructure consisted of $65 \times 65 \times 65$ pixels. This corresponds to a size of 20 μm × 20 μm × 20 μm. One realization of the 3D microstructure reconstructed using the GRF is shown in Fig. 31.

Nonlinear dimension reduction and construction of the topological model [33]: The GRF based reconstruction detailed above was used to generate a set of $N = 1000$ samples of two-phase microstructure. Each microstructure is represented as a $65 \times 65 \times 65$ pixel image. The three-point correlations of all of these samples are calculated. Based on the calculated S_3, the pairwise distance matrix $\mathbf{P}$ is computed. From this, the geodesic distance matrix $\mathbf{M}$ and the graph G [33] are computed. These are used to estimate the optimal dimensionality of the low-dimensional space. The dimensionality was estimated to be $d = 9$. Multi-dimensional scaling (MDS) is performed using the geodesic distance matrix $\mathbf{M}$. The nine largest eigenvalues and their corresponding eigenvectors are used to represent the input samples. The low-dimensional region $\mathcal{A}$ is constructed as the convex hull of these $N(= 1000)$ nine-dimensional points ξ_i. This region coupled with the suitable mappings $(\mathcal{G}_{\mathcal{A}}^{\mathcal{M}}(\xi))$ define the reduced-order stochastic input model $\mathcal{F} : \mathcal{A} \to \mathcal{M}$.

4.2.2. Representing the stochastic flux

Parametric representation of deterministic functions: Consider a deterministic heat flux $q(y,z)$, $(y,z) \in \partial\mathcal{D}_h$. There are many methodologies to parametrically represent $q(y,z)$ (wavelets, polynomials, Bézier surfaces among others). Consider a Bézier curve representation of the heat flux $q(y,z)$. A Bézier curve is a parametric

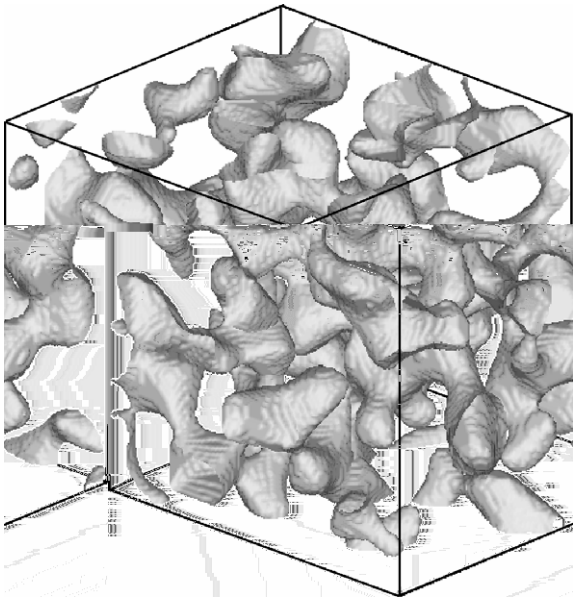

Fig. 31. One instance (realization) of the two-phase microstructure.

curve, used in representing surfaces in various CAD and engineering practices. A given Bézier curve of order (m_x, m_y) is defined by a set of $(m_x + 1) \times (m_y + 1)$ control points $P_{i,j}$. It maps the region $[0,1]^2$ into a smooth-continuous surface. A Bézier representation of the heat flux, $q(y,z)$, is defined (in terms of the normalized coordinates, u, v) as

$$q(u,v) = \sum_{i=0}^{m_x} \sum_{j=0}^{m_y} P_{i,j} B_i^{m_x}(u) B_j^{m_y}(v), \tag{43}$$

where

$$B_i^{m_x}(u) = \frac{m_x!}{i!(m_x - i)!} u^i (1-u)^{m_x - i}. \tag{44}$$

Parametric representation of stochastic functions: A deterministic function $q(y,z)$ is represented in terms of the $(m_x + 1) \times (m_y + 1)$ deterministic parameters $P_{i,j}$. Note that the parameters $P_{i,j}$ are completely independent of each other. Now consider the stochastic function $q(y,z,\boldsymbol{\xi})$. A general parametric representation of the stochastic function can be constructed by simply considering the deterministic parameters $P_{i,j}$ to now be independent random variables. The stochastic flux can be written as

$$q(u,v,\boldsymbol{\xi}) = \sum_{i=0}^{m_x} \sum_{j=0}^{m_y} P_{i,j}(\xi_{(i,j)}) B_i^{m_x}(u) B_j^{m_y}(v), \tag{45}$$

where $\boldsymbol{\xi} = \{\xi_{(i,j)}\}, i = 1, \ldots, m_x, j = 1, \ldots, m_y$. Following the discussion in Section 3, we represent each stochastic $P_{i,j}$ using a finite number, n_q of deterministic realizations as $P_{i,j}(\xi_{(i,j)}) = \sum_{k=1}^{n_q} P_{i,j}^k L_k(\xi_{(i,j)})$. The stochastic heat flux is then represented as

$$q(u,v,\boldsymbol{\xi}) = \sum_{k=1}^{n_q} \left[\sum_{i=0}^{m_x} \sum_{j=0}^{m_y} P_{i,j}^k B_i^{m_x}(u) B_j^{m_y}(v) \right] L_k(\boldsymbol{\xi}). \tag{46}$$

The stochastic flux $q(y,z,\boldsymbol{\xi})$ is represented by the $(m_x + 1) \times (m_y + 1) \times n_q$ scalars $\{P_{i,j}^k\}$.

4.2.3. The optimization problem: computational details and results

The computational domain is normalized to $[0,1]^3$ and a $64 \times 64 \times 64$ hexahedral discretization of the domain is used for all the computations. Each node in the mesh corresponds to a phase – silver or tungsten. The thermal conductivity assigned to each node is normalized to 1.0 or $\alpha_{\text{silver}}/\alpha_{\text{tungsten}}$.

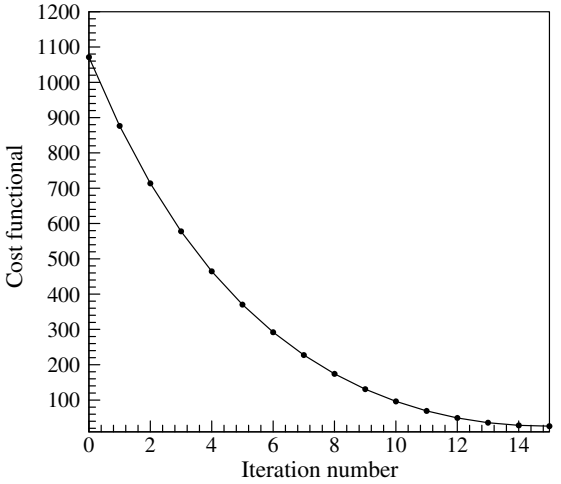

Fig. 32. Reduction in the cost functional during the optimization procedure.

The two stochastic spaces, representing the topological variation and the heat flux variation are represented using a collocation based representation, The topological variations is represented using 1177 sampling points ($n_\alpha = 1177$). This corresponds to a level 3 interpolation of the 9D stochastic space. The heat flux variation is represented using 65 sampling points ($n_q = 65$). This corresponds to a level 6 interpolation. The total number of design variables are $n_q \times m_x \times m_y = 65 \times 5 \times 5 = 1625$. Each direct problem involves the solution of $n_q \times n_\alpha = 76,505$ deterministic PDEs. The sensitivity temperature is computed by the solution of $m_x \times m_y \times n_\alpha = 29,425$ deterministic continuum sensitivity equations. Each deterministic equation is solved in a $n_{\text{dof}} = 65 \times 65 \times 65$ hexahedral grid. The total number of unknowns solved for in each iteration is $n_{\text{dof}} \times n_q \times n_\alpha \sim 2.1 \times 10^{10}$. Each iteration took about 9 h to complete on 46 nodes of our in-house Linux based super computing cluster. Fig. 32 plots the reduction in the cost functional as the number of iterations.

Fig. 33 shows some first- and second-order statistics of the designed heat flux. Fig. 33 (top left) plots the (optimal) mean (expectation of the) heat flux that should be applied to achieve the desired thermal profile. Fig. 33 (top right) plots the bounds on the heat flux. These bounds are computed as +/− one standard deviation from the mean heat flux. The lower- and upper-bounds on the heat flux are plotted again for clarity in Fig. 33 (bottom left, right). Notice that the bounding curves become almost identical towards the top half of the device (increasing z). This suggests that the range of the PDF of the heat flux applied at points on the top half of the device ($z > 10$ μm) becomes quite small.

The PDF of the optimal heat flux to be applied at six points (with increasing height) lying on the mid-plane of the left boundary are computed. Denote these points as Pt A–Pt F (see Fig. 28). These points have constant $(x, y) = (0, 9.375)$μm coordinates. The z coordinates of the points are $z = 0, 3.125, 6.24, 9.375, 12.5, 18.75$ μm, respectively. The PDF of the corresponding heat fluxes are plotted in Fig. 34. Notice the gradual shifting of the peak of the PDF to the right. More interestingly, notice that the designed PDF becomes narrower with increasing height. What does this mean physically? Our design criterion is to maintain a specific temperature profile (within a given tolerance) at the right wall in the presence of topological uncertainties. As stated before, the thermal variations on the right boundary are a consequence of the combined effect of the variations in the underlying topology as well as variations in the applied heat flux. Now, an almost deterministic heat flux in some regions (that is, when the PDF of the heat flux in some regions becomes very peaked) signifies that the underlying topological uncertainties by themselves cause such large variation in the thermal profile that there can only be a very small allowance for variation in the designed heat flux, if the design criterion have to be satisfied.

Several significant conclusions about the framework can be drawn from this observation: (1) the design framework automatically accounts for the effects of multiple sources of uncertainties and (2) the stochastic optimization/stochastic sensitivity framework can be naturally used to quantitatively predict *which source of uncertainty* causes large variabilities in the measured/observed dependant variable. This can have major

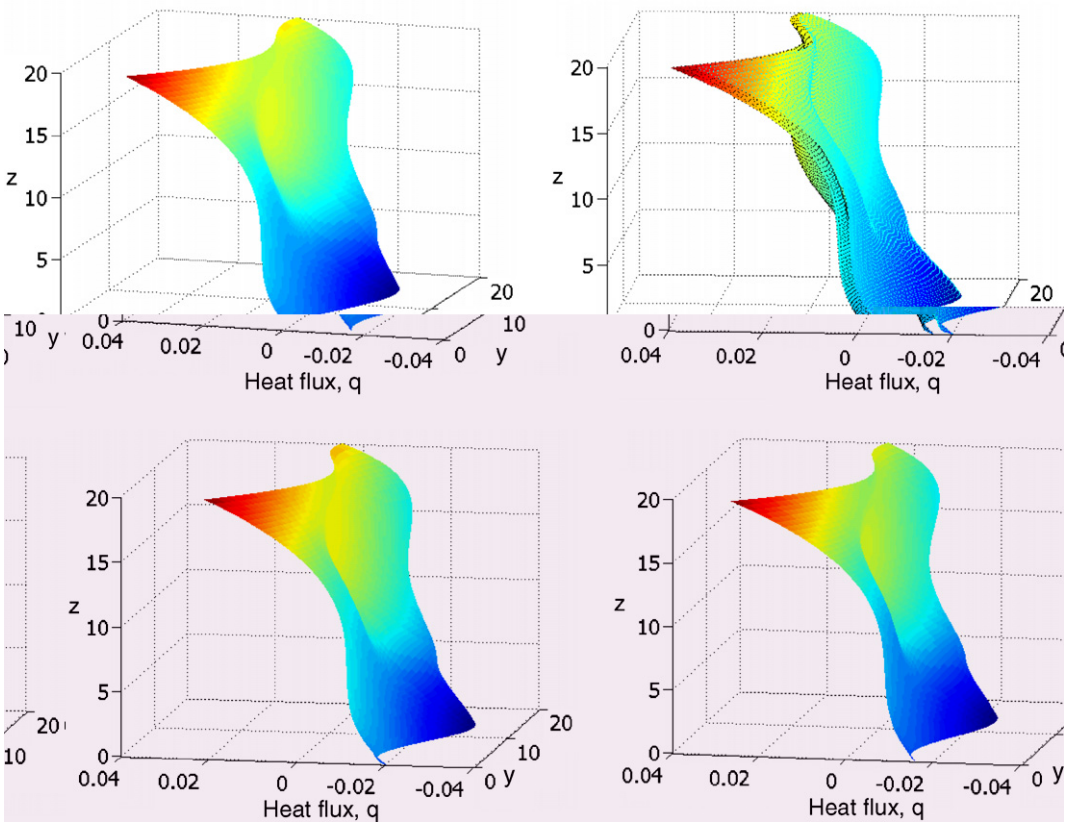

Fig. 33. The optimal flux ($n_q = 65$, and (5,5) Bézier parameters): Top left: The mean value of the designed flux. Top right: The bounding heat flux surfaces (mean +/− one standard deviation). The higher and lower bounding surfaces are re-plotted separately in the bottom row for clarity.

ramifications when there are multiple sources of uncertainties and performing experiments to reduce all these uncertainties is not feasible.

To see if the designed heat flux satisfies the design criterion, the PDFs of temperature at six points on the right boundary are computed. These points (denoted as Pt A′–Pt F′) lie on the mid-plane of the right boundary (see Fig. 28). These points have constant $(x, y) = (20, 9.375)\mu m$ coordinates. The z-coordinates of the points are $z = 0, 3.125, 6.24, 9.375, 12.5, 18.75$ μm, respectively. The PDFs of the corresponding heat fluxes are plotted in Fig. 35. Note the steady decrease in the mean value of the PDFs as the corresponding height of the sampling point is increased. This variation closely follows the desired thermal profile (linearly varying thermal profile). Also note that the standard deviation of these PDFs are around 0.05 which is the desired deviation. This can be clearly seen in Fig. 36 which plots the standard deviation of temperature at all the nodal points on the right boundary. Notice that the desired standard deviation of 0.05 is achieved in almost the whole of the region. Fig. 37 plots the corresponding mean temperature at all the nodal points on the right boundary. There is a linear variation of the temperature along the z-direction. The effect of the topological uncertainties is to make the mean temperature profile highly anisotropic and non-uniform.

4.2.4. Convergence analysis

We conclude this example with a discussion of the convergence of the stochastic design solution with respect to two quantities, namely

- Convergence with respect to the approximation of the design space.
- Convergence with respect to the number of Bézier control points.

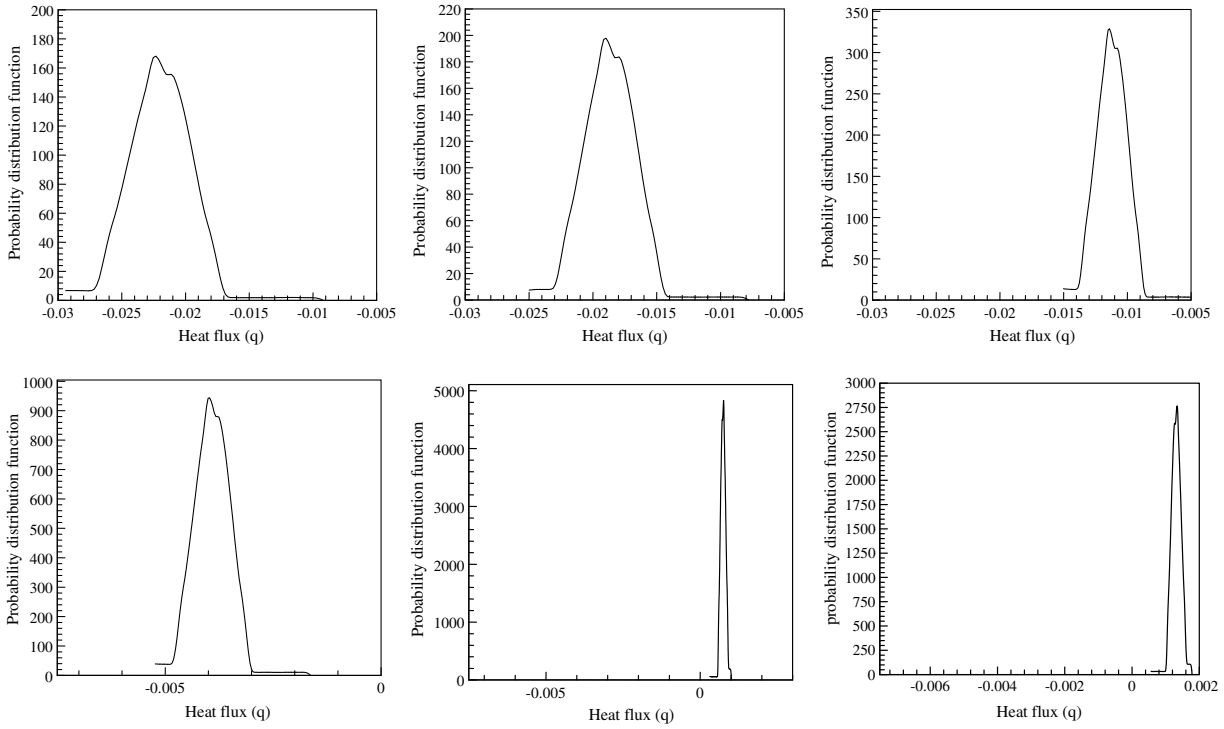

Fig. 34. The PDF of the optimal heat flux at six locations on the left boundary (on the mid-plane with $z = 0, 3.125, 6.24, 9.375, 12.5, 18.75$ μm, respectively). Note the gradual shift in the peak of the PDF to the right.

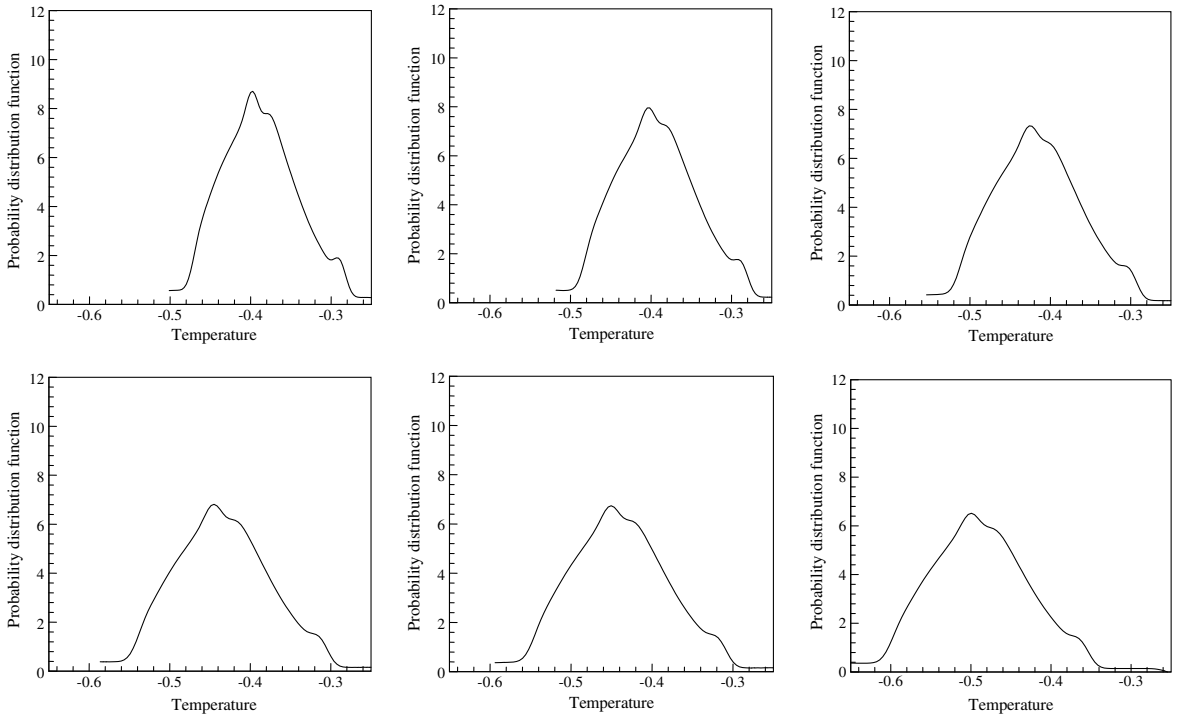

Fig. 35. The PDF of the temperature at six locations on the right boundary (on the mid -plane with $z = 0, 3.125, 6.24, 9.375, 12.5, 18.75$ μm, respectively). Note the gradual shift in the peak of the PDF to the left.

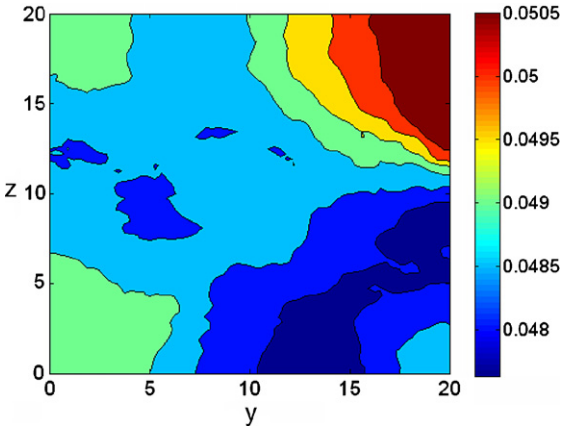

Fig. 36. Standard deviation of the temperature on the right boundary.

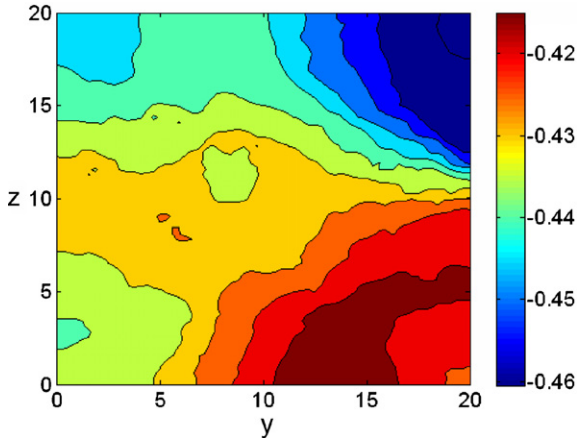

Fig. 37. Mean temperature on the right boundary.

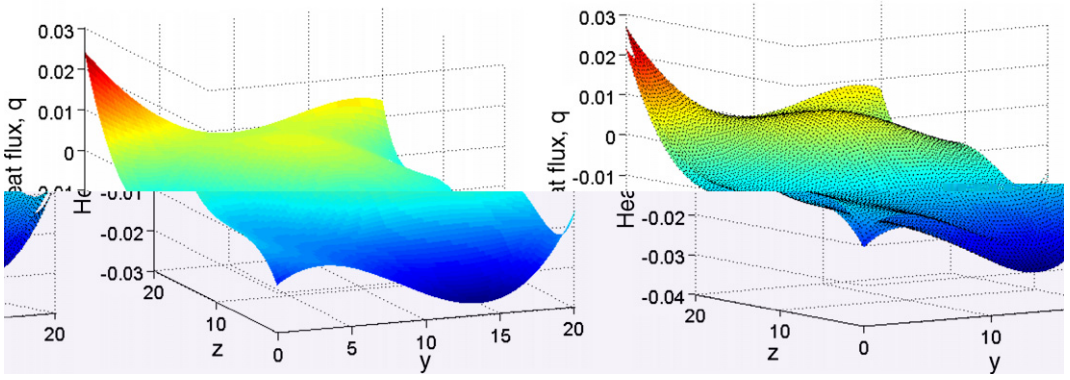

Fig. 38. The optimal flux with a finer discretization of the stochastic space ($n_q = 127$, and (5,5) Bézier parameters): the figure on the left plots the mean heat flux. The figure on the right plots the bounding heat flux (mean +/− one standard deviation).

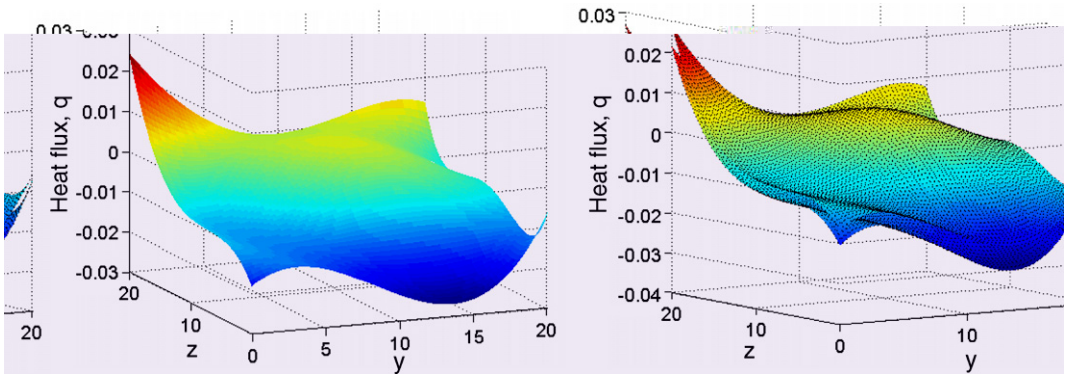

Fig. 39. The optimal flux with a finer Bézier discretization of the heat flux surface ($n_q = 65$, and (9,9) Bézier parameters): the figure on the left plots the mean heat flux. The figure on the right plots the bounding heat flux (mean +/− one standard deviation).

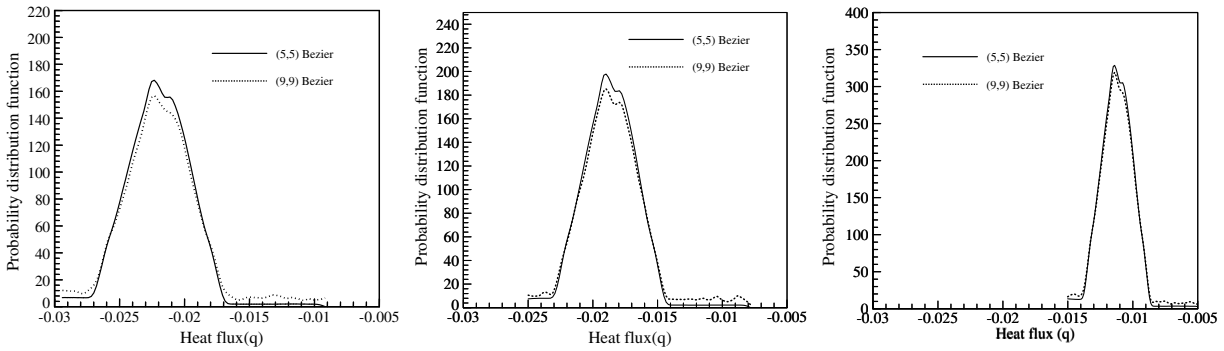

Fig. 40. Comparison of the reconstructed PDFs at three points on the left boundary (Pt A, Pt B, Pt C).

In the first convergence study, the representation of the designed stochastic space Ω_q is changed. Instead of utilizing $n_q = 65$ collocation points (Fig. 33), $n_q = 127$ collocation points are utilized. The optimization problem now is used to design $n_q \times m_x \times m_y = 127 \times 5 \times 5 = 3175$ variables. Fig. 38 plots the corresponding optimal flux. The difference between the two fluxes is of the order of $< 0.1\%$ of the $n_q = 65$ heat flux. The difference is defined as $e = \sum_{i=1}^{nno} (\Upsilon^{-1}[q_1(\boldsymbol{x}_i), u] - \Upsilon^{-1}[q_2(\boldsymbol{x}_i), u])^2$, where nno is the number of nodes on the right wall, $\Upsilon^{-1}(\cdot, u)$ is the corresponding inverse CDF of the designed heat flux at each nodal point on the left boundary and q_1 denotes the solution with $n_q = 65$, while q_2 denotes the solution with $n_q = 127$.

Since we utilize Bézier surfaces to parameterize each of these n_q deterministic heat fluxes, in the second convergence study, the parametric representation (i.e. number of Bézier control points) of the heat flux is refined. Instead of utilizing $(m_x, m_y) = (5, 5)$ control points, $(m_x, m_y) = (9, 9)$ control points are utilized. The optimization problem now is used to design $n_q \times m_x \times m_y = 65 \times 9 \times 9 = 5265$ variables. Fig. 39 plots the corresponding optimal flux. The difference between the two fluxes is of the order of $< 0.7\%$ of the $(m_x, m_y) = (5, 5)$ heat flux representation.

A comparison of the reconstructed PDF of heat flux at three points (Pt A, Pt B and Pt C defined earlier) is shown in Fig. 40.

5. Summary

The underlying idea developed in the present paper is a scalable methodology that provides the ability to perform design/estimation in the presence of multiple sources of uncertainties while relying purely on deterministic simulators. The reliance on purely deterministic simulators is accomplished through constructing the

solution of the stochastic direct problem using sparse grid interpolation strategies. This effectively removes the necessity for overhauling and rewriting of complex deterministic codes to use them in a stochastic setting.

Several issues related to existence and uniqueness of stochastic parabolic and elliptic equations are considered and discussed in the present work. A physically motivated framework (based on arguments of measurability/observability of random processes) is used to pose the stochastic inverse problem. A novel stochastic sensitivity technique is developed that effectively results in the construction of the stochastic sensitivity solution via repeated calls to deterministic sensitivity/direct problems. By using the sparse grid interpolation formulation, the stochastic optimization problem is converted to a deterministic optimization problem in a larger-dimensional space. This naturally allows us to utilize mature deterministic optimization algorithms to solve the stochastic optimization. We have also alluded to a promising extension to this stochastic optimization technique – a hierarchical stochastic optimization algorithm which seems to provide significant computational gains.

We showcased the effectiveness and scalability of the developed framework by solving realistic inverse and design problems in two- and three-dimensions. In the first example, we showcased the ability of the framework to solve *inference problems*. That is, given some sensor data, the framework was utilized to reconstruct/estimate the heat flux that resulted in that sensor measurement. This can naturally incorporate the problem of estimation in the presence of sensor error. Similar to Bayesian inference, the sensor error can be assumed to be an additive noise factor with some distribution. This distribution can be the input data to the stochastic optimization framework, which reconstructs the PDF/moments of the heat flux that result in the given input distribution. However, it has to be noted that Bayesian techniques to deal with measurement data provided in the form of a PDF have not yet been explored. In the second example, we showcased the ability of the framework to solve *design problems*. That is, given a desired distribution (via PDF or moments) of the dependant variable (temperature), the distribution (PDF/moments) of the heat flux that result in the desired temperature distribution is constructed. This framework naturally accounts for information in terms of single measurements (with assumed additive noise), statistics of the measurements and as PDF of the measurement (or desired variability).

We are currently investigating extensions/improvements of the proposed stochastic optimization framework along with the application of this framework to the robust design of materials processes as well as stochastic estimation in geological engineering.

Acknowledgements

This research was supported by the Computational Mathematics program of AFOSR (grant F49620-00-1-0373). The computing was conducted using the resources of the Cornell Center for Advanced Computing.

References

[1] R.G. Ghanem, P.D. Spanos, Stochastic Finite Elements: A Spectral Approach, Dover Publications, 1991.
[2] R. Ghanem, Probabilistic characterization of transport in heterogeneous porous media, Comput. Methods Appl. Mech. Eng. 158 (1998) 199–220.
[3] R. Ghanem, A. Sarkar, Mid-frequency structural dynamics with parameter uncertainty, Comput. Methods Appl. Mech. Eng. 191 (2002) 5499–5513.
[4] D. Xiu, G.E. Karniadakis, Modeling uncertainty in steady-state diffusion problems via generalized polynomial chaos, Comput. Methods Appl. Mech. Eng. 191 (2002) 4927–4948.
[5] D. Xiu, G.E. Karniadakis, Modeling uncertainty in flow simulations via generalized polynomial chaos, J. Comput. Phys. 187 (2003) 137–167.
[6] I. Babuska, R. Tempone, G.E. Zouraris, Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation, Comput. Methods Appl. Mech. Eng. 194 (2005) 1251–1294.
[7] I. Babuska, R. Tempone, G.E. Zouraris, Galerkin finite elements approximation of stochastic finite elements, SIAM J. Numer. Anal. 42 (2004) 800–825.
[8] D. Xiu, D. Lucor, C.-H. Su, G.E. Karniadakis, Performance evaluation of generalized polynomial chaos, in: International Conference on Computational Science, Lecture Notes in Computer Science, vol. 2660, Springer, 2003, pp. 346–354.
[9] B. Ganapathysubramanian, N. Zabaras, Sparse grid collocation schemes for stochastic natural convection problems, J. Comput. Phys. 225 (2007) 652–685.

[10] D. Xiu, J.S. Hesthaven, High-order collocation methods for the differential equation with random inputs, SIAM J. Sci. Comput. 27 (2005) 1118–1139.
[11] D. Xiu, Efficient collocational approach for parametric uncertainty analysis, Commun. Comput. Phys. 2 (2) (2007) 293–309.
[12] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, J. Comput. Phys. 209 (2005) 617–642.
[13] B. Velamur Asokan, N. Zabaras, Using stochastic analysis to capture unstable equilibrium in natural convection, J. Comput. Phys. 208 (2005) 134–153.
[14] I. Babuska, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, ICES Report 05-47, 2005.
[15] F. Nobile, R. Tempone, C.G. Webster, A sparse grid stochastic collocation method for elliptic partial differential equations with random input data, preprint.
[16] T. Gerstner, M. Griebel, Numerical integration using sparse grids, Numer. Algor. 18 (1998) 209–232.
[17] V. Barthelmann, E. Novak, K. Ritter, High-dimensional polynomial interpolation on sparse grids, Adv. Comput. Math. 12 (2000) 273–288.
[18] A. Klimke, Uncertainty modeling using fuzzy arithmetic and sparse grids, Ph.D. Thesis, Universitt Stuttgart, Shaker Verlag, Aachen, 2006.
[19] O.M. Alifanov, Inverse Heat Transfer Problems, Springer-Verlag, Berlin, 1994.
[20] A. Kirsch, An Introduction to the Mathematical Theory of Inverse Problems, Springer, 1996.
[21] J. Kaipio, E. Somersalo, Statistical and Computational Inverse Problems, Springer, New York, 2005.
[22] J. Wang, N. Zabaras, Hierarchical Bayesian models for inverse problems in heat conduction, Inverse Probl. 21 (2005) 183–206.
[23] J. Wang, N. Zabaras, A Bayesian inference approach to the stochastic inverse heat conduction problem, Int. J. Heat Mass Transfer 47 (2004) 3927–3941.
[24] J. Wang, N. Zabaras, Using Bayesian statistics in the estimation of heat source in radiation, Int. J. Heat Mass Transfer 48 (2005) 15–29.
[25] J. Wang, N. Zabaras, A Markov random field model to contamination source identification in porous media flow, Int. J. Heat Mass Transfer 49 (2006) 939–950.
[26] V.A. Badrinarayanan, N. Zabaras, Stochastic inverse heat conduction using a spectral approach, Int. J. Numer. Methods Eng. 60 (2004) 1569–1593.
[27] M. Loève, Probability Theory, fourth ed., Springer-Verlag, Berlin, 1977.
[28] Ch. Schwab, R.A. Todur, Sparse finite elements for stochastic elliptic problems – higher moments, Computing 71 (1) (2003) 43–63.
[29] Ch. Schwab, R.A. Todur, Sparse finite elements for elliptic problems with stochastic loading, Numer. Math. 95 (2001) 707–734.
[30] T. Petersdorff, Ch. Schwab, Sparse finite element methods for operator equations with stochastic data, Appl. Math. 51 (2006) 145–180.
[31] D.G. Aronson, Non-negative solutions of linear parabolic equations, Ann. Scuola Norm. Sup. Pisa 22 (1968) 607–694.
[32] N.A. Eklund, Existence and representation of solutions of parabolic equations, Proc. Am. Math. Soc. 47 (1975) 137–142.
[33] B. Ganapathysubramanian, N. Zabaras, A non-linear dimension reduction methodology for generating data-driven stochastic input models, J. Comput. Phys., submitted for publication.
[34] C. Desceliers, R. Ghanem, C. Soize, Maximum likelihood representation of stochastic chaos representations from experimental data, Int. J. Numer. Methods Eng. 66 (2006) 978–1001.
[35] S.A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Soviet Math. 4 (1963) 240–243.
[36] S. Umekawa, R. Kotfila, O.D. Sherby, Elastic properties of a tungsten–silver composite above and below the melting point of silver, J. Mech. Phys. Solids 13 (1965) 229–230.