

## INVERSE PROBLEMS IN HEAT TRANSFER

Nicholas Zabaras

Materials Process Design and Control Laboratory

Sibley School of Mechanical and Aerospace Engineering

Cornell University

Ithaca, New York, USA

### Contents

17.1 Introduction . . . . .	2
17.2 THE INVERSE HEAT-CONDUCTION PROBLEM - A SPECTRAL STOCHASTIC APPROACH . . . . .	8
17.2.1 Introduction: Representation of random variables . . . . .	9
17.2.2 The stochastic inverse heat-conduction problem (SIHCP): Problem definition . . . . .	11
17.2.3 Adjoint problem and gradient calculations . . . . .	15
17.2.4 Numerical Implementation . . . . .	17
17.2.5 Numerical results with the spectral approach . . . . .	21
17.2.6 Further developments on the use of spectral stochastic methods in inverse modeling . . . . .	25
17.3A BAYESIAN APPROACH TO THE INVERSE HEAT-CONDUCTION PROBLEM . . . . .	25
17.3.1 Revisiting the IHCP . . . . .	26
17.3.2 Bayesian inverse formulation . . . . .	29
17.3.3 MCMC sampler . . . . .	31
17.3.4 Numerical results with a Bayesian approach . . . . .	34
17.3.5 Further work on using Bayesian inference for inverse thermal problems . . . . .	37

## 17.1 Introduction

The term *direct problems* is usually referred to problems in mathematical physics that are defined by one or more coupled integral, partial, or ordinary differential equations (related to conservation principles), complete description of the coefficients in these equations (usually related to material data), and “sufficient” boundary and initial conditions for each of the main fields (such as temperature and velocity) [1, 2]. What defines ‘sufficient’ boundary conditions varies from the type of problem at hand. For example, in a transient heat-conduction problem, temperature or flux boundary conditions need to be applied to each part of the boundary [3]. Such conditions, in addition to “well defined” thermal conductivity, specific heat, density, and initial temperature, define a direct heat-conduction problem. These direct problems are often called “forward problems”, implying that their solution can be computed from one time step to the next. Direct problems are characterized by the existence of a unique solution that is stable to perturbations in the given data (material properties, boundary and initial conditions) assuming that these data obey certain “regularity conditions”. It is customary to refer to such direct problems as *well-posed* problems [4].

If any of the conditions needed to define a direct problem are unknown, an *inverse problem* results. For example, an inverse problem will result if insufficient boundary conditions are provided or if the initial conditions are not known or when a source term is not known or when some of the material data are not given. To allow a solution to such inverse problems, additional conditions will obviously need to be provided. Such conditions may, for example, include partial information of the unknown fields (e.g., temperature) resulting from sensor (experimental) data at distinct points in the domain and time. Hadamard classified inverse problems as *ill-posed* in the sense that conditions of existence and uniqueness of solution are not necessarily satisfied and that the solution may be unstable to perturbations in input data [5].

In many respects, inverse problems driven by discrete measurement data (or desired data) are much more practical from engineering point of view than direct analysis problems. For example, many design problems where a design objective needs to be specified with a proper choice of input conditions (design variables) can be posed as inverse problems [6]. Similarly, using experimental data from sensor measurements to compute unknown boundary conditions or source terms constitutes an inverse problem [7]. Estimation of material properties (e.g., conductivity) given indirect measurements of a field (e.g., temperature) depending on these properties is another inverse (parameter estimation) problem [8].

Inverse problems have practical implications in thermal transport systems which involve conduction, convection and radiation. In thermal radiation [9], for example, identifying the distribution of the radiation source has been stimulated by a wide range of applications, including thermal control in space technology, combustion, high-temperature forming and coating technology, solar energy utilization, high-temperature engines, and furnace technology [10, 11].

Inverse problems of heat-conduction, or IHCPs, can be subdivided into three categories: boundary problems, retrospective problems, and identification problems [12]. Here, emphasis is given to boundary heat-flux reconstruction in a conducting solid given temperature measurements at various points within the solid. This is by far the most popular of all inverse problems as its applications extend in many areas of engineering, including thermal processing of materials, thermal monitoring in nuclear engineering, and crystal growth and solidification processes [7, 12].

Various methodologies have been proposed and successfully been implemented for the solution of the IHCP defined in Fig. 17.1 [7, 12, 13]. Several of these techniques involve restatement of the ill-posed inverse problem as a conditionally well-posed functional optimization problem. For example, in the work of Alifanov [12], the inverse heat-

conduction problem of Fig. 17.1 is stated as an optimization problem as follows:

$$\begin{aligned} \min_{q_0(t)} J(q_0) &= \min_{q_0(t)} \frac{1}{2} \|T(d, t; q_0) - Y(t)\|_{L_2(\mathcal{T})}^2 \\ &= \min_{q_0(t)} \frac{1}{2} \int_{\mathcal{T}} \{T(d, t; q_0) - Y(t)\}^2 dt \end{aligned} \quad (17.1)$$

where  $T(d, t; q_0)$  is the solution of the parametric direct heat-conduction problem at the location  $d$  of the sensor where the temperature data  $Y(t)$  are collected. This well-posed direct problem is defined for each (guessed) heat-flux  $q_0$  applied on the left wall, with insulated boundary conditions on the right wall and all material properties and initial conditions given. It is to be noted that the measured/desired temperature response  $Y(t)$  does not play a part in the definition of this direct problem. Based on Eq. (17.1), one is looking for a heat-flux  $q_0$  so that the solution of the direct problem at  $x = d$  matches, in the  $L_2$  norm sense, the given data. This optimization problem can be solved in either finite or infinite-dimensional spaces. In the first approach, an a priori parametrization is chosen for the boundary heat-flux. If the heat-flux function itself is considered as the design parameter, then the optimization problem has to be solved in an appropriate infinite-dimensional function space. Examples of such an approach include the iterative regularization method [12, 14] where gradient-based optimization techniques are introduced using continuum or discrete sensitivity and/or adjoint problems [14, 15]. For example, in the infinite-dimensional approach, the adjoint fields are used to define the search direction, whereas the sensitivity fields are used to compute the step size in the search direction [15].

The ill-posed nature of these inverse problems can be addressed using appropriate techniques, including Tikhonov regularization [16]. For example, the previous optimization problem is restated as the minimization of the cost functional  $\mathcal{J}(q_0)$  defined as follows:

$$\mathcal{J}(q_0) = \frac{1}{2} \|T(d, t; q_0) - Y(t)\|_{L_2(\mathcal{T})}^2 + \alpha \|q_0\|_{H_1}$$

$$= \frac{1}{2} \int_{\mathcal{T}} [T(d, t; q_0) - Y(t)]^2 dt + \alpha \int_{\mathcal{T}} [q_0^2 + \left(\frac{dq_0}{dt}\right)^2] dt \quad (17.2)$$

The second term on the right hand side of this equation introduces a form of temporal regularization of the heat-flux  $q_0$  (in simple terms it smooths variations in  $q_0$  and its time derivative). The choice of the regularization parameter  $\alpha$  in the above functional is not a trivial one. The discrepancy principle of Morozov [17] states that  $\alpha$  should be chosen such that the error in the estimation of  $q_0$  (expressed in terms of the cost functional) is consistent with the measurement error in the  $Y(t)$  data. This technique is not simple to enforce (as in many occasions the error in the data is not known) and thus  $\alpha$  is usually chosen from the range of values that lead to smooth solutions when minimizing the cost functional in the equation above. Note that similar forms of spatial regularization are more commonly used for multi-dimensional inverse heat-conduction problems to constrain the spatial variation of the heat-flux on the boundary it is applied.

An alternative form of temporal regularization (often used in addition to spatial Tikhonov regularization) is the regularization introduced by the *future time method* of Beck that is appropriate for a sequential (one time step at a time) estimation of the heat-flux  $q_0$  [7]. Let us consider that the heat-flux  $q_0$  is assumed to vary in a piecewise constant manner (as we discussed above with the function specification method). In particular, let us denote with  $q_M$  the value of the heat-flux in the time interval  $(t_{M-1}, t_M)$ . Let us consider that the heat-flux  $q_0$  has been estimated up to time  $t_{M-1}$ . To estimate the heat-flux  $q_M$ , we consider the objective function in Eq. (17.2) but defined over the time interval  $(t_M, t_{M+r-1})$  (usually the number  $r$  of future steps being small, of the order of 1, 2, or 3, depending on the size of the time step  $\Delta t$ ). Our objective function that we need to minimize with respect to the heat-fluxes  $q_M, q_{M+1}, \dots, q_{M+r-1}$  now takes the following

form:

$$\mathcal{J}(q_M, q_{M+1}, \dots, q_{M+r-1}) = \frac{1}{2} \int_{t_M}^{t_{M+r-1}} [T(d, t; q_0) - Y(t)]^2 dt \quad (17.3)$$

To allow a sequential calculation of the heat-flux (i.e., to compute  $q_M$ ), Beck and colleagues introduced a simple but effective form of time regularization by *temporarily* constraining the variation of the heat-flux in time (Fig. 17.2), mainly by taking  $q_M = q_{M+1} = \dots = q_{M+r-1}$  [7]. With this constraint, Eq. (17.3) is only a function of  $q_M$  and gradient optimization techniques can be used for its solution. With a successful calculation of  $q_M$ , the calculation proceeds by one time step and is repeated. This technique has been applied successfully for sequential calculation of the boundary heat-flux in both one-dimensional and multi-dimensional problems [18].

In practice, experimentation, measurements, material data, and process conditions are always polluted with noise and estimation error. Hence, to obtain a solution robust to fluctuations in input data, it is necessary to incorporate these uncertainties in the inverse methodology. Even though a number of methods exist to address inverse heat-conduction problems (such as the methods discussed earlier), limited work is reported on how the error (uncertainty) in the given data propagates to the inverse solution. Uncertainties in the material data and the given thermal conditions are seldom taken into account. Most of the existent inverse techniques consider measurement noise, but complete probabilistic characterization of the inverse solution is not done. Addressing robustness of the solution to fluctuations in the given data or providing full probabilistic characterization of the solution requires integration of powerful uncertainty analysis techniques with innovative inverse analysis methodologies. A monograph by Kaipio and Somersalo provides such a comprehensive statistical approach to inverse problems with emphasis in electrical impedance tomography [19]. Stochastic or statistical applications to heat transfer have been more limited. Among others, Emery et al. have explored the

use of an extended maximum likelihood estimator (MLE) framework [20, 21] for inverse parameter estimation problems in heat-conduction systems.

In this chapter, we will not provide any further review of deterministic inverse methods for the IHCP since such methods are very well-documented elsewhere in the literature [7, 12, 13, 22, 23, 24]. These methods work very well and are simple to implement. Even though their role will remain strong in the future, we herein plan to put forward some modern ideas with the hope that they can provide an incentive for further developments and discussion in the solution of not only the IHCP but also of other more complex data-driven inverse problems in thermal sciences and engineering.

To introduce stochastic inverse modeling, we need to rely on techniques for forward stochastic modeling. Stochastic modeling and uncertainty analysis techniques for continuum systems have been developed considerably in the last two decades. Powerful techniques like the extended perturbation method [25, 26], the improved Neumann expansion method [27], and generalized polynomial chaos techniques [28]-[32] have been proposed and successfully used to analyze uncertainty propagation in various continuum systems. We will herein make use of the spectral stochastic finite-element method [33]-[35] and of Markov Chain Monte Carlo (MCMC) techniques [36]-[39].

Two new methods for addressing the IHCP in a fully stochastic setting are introduced. In the first part, we integrate the spectral stochastic finite-element method and the iterative regularization (functional optimization) technique to obtain a robust solution to the stochastic inverse heat-conduction problem (SIHCP) in the presence of uncertainties in input data. In the SIHCP, one computes the complete stochastic response of the boundary heat-flux in a conducting solid given the probability density function (PDF) of temperature at various points within the solid. Uncertainties can be directly incorporated in the inverse-problem statement by considering randomness in material data and the initial and boundary conditions. Since this approach is based on functional inter-

pretation of a stochastic process, it is elegantly integrated with a functional optimization technique such as the iterative regularization method of Alifanov [12]. The second technique is based on a Bayesian inference approach. Our approach will push forward in these directions and take advantage of recent advances in Bayesian computational statistics [40, 41] and spatial statistics [42] to provide robust formulations and efficient numerical algorithms for stochastic inverse heat-transfer problems.

## **17.2 THE INVERSE HEAT-CONDUCTION PROBLEM - A SPECTRAL STOCHASTIC APPROACH**

The sequencing of various sub-sections in this section are as follows: The functional interpretation of a random variable and various spectral expansion techniques for representing random variables are discussed in Section 17.2.1. In Section 17.2.2, the inverse problem is stated as an  $L_2$  optimization problem with appropriate definitions of the cost functional. The need for calculating the stochastic adjoint field is discussed and appropriate definition of the stochastic adjoint field is given. Further, a new interpretation to temperature sensitivity in the context of random processes is discussed. Application of the conjugate gradient method for the solution of the inverse problem is reviewed in Section 17.2.3. Due to the essential similarity in the three parametric direct, adjoint, and sensitivity subproblems, in Section 17.2.4, the solution methodology is discussed only for the direct problem. In Section 17.2.5, the algorithm is applied to various stochastic inverse heat-conduction problems. Finally, in Section 17.2.6, directions for further investigation in the use of spectral methods for stochastic inverse problems are suggested.

### 17.2.1 Introduction: Representation of random variables

A probability space can be specified as  $(\Omega, \mathcal{F}, \mathcal{P})$ , where  $\Omega$  is the sample space of elementary events,  $\mathcal{F}$  is the minimal  $\sigma$ -algebra of  $\Omega$ , and  $\mathcal{P}$  is a probability measure. It should be noted that the  $\sigma$ -algebra consists of all allowed permutations of the basic outcomes in  $\Omega$  [43].

In this context, a real-valued random variable can be interpreted as a function that maps each point on the sample space to a corresponding point on the real line according to the probability measure. In short, a real-valued random variable  $X$  can be written as  $X : (\Omega, \mathcal{F}, \mathcal{P}) \mapsto \mathbb{R}$ . For notational convenience, a random variable will be denoted as  $X(\Omega) \mapsto \mathbb{R}$  with the understanding that  $\Omega$  denotes  $(\Omega, \mathcal{F}, \mathcal{P})$ . Extending this interpretation to stochastic processes, a real-valued space-time stochastic process  $w(\mathbf{x}, \mathbf{t})$  with a known probability distribution can be written as a function  $w(\mathbf{x}, \mathbf{t}, \theta)$  with  $\mathbf{x}$ ,  $\mathbf{t}$ , and  $\theta$  denoting dependence on space, time, and the sample space, respectively. In short,  $w(\mathbf{x}, \mathbf{t}, \theta) : (\mathcal{D}, \mathcal{T}, \Omega) \mapsto \mathbb{R}$ , where,  $\mathcal{D}$  and  $\mathcal{T}$  denote the spatial and temporal domains, respectively. In subsequent sections, any quantity with a  $\theta$ -dependence represents a random quantity and  $\theta$  will be referred to as the random dimension.

In the spectral stochastic approach, random processes are expressed in a generalized Fourier-like expansion. This allows for a higher-order, computationally efficient representation of uncertainty compared to perturbation methods. The spectral stochastic framework is based on ideas of homogenous chaos [44]. We proceed to discuss two classical uncertainty representation techniques, the Karhunen-Loève expansion and the polynomial chaos expansion.

*Karhunen-Loève (KL) expansion:* A spectral expansion of the covariance function of a stochastic process is considered [43]. Let us denote the stochastic process as  $w(\mathbf{x}, \mathbf{t}, \theta)$  and its covariance function as  $\mathcal{R}(\mathbf{x}, \mathbf{t}_1, \mathbf{y}, \mathbf{t}_2)$ , where  $\mathbf{x}$ ,  $\mathbf{y}$  are spatial coordinates and  $\mathbf{t}_1$ ,

$t_2$  are temporal coordinates. By definition, the covariance function is real, symmetric, and positive definite. Thus, all its eigenfunctions are mutually orthogonal and form a complete set spanning the function space to which  $\mathcal{R}(\mathbf{x}, \mathbf{t}_1, \mathbf{y}, \mathbf{t}_2)$  belongs. The KL expansion can be written as

$$w(\mathbf{x}, \mathbf{t}, \theta) = \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i(\theta) \mathbf{f}_i(\mathbf{x}, \mathbf{t}) + \bar{\mathbf{w}}(\mathbf{x}, \mathbf{t}) \quad (17.4)$$

where  $\bar{\mathbf{w}}(\mathbf{x}, \mathbf{t})$  denotes the mean of the stochastic process and  $\{\xi_i(\theta)\}_{i=0}^{\infty}$  forms a set of independent random variables spanning the probability space  $(\Omega, \mathcal{F}, \mathcal{P})$ . The scalars  $\lambda_i$  and deterministic functions  $f_i(\mathbf{x}, \mathbf{t})$  are the eigen-pairs of the covariance function, i.e.,

$$\int_{\mathcal{T}} \int_{\mathcal{D}} \mathcal{R}(\mathbf{x}, \mathbf{t}_1, \mathbf{y}, \mathbf{t}_2) \mathbf{f}_n(\mathbf{x}, \mathbf{t}_1) d\mathbf{x} d\mathbf{t}_1 = \lambda_n \mathbf{f}_n(\mathbf{y}, \mathbf{t}_2) \quad (17.5)$$

The KL expansion also has an interesting property. Of all possible decompositions of a random process, the KL expansion is optimal in mean-square sense. However, its application is limited by the fact that the covariance function needs to be known a priori. This limitation is complemented by the use of another expansion technique known as the polynomial chaos expansion that is discussed next.

*Polynomial chaos expansion (PCE):* The classical polynomial chaos expansion employs Hermite polynomials with multidimensional  $\mathcal{N}(0, 1)$  random variables as a trial basis for the probability space  $(\Omega, \mathcal{F}, \mathcal{P})$  to represent a stochastic process [44]. Further, Cameron and Martin [45] showed that such a choice of basis leads to a spectral expansion that converges in an  $L_2$  sense to any random variable with finite variance.

Let us consider a set of independent, identically distributed (iid) Gaussian random variables as  $\{\xi_i(\theta)\}_{i=0}^{\infty}$ . The polynomial chaos representation of a random process can be

written as:

$$w(\mathbf{x}, \mathbf{t}, \theta) = \mathbf{a}_0(\mathbf{x}, \mathbf{t})\mathbf{H}_0 + \sum_{i_1=1}^{\infty} \mathbf{a}_{i_1}(\mathbf{x}, \mathbf{t})\mathbf{H}_1[\xi_{i_1}(\theta)] + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} \mathbf{a}_{i_1 i_2}(\mathbf{x}, \mathbf{t})\mathbf{H}_2[\xi_{i_1}(\theta), \xi_{i_2}(\theta)] + \dots \quad (17.6)$$

where  $H_p(\cdot)$ ,  $p = 0, 1, 2, \dots$  (popularly known as polynomial chaoses) are Hermite polynomials in  $(\xi_1(\theta), \dots)$ . Equation (17.6) can be re-written in a compact form:

$$w(\mathbf{x}, \mathbf{t}, \theta) = \sum_{i=0}^{\infty} \mathbf{w}_i(\mathbf{x}, \mathbf{t})\psi_i(\theta) \quad (17.7)$$

where there is one-to-one correspondence between the functionals  $H_p(\cdot)$  and  $\psi_i(\theta)$ . For further details, the reader is referred to [33]. It should be emphasized here that all information about stochasticity comes from the definition of the stochastic process.

KL and PCE are just two ways of representing the stochastic process using different bases of the probability space  $(\Omega, \mathcal{F}, \mathcal{P})$ . The PCE expansion can also be effectively used to model non-Gaussian processes. In more recent work [30, 32] uses the generalized Askey system of orthogonal polynomials [46] as trial bases for the probability space.

### 17.2.2 The stochastic inverse heat-conduction problem (SIHCP): Problem definition

Let  $\mathcal{D}$  be a bounded region in  $R^d$  with boundary  $\Gamma$  occupied by a material with random thermal conductivity  $k(\mathbf{x}, \theta)$  and heat capacity  $C(\mathbf{x}, \theta)$ . Let the boundary  $\Gamma$  be further sub-divided as  $\Gamma_h$  and  $\Gamma_0$  with  $\Gamma_h \cap \Gamma_0 = \emptyset$ , where,  $\Gamma_h$  is the part of the boundary  $\Gamma$  with known thermal boundary conditions (here, heat-flux). The heat-flux on the boundary  $\Gamma_0$  is considered unknown and is the main variable of the present inverse analysis. Extra thermal conditions in the form of measured temperature  $Y(\mathbf{x}, \mathbf{t}, \theta)$  are provided along a surface  $\Gamma_I$  within the conducting solid. A sketch of the problem definition is provided in Fig. 17.3. The objective of the present analysis is to calculate the unknown stochastic

flux on the boundary  $\Gamma_0$  that approximately yields the measured stochastic temperature  $Y(\mathbf{x}, \mathbf{t}, \theta)$  along  $\Gamma_I$ .

With this interpretation, the SIHCP can be seen as an extension of the classical inverse heat-conduction problem to the stochastic regime [31]. It is formally defined as follows:

$$C \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\mathcal{D}, \mathcal{T}, \Omega) \quad (17.8)$$

$$T(\mathbf{x}, \mathbf{0}, \theta) = \hat{T}(\mathbf{x}, \theta), \quad (\mathbf{x}, \theta) \in (\mathcal{D}, \Omega) \quad (17.9)$$

$$k \frac{\partial T}{\partial n} = \hat{f}(\mathbf{x}, \mathbf{t}, \theta), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_h, \mathcal{T}, \Omega) \quad (17.10)$$

$$k \frac{\partial T}{\partial n} = q_0(\mathbf{x}, \mathbf{t}, \theta), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_0, \mathcal{T}, \Omega) \quad (q_0 \text{ unknown}) \quad (17.11)$$

$$T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0) \simeq Y(\mathbf{x}, \mathbf{t}, \theta), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_I, \mathcal{T}, \Omega) \quad (17.12)$$

where  $C(= \rho C_p)$  is the volumetric heat capacity and  $k$  the conductivity.

In this case of a stochastic inverse analysis, randomness in the temperature readings  $Y(\mathbf{x}, \mathbf{t}, \theta)$  along  $\Gamma_I$  results from experimental noise, sensor statistics, and other non-quantifiable sources. This randomness coupled with uncertainties in material data make it impossible to exactly determine the unknown boundary flux. Thus, the analysis attempts to reconstruct the statistics of the unknown boundary heat-flux.

Even a deterministic flux applied along  $\Gamma_0$  leads to a stochastic temperature response along  $\Gamma_I$ . This is due to propagation of uncertainty in material data and given thermal conditions to the temperature solution. Thus, a SIHCP statement should allow for variabilities in the temperature along  $\Gamma_I$ . This is achieved by specifying a stochastic desired temperature profile along  $\Gamma_I$ . Also, the inverse objective is modified to account for variabilities in the desired (heat-flux) unknowns. Note that effectively the SIHCP is a well-posed extension of the IHCP in a larger space of probability distributions.

In this work, it is assumed that a quasi-solution to the inverse problem exists in the

sense of Tikhonov [16]. In particular, one looks for a flux  $\bar{q}_0(\mathbf{x}, \mathbf{t}, \theta) \in \mathbf{L}_2(\Gamma_0 \times \mathcal{T} \times \Omega)$  such that:

$$\mathcal{J}(\bar{q}_0) \leq \mathcal{J}(q_0), \quad \forall q_0 \in L_2(\Gamma_0 \times \mathcal{T} \times \Omega) \quad (17.13)$$

where,  $L_2(\Gamma_0 \times \mathcal{T} \times \Omega)$  is the space of all mean-square integrable stochastic processes defined over the spatial and temporal domain  $\Gamma_0$  and  $\mathcal{T}$ . Further,  $\mathcal{J}(q_0)$ , the objective function, is chosen to be the  $L_2$  norm of the error between estimated stochastic temperature and measured/desired temperature along the surface  $\Gamma_I$ :

$$\mathcal{J}(q_0) = \frac{1}{2} \|T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0) - \mathbf{Y}(\mathbf{x}, \mathbf{t}, \theta)\|_{\mathbf{L}_2(\Gamma_I \times \mathcal{T} \times \Omega)}^2 \quad (17.14)$$

$$= \frac{1}{2} \int_{\mathcal{T}} \int_{\Gamma_I} \int_{\Omega} [T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0) - \mathbf{Y}(\mathbf{x}, \mathbf{t}, \theta)]^2 d\mathbf{P} d\mathbf{x} dt \quad (17.15)$$

where  $T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0) \equiv \mathbf{T}(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0(\mathbf{x}, \mathbf{t}, \theta))$  is the solution of the parametric direct stochastic heat-conduction problem defined in Box I and  $\int_{\Omega} \bullet dP$  denotes an integral with respect to the probability measure on  $(\Omega, \mathcal{F}, \mathcal{P})$ . It is to be noted that the measured/desired temperature response  $Y(\mathbf{x}, \mathbf{t}, \theta)$  does not play a part in the definition of the direct problem.

The main difficulty in solving the optimization problem defined in Eq. (17.14) is the calculation of the gradient  $\mathcal{J}'(q_0)$  of the cost functional in the function space  $L_2(\Gamma_0 \times \mathcal{T} \times \Omega)$ . Typically, this gradient is evaluated in a distributional or weak sense. To this end, let us introduce the directional derivative  $D_{\Delta q_0} \mathcal{J}(q_0) \equiv (\mathcal{J}'(q_0), \Delta q_0)_{L_2(\Gamma_0 \times \mathcal{T} \times \Omega)}$ . Using Eq. (17.14), one can write the following:

$$\begin{aligned} D_{\Delta q_0} \mathcal{J}(q_0) &\equiv (\mathcal{J}'(q_0), \Delta q_0)_{L_2(\Gamma_0 \times \mathcal{T} \times \Omega)} \\ &= (T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0) - \mathbf{Y}(\mathbf{x}, \mathbf{t}, \theta), \Theta(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0, \Delta \mathbf{q}_0))_{\mathbf{L}_2(\Gamma_I \times \mathcal{T} \times \Omega)} \end{aligned} \quad (17.16)$$

where the sensitivity temperature field  $\Theta(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0, \Delta \mathbf{q}_0) \equiv \mathbf{D}_{\Delta \mathbf{q}_0} \mathbf{T}(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0)$  is defined as

the linear part in  $\Delta q_0$  in the expansion of the process  $T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0 + \Delta \mathbf{q}_0)$  i.e.

$$T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0 + \Delta \mathbf{q}_0) = \mathbf{T}(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0) + \mathbf{D}_{\Delta \mathbf{q}_0} \mathbf{T}(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0) + \mathcal{O}(\|\Delta \mathbf{q}_0\|_{\mathbf{L}_2(\Gamma_0 \times \mathcal{T} \times \Omega)}^2) \quad (17.17)$$

*The continuum stochastic sensitivity method:* Interpreting uncertainties as fluctuations, the direct stochastic heat-conduction problem is often considered to possess information about both deterministic direct and sensitivity problems. This kind of reasoning is incorrect since in practice, it is common to come across fluctuations of 10% of mean. Here, uncertainties are no longer small perturbations. Further, the mean solution of the direct stochastic heat-conduction problem can be totally different than the solution of the direct deterministic heat-conduction problem.

Thus, to provide a mathematically sound sensitivity analysis technique, a continuum stochastic sensitivity method (CSSM) is introduced. Here, the sensitivity of the temperature with respect to the design flux  $q_0$  is interpreted as the perturbation in the PDF of the temperature due to perturbations in PDF of  $q_0$ . Thus, the equations governing the sensitivity problem are obtained by differentiating the direct stochastic problem with respect to perturbations in the unknown stochastic flux. The resultant sensitivity equations are summarized in Box II.

**Box I:** Direct problem to define  $T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0)$

$$C \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\mathcal{D}, \mathcal{T}, \Omega) \quad (17.18)$$

$$T(\mathbf{x}, \mathbf{0}, \theta) = \hat{T}(\mathbf{x}, \theta), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\mathcal{D}, \Omega) \quad (17.19)$$

$$k \frac{\partial T}{\partial n}(\mathbf{x}, \mathbf{t}, \theta) = q_0(\mathbf{x}, \mathbf{t}, \theta), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_0, \mathcal{T}, \Omega) \quad (17.20)$$

$$k \frac{\partial T}{\partial n}(\mathbf{x}, \mathbf{t}, \theta) = \hat{f}(\mathbf{x}, \mathbf{t}, \theta), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_h, \mathcal{T}, \Omega) \quad (17.21)$$

It is clear from the definition of the directional derivative of the cost functional Eq. (17.16) that the calculation of the gradient  $\mathcal{J}'(q_0)$  requires the evaluation of the adjoint to the sensitivity operator. The definition of the continuum adjoint problem is undertaken next.

**Box II:** Sensitivity problem to define  $\Theta(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0, \Delta \mathbf{q}_0)$

$$C \frac{\partial \Theta}{\partial t} = \nabla \cdot (k \nabla \Theta), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\mathcal{D}, \mathcal{T}, \Omega) \quad (17.22)$$

$$\Theta(\mathbf{x}, \mathbf{0}, \theta; \mathbf{q}_0, \Delta \mathbf{q}_0) = 0, \quad (\mathbf{x}, \theta) \in (\mathcal{D}, \Omega) \quad (17.23)$$

$$k \frac{\partial \Theta}{\partial n}(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0, \Delta \mathbf{q}_0) = \Delta q_0(\mathbf{x}, \mathbf{t}, \theta), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_0, \mathcal{T}, \Omega) \quad (17.24)$$

$$k \frac{\partial \Theta}{\partial n}(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0, \Delta \mathbf{q}_0) = 0, \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_h, \mathcal{T}, \Omega) \quad (17.25)$$

### 17.2.3 Adjoint problem and gradient calculations

Let us denote the adjoint temperature variable as  $\phi(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0)$ . The adjoint operator is defined from the following equality:

$$(\mathcal{L}^*(\phi), \Theta)_{L_2(\mathcal{D} \times \mathcal{T} \times \Omega)} = (\phi, \mathcal{L}(\Theta))_{L_2(\mathcal{D} \times \mathcal{T} \times \Omega)} \equiv 0 \quad (17.26)$$

where the inner product is here defined as

$$(f, g)_{L_2(\mathcal{D} \times \mathcal{T} \times \Omega)} = \int_0^{t_{\max}} \int_{\mathcal{D}} \int_{\Omega} f g \, dP \, d\mathbf{x} \, dt \quad (17.27)$$

with the processes  $f, g$  defined on the domain of interest  $\mathcal{D}$  and belonging to the function space  $L_2(\mathcal{D} \times \mathcal{T} \times \Omega)$  and  $\mathcal{T}$  is defined as the time interval  $(0, t_{\max})$ . From Eq. (17.26), one can obtain

$$(\phi, \mathcal{L}(\Theta))_{L_2(\mathcal{D} \times \mathcal{T} \times \Omega)} = \int_0^{t_{\max}} \int_{\mathcal{D}} \int_{\Omega} \left( C \frac{\partial \Theta}{\partial t} - \nabla \cdot (k \nabla \Theta) \right) \phi \, dP \, d\mathbf{x} \, dt \quad (17.28)$$

The first term on the right-hand side, with integration by parts, takes the following form:

$$\int_0^{t_{\max}} \int_{\mathcal{D}} \int_{\Omega} C \frac{\partial \Theta}{\partial t} \phi \, dP \, d\mathbf{x} \, dt = \int_{\mathcal{D}} \int_{\Omega} [C\Theta\phi]_0^{t_{\max}} \, dP \, d\mathbf{x} - \int_0^{t_{\max}} \int_{\mathcal{D}} \int_{\Omega} C \frac{\partial \phi}{\partial t} \Theta \, dP \, d\mathbf{x} \, dt \quad (17.29)$$

Simplifying the second term on the right-hand side of Eq. (17.28) results in

$$\begin{aligned} & - \int_0^{t_{\max}} \int_{\mathcal{D}} \int_{\Omega} \nabla \cdot (k\nabla\Theta)\phi \, dP \, d\mathbf{x} \, dt = - \int_0^{t_{\max}} \int_{\mathcal{D}} \int_{\Omega} \nabla \cdot (k\nabla\phi)\Theta \, dP \, d\mathbf{x} \, dt \\ & - \int_0^{t_{\max}} \int_{\Gamma} \int_{\Omega} k \frac{\partial \Theta}{\partial n} \phi \, dP \, d\mathbf{x} \, dt + \int_0^{t_{\max}} \int_{\Gamma} \int_{\Omega} k \frac{\partial \phi}{\partial n} \Theta \, dP \, d\mathbf{x} \, dt \end{aligned} \quad (17.30)$$

Using Eqs. (17.29) and (17.30) and  $(\phi, \mathcal{L}(\Theta))_{L_2(\mathcal{D} \times \mathcal{T} \times \Omega)} = 0$  in Eq. (17.28), the following adjoint-field governing equation is obtained:

$$C \frac{\partial \phi}{\partial t} = -\nabla \cdot (k\nabla\phi) \quad (17.31)$$

along with the final time condition

$$\phi(\mathbf{x}, \mathbf{t}_{\max}, \theta) = \mathbf{0} \quad (17.32)$$

The remaining boundary integral terms can be simplified as follows

$$\int_0^{t_{\max}} \int_{\Gamma_0} \int_{\Omega} \Delta q_0 \phi \, dP \, d\mathbf{x} \, dt = \int_0^{t_{\max}} \int_{\Gamma_I} \int_{\Omega} \left[ k \frac{\partial \phi}{\partial n} \right]_{\Gamma_I} \Theta \, dP \, d\mathbf{x} \, dt \quad (17.33)$$

with the condition

$$k \frac{\partial \phi}{\partial n}(\mathbf{x}, \mathbf{t}, \theta) = \mathbf{0}, \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma, \mathcal{T}, \Omega) \quad (17.34)$$

The symbol  $[\cdot]_{\Gamma_I}$  is used here to denote imposed discontinuity on  $\Gamma_I$ . The adjoint variable  $\phi$  is taken to be continuous on  $\Gamma_I$  but with discontinuous first derivative. The following

jump condition across the boundary  $\Gamma_I$  is introduced:

$$\left[ k \frac{\partial \phi}{\partial n}(\mathbf{x}, \mathbf{t}, \theta) \right]_{\Gamma_I} = T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0) - \mathbf{Y}(\mathbf{x}, \mathbf{t}, \theta), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_I, \mathcal{T}, \Omega) \quad (17.35)$$

Substituting Eq. (17.35) in Eq. (17.33) and comparing with Eq. (17.16), one obtains the following relation for the gradient of the cost functional  $\mathcal{J}'(q_0)$ :

$$\mathcal{J}'(q_0) = \phi(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_0, \mathcal{T}, \Omega) \quad (17.36)$$

The equations defining the adjoint variable  $\phi(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0)$  are summarized in Box III.

**Box III:** Adjoint problem to define  $\phi(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0(\mathbf{x}, \mathbf{t}, \theta))$

$$C \frac{\partial \phi}{\partial t} = -\nabla \cdot (k \nabla \phi) \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\mathcal{D}, \mathcal{T}, \Omega) \quad (17.37)$$

$$\phi(\mathbf{x}, \mathbf{t}_{\max}, \theta) = 0, \quad (\mathbf{x}, \theta) \in (\mathcal{D}, \Omega) \quad (17.38)$$

$$k \frac{\partial \phi}{\partial n}(\mathbf{x}, \mathbf{t}, \theta) = 0, \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma, \mathcal{T}, \Omega) \quad (17.39)$$

$$\left[ k \frac{\partial \phi}{\partial n}(\mathbf{x}, \mathbf{t}, \theta) \right]_{\Gamma_I} = T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0) - \mathbf{Y}(\mathbf{x}, \mathbf{t}, \theta), \quad (\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_I, \mathcal{T}, \Omega) \quad (17.40)$$

#### 17.2.4 Numerical Implementation

After obtaining the gradient  $\mathcal{J}'(q_0)$  of the cost functional  $\mathcal{J}(q_0)$  by the adjoint method, any of the standard optimization algorithms can be used to solve the inverse/design optimization problem defined by Eq. (17.15). Here, the conjugate gradient method (CGM) is used to minimize  $\mathcal{J}(q_0)$  [15]. The overall algorithm of computing the optimal heat-flux  $\bar{q}_0 \in L_2(\Gamma_0 \times \mathcal{T} \times \Omega)$  is summarized in Box IV. It is to be noted that for each CGM iteration three continuum stochastic subproblems, viz. the direct, sensitivity, and adjoint problems, are solved.

By comparing Eqs. (17.18), (17.22), and (17.37), it can be inferred that the governing stochastic partial differential equation is the same for the direct, sensitivity, and adjoint problems. In the adjoint problem the backward heat-conduction problem is solved by making the substitution  $\tau = t_{\max} - t$ . Thus, the derivation of the weak statement is common for all three subproblems. Thus, in this section, a weak form for the direct problem is introduced followed by a discussion of its implementation using the spectral stochastic finite-element method. Implementation of boundary heat-flux conditions where the subproblems differ is highlighted.

From Eq. (17.18), the weak form for the direct problem may be written as follows: Find  $T \in H^1(\mathcal{D} \times \mathcal{T} \times \Omega)$  such that for all  $w \in H^1(\mathcal{D} \times \mathcal{T} \times \Omega)$ , the following is satisfied

$$a(T, w) = (f, w) \quad (17.41)$$

where the bilinear form  $a(T, w)$  is given as follows:

$$a(T, w) = \int_{\mathcal{D}} \int_{\Omega} \left\{ C(\mathbf{x}, \theta) \frac{\partial \mathbf{T}}{\partial \mathbf{t}} \mathbf{w} + \mathbf{k}(\mathbf{x}, \theta) \nabla \mathbf{T} \cdot \nabla \mathbf{w} \right\} dP \, dx \quad (17.42)$$

The linear form  $(f, w)$  for the direct subproblem takes the following form:

$$(f, w) = \int_{\Gamma_0} \int_{\Omega} q_0(\mathbf{x}, \mathbf{t}, \theta) \mathbf{w} \, dP \, dx + \int_{\Gamma_h} \int_{\Omega} \hat{\mathbf{f}}(\mathbf{x}, \mathbf{t}, \theta) \mathbf{w} \, dP \, dx \quad (17.43)$$

The weak statements are implemented numerically using a spectral stochastic finite-element approach in which the weighting function  $w$  is assumed to be a product of a standard Galerkin weighting function and a polynomial chaos. This specification is pseudo-spectral in nature, since, Galerkin weighting functions are local polynomials, whereas polynomial chaoses are global polynomials. In particular, the following func-

tional form is chosen for  $w$ :

$$w_m(\mathbf{x}, \theta) = \hat{\mathbf{w}}_\alpha(\mathbf{x})\psi_r(\theta) \quad (17.44)$$

where  $\hat{w}_\alpha(\mathbf{x})$  is a standard Galerkin weighting function,  $m = (\alpha - 1)P + r$  and  $P$  is the number of terms in the polynomial chaos expansion of temperature.

Similarly, the following finite-element discretization is chosen for temperature:

$$T(\mathbf{x}, \mathbf{t}, \theta) = \sum_{n=0}^N \mathbf{w}_n(\mathbf{x}, \theta) \mathbf{T}_n(\mathbf{t}) \quad (17.45)$$

where,  $N = \text{nbf} \times P$  and  $\text{nbf}$  is the number of basis functions used in spatial discretization and  $P$  is the number of polynomial chaoses used. The weighting function  $w_n(\mathbf{x}, \theta)$  is chosen to be of the same functional form as  $w_m(\mathbf{x}, \theta)$ :

$$w_n(\mathbf{x}, \theta) = \hat{\mathbf{w}}_\beta(\mathbf{x})\psi_s(\theta) \quad (17.46)$$

where,  $n = (\beta - 1)P + s$ . Using the above relations, Eq. (17.41) can be simplified as follows:

$$\begin{aligned} & \sum_{e=1}^{N_e} \left[ \int_{\mathcal{D}^e} \langle C(\mathbf{x}, \theta) \mathbf{w}_m(\mathbf{x}, \theta) \mathbf{w}_n(\mathbf{x}, \theta) \rangle d\mathbf{x} \frac{\partial \mathbf{T}_n}{\partial \mathbf{t}} + \int_{\mathcal{D}^e} \langle \mathbf{k}(\mathbf{x}, \theta) \nabla \mathbf{w}_m(\mathbf{x}, \theta) \cdot \nabla \mathbf{w}_n(\mathbf{x}, \theta) \rangle d\mathbf{x} \mathbf{T}_n \right] \\ &= \sum_{e=1}^{N_e} \left[ \int_{\Gamma_0^e} \langle w_m(\mathbf{x}, \theta) \mathbf{q}_0(\mathbf{x}, \mathbf{t}, \theta) \rangle d\mathbf{x} + \int_{\Gamma_h^e} \langle \mathbf{w}_m(\mathbf{x}, \theta) \hat{\mathbf{f}}(\mathbf{x}, \mathbf{t}, \theta) \rangle d\mathbf{x} \right] \quad (17.47) \end{aligned}$$

Eq. (17.47) can be further simplified:

$$\begin{aligned}
& \sum_{e=1}^{\text{Nel}} \mathbf{A} \left[ \int_{\mathcal{D}^e} \hat{w}_\alpha \hat{w}_\beta \langle C(\mathbf{x}, \theta) \psi_r(\theta) \psi_s(\theta) \rangle \, d\mathbf{x} \frac{\partial \mathbf{T}_{\beta s}}{\partial \mathbf{t}} + \int_{\mathcal{D}^e} \nabla \hat{\mathbf{w}}_\alpha \cdot \nabla \hat{\mathbf{w}}_\beta \langle \mathbf{k}(\mathbf{x}, \theta) \psi_r(\theta) \psi_s(\theta) \rangle \, d\mathbf{x} \mathbf{T}_{\beta s} \right] \\
& = \sum_{e=1}^{\text{Nel}} \mathbf{A} \left[ \int_{\Gamma_0^e} \hat{w}_\alpha \langle q_0(\mathbf{x}, \mathbf{t}, \theta) \psi_r(\theta) \rangle \, d\mathbf{x} + \int_{\Gamma_h^e} \hat{\mathbf{w}}_\alpha \langle \hat{\mathbf{f}}(\mathbf{x}, \mathbf{t}, \theta) \psi_r(\theta) \rangle \, d\mathbf{x} \right] \quad (17.48)
\end{aligned}$$

where,  $T_{\beta s} \equiv T_n$  with the understanding that  $n = (\beta - 1)P + s$ . The quantities in brackets, e.g.  $\langle \cdot \rangle$ , denote averages with respect to the probability measure  $dP$ .

To facilitate the use of standard FEM algorithms and solvers, the stochastic solution is viewed as a vector of  $P$  unknowns per computation node. The averages, e.g.,  $\langle C(\mathbf{x}, \theta) \psi_r(\theta) \psi_s(\theta) \rangle$ , are simplified as follows:

$$\langle C(\mathbf{x}, \theta) \psi_r(\theta) \psi_s(\theta) \rangle = \sum_{i=0}^N \mathbf{C}_i(\mathbf{x}, \theta) \langle \xi_i(\theta) \psi_r(\theta) \psi_s(\theta) \rangle \quad (17.49)$$

The averages  $\langle \xi_i(\theta) \psi_r(\theta) \psi_s(\theta) \rangle$  are precalculated and stored in a digital database for further use. ‘ $N$ ’ in the equation above is the number of terms used in the KL expansion of  $C(\mathbf{x}, \theta)$ .

**Box IV:** The conjugate gradient optimization algorithm

I: Make an initial guess  $q_0^0(\mathbf{x}, \mathbf{t}, \theta) \in L_2(\Gamma_0 \times \mathcal{T} \times \Omega)$ . Set  $\mathbf{k}=0$ .

II: Calculate the conjugate search direction  $p^k(\mathbf{x}, \mathbf{t}, \theta)$ ,  $(\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_0, \mathcal{T}, \Omega)$ .

1. Solve the direct problem for  $T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0^k)$ .
2. Compute the residual  $T(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0^k) - \mathbf{Y}(\mathbf{x}, \mathbf{t}, \theta)$ ,  $(\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_I, \mathcal{T}, \Omega)$ .
3. Solve the adjoint problem backwards in time for  $\phi(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0^k)$ .
4. Set  $\mathcal{J}'(q_0^k) = \phi(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0^k)$ ,  $(\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_0, \mathcal{T}, \Omega)$ .
5. Set  $\gamma^k = 0$  if  $\mathbf{k}=0$ ; otherwise:

$$\gamma^k = \frac{(\mathcal{J}'(q_0^k), \mathcal{J}'(q_0^k) - \mathcal{J}'(q_0^{k-1}))_{L_2(\Gamma_0 \times \mathcal{T} \times \Omega)}}{\|\mathcal{J}'(q_0^{k-1})\|_{L_2(\Gamma_0 \times \mathcal{T} \times \Omega)}^2}$$

6. Define  $p^k(\mathbf{x}, \mathbf{t}, \theta)$ : If  $\mathbf{k}=0$ ,  $p^0 = -\mathcal{J}'(q_0^0)$ ;  
otherwise  $p^k = \gamma^k p^{k-1} - \mathcal{J}'(q_0^k)$

III: Calculate the optimal step size  $\alpha^k$ :

1. Solve the sensitivity problem for  $\Theta(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0^k, \mathbf{p}^k)$ .
2. Calculate  $\alpha^k$  by  $\alpha^k = \frac{-(\mathcal{J}'(q_0^k), \mathbf{p}^k)_{L_2(\Gamma_0 \times \mathcal{T} \times \Omega)}}{\|\Theta(\mathbf{x}, \mathbf{t}, \theta; \mathbf{q}_0^k, \Delta \mathbf{q}_0^k)\|_{L_2(\Gamma_I \times \mathcal{T} \times \Omega)}^2}$

IV: Update  $q_0^{k+1}(\mathbf{x}, \mathbf{t}, \theta) = \mathbf{q}_0^k(\mathbf{x}, \mathbf{t}, \theta) + \alpha^k \mathbf{p}^k(\mathbf{x}, \mathbf{t}, \theta)$   $(\mathbf{x}, \mathbf{t}, \theta) \in (\Gamma_0, \mathcal{T}, \Omega)$ .

V: If  $\|q_0^{k+1} - q_0^k\|_{L_2(\Gamma_0 \times \mathcal{T} \times \Omega)} < \epsilon$  (specified tolerance), Stop.

Otherwise; set  $\mathbf{k}=\mathbf{k}+1$  and go to Step II.

### 17.2.5 Numerical results with the spectral approach

Two one-dimensional SIHCP examples are reported in this section; two-dimensional applications can be found in [31]. In all examples reported here, dimensionless quantities are considered. In particular, we consider  $C = 1$ ,  $k = 1$ , with both  $C$  and  $k$  taken as deterministic properties. Inverse heat-conduction problems that consider uncertainty in  $k$  or  $C$  are examined in [31].

Example 1: Identification of a triangular heat-flux profile Here, the triangular flux problem suggested in [[7], Chapter 5] is addressed using the proposed stochastic inverse analysis methodology. A bar of unit length that is insulated at the right end  $x = 1$  is considered. The sensor readings are obtained by applying a time varying flux  $q_{tri}^+(t)$  at the left end of the bar of the following functional form (see Fig. 17.4):

$$q_{tri}^+ = \begin{cases} 0, & 0 \leq t \leq 0.1 \\ 2.5t - 0.25, & 0.1 < t \leq 0.5 \\ 2.25 - 2.5t, & 0.5 < t \leq 0.9 \\ 0, & 0.9 < t \leq 1.0 \end{cases} \quad (17.50)$$

The readings are then polluted with a Gaussian process with the following correlation kernel

$$R_{hh}(t_1, t_2) = \exp\left(-\frac{|t_1 - t_2|}{L}\right) \quad (17.51)$$

where  $L$  is a quantity denoting the correlation length. This is a normalized measure of how far the temperature at a given time point affects the future temperature readings. With a correlation length of 0.8, it was found that to represent the Gaussian process, a four-dimensional Karhunen-Loève expansion was enough. Thus, the desired experimental temperature readings were obtained by considering 15 different realizations of the following process:

$$T(d, t; q_{tri}^+) + \sum_{i=1}^4 \sigma f_i(t) \xi_i(\theta) \quad (17.52)$$

where,  $f_i(t)$  are the eigenvectors of the correlation kernel and  $\sigma$  is the measurement noise level (here taken as the standard deviation of the instrument used for measuring temperature).

The 15 sets of experimental readings thus generated constitute the input for the inverse

analysis. However, because a spectral stochastic approach is used for solving the problem, we need to again represent the uncertainty in the 15 experimental sets of readings in a spectral form. This was done by finding the covariance of the readings and then obtaining the KL expansion (again a four-dimensional approximation). We emphasize that the final spectral expansion obtained from data is not the same as the spectral expansion used to generate the data. This is due to the loss of information when we consider finite data sets.

We consider two different cases

1. Case I : Low instrument noise level  $\sigma = 0.001$
2. Case II : High instrument noise level  $\sigma = 0.01$

The results for case I are shown in Fig. 17.5. It can be seen that the estimation of mean heat-flux is excellent. Further, the effect of other uncertainty modes of the estimated heat-flux is also explicitly calculated. As the noise increases however, the effect of uncertainty increases nonlinearly. The results are shown in Fig. 17.6. Though the estimate of mean heat-flux is excellent, the polynomial chaos terms in the expansion of heat-flux are also comparable to the mean heat-flux, hence, the estimate possesses a large variability. This example shows the robustness of the technique even with higher uncertainty levels in measurement. The technique not only gives an accurate estimate of the mean but also points to those uncertainty modes that affect the estimate in a drastic way.

The estimated flux above was obtained in a polynomial chaos expansion (using Hermite polynomials in  $(\xi_1, \dots, \xi_4)$ , the independent Gaussian random variables used to model measurement uncertainty). The form of the estimated heat-flux is as follows:

$$q(t) = E[q(t)] + \sum_{i=1}^4 q_i(t)\xi_i \quad (17.53)$$

It should be noted here that Hermite polynomials of degree one were found to be enough to characterize the uncertainty in the estimated heat-flux. Further, the standard deviation of the estimated heat-flux can be written as

$$\sigma(t) = \sum_{i=1}^4 [q_i(t)]^2 \quad (17.54)$$

The 90% probability bounds for the estimated heat-flux were evaluated using the rule of thumb ( $E[q(t)] - 1.6\sigma, E[q(t)] + 1.6\sigma$ ). Figure 17.7 shows the computed probability bounds for cases I and II, respectively.

**Example 2: Identification of an impulse heat-flux profile** Here, the impulse flux problem with two peaks is addressed using the proposed stochastic inverse analysis methodology. A bar of unit length that is insulated at the right end  $x = 1$  is considered. The sensor readings are obtained by applying a time-varying flux  $q_{tp}^+(t)$  at the left end of the bar of the following functional form:

$$q_{tp}^+ = \begin{cases} 0, & 0 \leq t \leq 0.1 \\ 1.0, & 0.1 < t \leq 0.4 \\ 0, & 0.4 < t \leq 0.6 \\ 1, & 0.6 < t \leq 0.9 \\ 0, & 0.9 < t \leq 1.0 \end{cases} \quad (17.55)$$

This problem is more difficult due to the presence of discontinuities in the surface heat-flux. Here we consider only the lower instrument noise level case (case I). The procedure for obtaining the experimental readings and their spectral stochastic representation is the same as in the previous example. The results are shown in Fig. 17.8. Again the mean solution is excellent and the higher-order uncertainty modes in the polynomial chaos expansion of the estimated heat-flux are obtained explicitly. The probability bounds computed for this solution are shown in Fig. 17.9.

### **17.2.6 Further developments on the use of spectral stochastic methods in inverse modeling**

The developments presented here provide only the early steps toward a fully stochastic framework for addressing the SIHCP. Many developments need to be explored to allow these techniques to work efficiently in both infinite- and finite-dimensional solutions. The optimal form and terms of the polynomial chaos expansions for different stochastic processes is still an unresolved issue and at present is addressed on a case by case basis. Exploring stochastic optimization with constraints and addressing inverse thermal problems defined by a coupled set of PDEs is still an unexplored area. Tikhonov type of regularization seems to play a minor or no role in these formulations that seem to be self-regularized by the relaxed input data and computed solution in larger probability spaces.

### **17.3 A BAYESIAN APPROACH TO THE INVERSE HEAT-CONDUCTION PROBLEM**

A new stochastic outlook to inverse thermal problems has recently been introduced using Bayesian inference [47, 48], which can account for uncertainties in the measurement data and is able to provide point estimates to the inverse solution with associated probabilistic characteristics. Bayesian inference incorporates a priori information regarding the unknowns and the system uncertainties into a prior distribution model that is then combined with the likelihood to formulate the posterior probability density function (PPDF) [40, 41]. The method regularizes the ill-posed inverse problem through prior distribution modeling [49, 50] and, in addition, provides means to estimate the statistics of uncertainties [48].

With the recent propagation of efficient sampling methods, such as Markov Chain

Monte Carlo (MCMC) [39], the application of Bayesian inference to engineering inverse problems becomes feasible. MCMC provides a large sample data set drawn from the PPDF. The samples can then be used to approximate the expectation of any function of the random unknowns.

The remainder of this section is organized in the following sequence. Section 17.3.1 re-introduces the SIHCP with outlines of the discrete measurement case. The formulation of the likelihood is presented in Section 17.3.2 together with the prior distribution model and the PPDF under a Bayesian inference framework. The design of the MCMC sampler is discussed in Section 17.3.3, including the exploration of the posterior state space. In Section 17.3.4, various examples are provided. Finally, Section 17.3.5 summarizes the observations and potential of this Bayesian inference approach to the solution of inverse heat-transfer problems.

### 17.3.1 Revisiting the IHCP

In this section, the stochastic inverse heat-conduction problem is restated to emphasize the discrete nature of the temperature measurements, but more importantly to introduce a statistical (rather than a stochastic) approach to the inverse problem. The inverse heat-conduction problem of interest is defined with the following equations (see Fig. 17.10):

$$C \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T), \quad \text{in } \mathcal{D}, \quad t \in [0, t_{\max}] \quad (17.56)$$

$$T(\mathbf{x}, t) = T_g, \quad \text{on } \Gamma_g, \quad t \in [0, t_{\max}] \quad (17.57)$$

$$k \frac{\partial T(\mathbf{x}, t)}{\partial \mathbf{n}} = q_h, \quad \text{on } \Gamma_h, \quad t \in [0, t_{\max}] \quad (17.58)$$

$$T(\mathbf{x}, 0) = \hat{T}(\mathbf{x}), \quad \text{in } \mathcal{D} \quad (17.59)$$

The known boundary conditions include the heat-flux  $q_h$  and the temperature  $T_g$  on the subsets  $\Gamma_h$  and  $\Gamma_g$ , respectively, of  $\Gamma$ . Finally,  $\hat{T}$  is the known initial temperature field.

The main objective of the IHCP is the calculation of the heat-flux  $q_0$  on  $\Gamma_0 \times [0, t_{\max}]$

( $\Gamma_g \cup \Gamma_h \cup \Gamma_0 = \Gamma$ ):

$$k \frac{\partial T(\mathbf{x}, t)}{\partial \mathbf{n}} = q_0, \text{ (unknown) on } \Gamma_0, \quad t \in [0, t_{\max}] \quad (17.60)$$

Approximations of the temperature field are also available at  $M$  points (temperature sensor locations) within the domain:

$$Y_i^{(j)} \simeq T(\hat{\mathbf{x}}_i, \hat{t}_j), \quad i = 1, \dots, M, \quad j = 1, \dots, N \quad (\hat{t}_N = t_{\max}) \quad (17.61)$$

Equations (17.56)–(17.60) define a well-posed direct heat-conduction problem for each guessed heat-flux  $q_0$  on  $\Gamma_0 \times [0, t_{\max}]$ . Let us denote the solution to this direct problem as  $T(\mathbf{x}, t; q_0)$ . In the IHCP, one is interested in computing  $q_0$  so that in some sense  $T(\hat{\mathbf{x}}_i, \hat{t}_j; q_0)$  matches the given temperature measurements  $Y$  (vector form of  $Y_i^{(j)}$  in Eq. (17.61)). Let us denote with  $F(q_0)$  the vector of computed temperatures at the sensor locations  $\hat{\mathbf{x}}_i$ ,  $i = 1, \dots, M$ , at times  $\hat{t}_j$ ,  $j = 1, \dots, N$  ( $\hat{t}_N = t_{\max}$ ). The operator  $F$  in general needs to be computed numerically through techniques such as finite-element analysis.

In the present implementation of the IHCP, the unknown heat-flux  $q_0(x, t)$  is discretized linearly in space and time using finite-element interpolation for the grid and time-stepping that is also used in the direct heat-conduction analysis. However, the space/time discretization used in the direct problem (state space) is generally finer than that used in the discretization of  $q_0$  (parameter space). Thus, the unknown  $q_0$  can be written as

$$q_0(\mathbf{x}, t) = \sum_{i=1}^m \theta_i w_i(\mathbf{x}, t) \quad (17.62)$$

where  $w_i$ 's are the pre-defined basis functions. The IHCP is then transformed to the

estimation of the weights  $\theta'_i$ s. These weights are considered to be represented by an unknown random vector  $\theta$  of length  $m$ .

Let us denote with  $\omega_m$  the sensor uncertainty (sensor noise). Then one looks for the vector  $\theta$  such that

$$Y \simeq F(\theta) + \omega_m \quad (17.63)$$

The statistics of noise may be known a priori or unknown as in most real-world engineering systems. In both cases, either to improve knowledge of noise distribution (known a priori) or to capture this noise information (totally unknown), the parameters controlling noise distribution can be incorporated into the unknown  $\theta$ . For example, in many engineering applications, the measurement noise is modeled as stationary zero-mean white noise with Gaussian distribution, hence, the probability density function of  $\omega_m$  is determined by a single unknown parameter  $\sigma$ , which is the standard deviation (std). An augmented unknown vector  $\theta_{aug} = [\theta \ \sigma]^T$  can be formed to be estimated from observation  $Y$ . In principle, other system uncertainties can be treated as unknowns and be computed from the given data as well.

Direct inversion of Eq. (17.63) to compute  $\theta$  is not feasible as it leads to an ill-posed system of equations. In most deterministic approaches to the IHCP, it is assumed that a quasi-solution to the inverse problem exists in the sense of Tikhonov [16]. In particular, one looks for a flux  $\bar{q}_0(\mathbf{x}, t) \in L_2(\Gamma_0 \times [0, t_{\max}])$  such that:

$$\mathcal{J}(\bar{q}_0) \leq \mathcal{J}(q_0), \quad \forall q_0 \in L_2(\Gamma_0 \times [0, t_{\max}]) \quad (17.64)$$

where,  $L_2(\Gamma_0 \times [0, t_{\max}])$  is the space of all square integrable functions defined over the spatial and temporal domains  $\Gamma_0$  and  $[0, t_{\max}]$ , respectively. The objective function  $\mathcal{J}(q_0) \equiv \mathcal{J}(\theta)$  to be minimized is usually chosen as the  $L_2$  norm of the error between the

estimated and measured temperatures along the sensor locations:

$$\begin{aligned}\mathcal{J}(q_0) &= \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^N [T(\hat{\mathbf{x}}_i, \hat{t}_j; q_0) - Y(\hat{\mathbf{x}}_i, \hat{t}_j)]^2 \\ &= \frac{1}{2} \|F(\theta) - Y\|_{L_2}^2\end{aligned}\tag{17.65}$$

where the solution  $T(\mathbf{x}, t; q_0)$  of the parametric direct problem was defined earlier.

### 17.3.2 Bayesian inverse formulation

Using the Bayesian inference approach, the inverse problem is transformed to the estimation of the joint PDF of a discrete stochastic process  $\{\theta_i, i = 1 : m\}$ . The probability density function of  $\theta$  (vector form of  $\{\theta_i, i = 1 : m\}$ ) given  $Y$  can be written according the Bayes's formula as

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{p(Y)}\tag{17.66}$$

where the conditional PDF  $p(\theta|Y)$  is called the posterior probability density function (PPDF),  $p(Y|\theta)$  is the likelihood function, and the marginal PDF  $p(\theta)$  is called the prior PDF. Once the PPDF is known, various point estimators can be computed, such as the 'Maximum A Posteriori' (MAP) estimator:

$$\hat{\theta}_{\text{MAP}} = \underset{\theta}{\text{argmax}} p(\theta|Y)\tag{17.67}$$

and the posterior mean estimator:

$$\hat{\theta}_{\text{postmean}} = E \theta|Y\tag{17.68}$$

As a normalizing constant, the knowledge of  $p(Y)$  can be avoided if the posterior state

space can be exploited up to the normalizing constant. This is actually true for the numerical sampling strategies adopted in the current work. Therefore, the PPDF can be evaluated as

$$p(\theta|Y) \propto p(Y|\theta)p(\theta) \quad (17.69)$$

The likelihood function can be obtained from the relationship

$$Y = F(\theta) + \omega_m \quad (17.70)$$

where  $F$  is the vector of computed temperatures given a guessed heat-flux (Eq. (17.62)).  $F_i$  represents the temperature at the same location and time as  $Y_i$  does. In this work, we regard  $\omega_m$  as the measurement noise generated from a normal distribution with mean 0 and standard deviation  $\sigma$  (white noise). Subsequently, the likelihood can be written as

$$p(Y|\theta) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left\{-\frac{[Y - F(\theta)]^T [Y - F(\theta)]}{2\sigma^2}\right\} \quad (17.71)$$

The prior distribution reflects the knowledge, if there is any, of the heat-flux, before  $Y$  is gathered. From an inverse point of view, the prior distribution model provides regularization to the ill-posed inverse problem [48]. In the current study, a specific form of Markov Random Field (MRF) [51] is adopted for the prior modeling of  $\theta$  by treating the temporal direction as another (additional) “spatial” coordinate. In general, the MRF can be mathematically expressed as follows:

$$p(\theta) \propto \exp\left\{-\sum_{i \sim j} W_{ij} \Phi(\gamma(\theta_i - \theta_j))\right\} \quad (17.72)$$

where  $\gamma$  is a scaling parameter,  $\Phi$  is an even function that determines the specific form of the MRF, the summation is over all pairs of sites  $i \sim j$  that are defined as neighbors, and  $W'_{ij}$ s are specified non-zero weights [41]. Let  $\Phi(u) = \frac{1}{2}u^2$ , the MRF can then be

rewritten as

$$p(\theta) \propto \lambda^{m/2} \exp\left(-\frac{1}{2}\lambda\theta^T W\theta\right) \quad (17.73)$$

In the one-parameter model of Eq. (17.73), the entries of the  $m \times m$  matrix  $W$  are determined as,  $W_{ij} = n_i$  if  $i=j$ ,  $W_{ij} = -1$  if  $i$  and  $j$  are adjacent, and as 0 otherwise.  $n_i$  is the number of neighbors adjacent to site  $i$ .  $\lambda$  is a constant that controls the scaling of distribution of  $\theta$ . This MRF model is equivalent to Tikhonov regularization provided the measurement noise is Gaussian and the objective is to maximize the posterior probability (MAP) [48].

With the specified likelihood function in Eq. (17.71) and prior PDF in Eq. (17.73), the PPDF for the inverse problem can then be formulated as

$$p(\theta|Y) \propto \exp\left\{-\frac{1}{2\sigma^2}[F(\theta) - Y]^T[F(\theta) - Y]\right\} \cdot \exp\left(-\frac{1}{2}\lambda\theta^T W\theta\right) \quad (17.74)$$

Both point estimates of MAP [Eq. (17.67)] and posterior mean [Eq. (17.68)] and associated probability bounds are computed based on this formulation.

### 17.3.3 MCMC sampler

The above-introduced PPDF has an implicit form (likelihood has a numerical solver), and, hence, can be explored numerically only through Monte Carlo simulation. The idea of general Monte Carlo simulation is to approximate the expectation or higher-order statistics of any function  $f(\theta)$  of the random variable  $\theta$  by the sample mean and sample statistics from a large set of identically independent distributed (iid) samples  $\{\theta^{(i)}, i = 1 : L\}$  drawn from the target probability density function (PDF)  $p(\theta)$ ,

$$E_L f(\theta) = \frac{1}{L} \sum_{i=1}^L f(\theta^{(i)}) \xrightarrow{L \rightarrow \infty} E f(\theta) = \int f(\theta)p(\theta)d\theta \quad (17.75)$$

The convergence is guaranteed by the strong law of large numbers. Therefore, the posterior mean estimate of Eq. (17.74) can be obtained through the above approximation, and the MAP estimate can be approximated as

$$\hat{\theta}_{MAP} = \underset{\theta^{(i)}}{\operatorname{argmax}} p(\theta^{(i)}) \quad (17.76)$$

Notice that the target PDF in this study is the PPDF  $p(\theta|Y)$ .

The key step in Monte Carlo simulation is to draw the sample set from the high dimensional and implicit PPDF. Markov Chain Monte Carlo (MCMC) is a strategy for generating samples  $\theta^{(i)}$  while exploring the state space of  $\theta$  using a Markov chain mechanism [39, 51, 52, 53]. This mechanism is constructed so that the samples  $\theta^{(i)}$  mimic samples drawn from the target distribution  $p(\theta)$ . Note that one uses MCMC when it is not possible to draw directly samples from  $p(\theta)$ , but can evaluate  $p(\theta)$  up to a normalizing constant.

The Gibbs sampler is a widely used MCMC algorithm. For an  $m$ -dimensional random vector  $\theta$ , the full conditional PDF is defined as  $p(\theta_i|\theta_{-i})$ , where  $\theta_{-i}$  stands for  $\{\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_m\}$ . When the full conditional PDF is known, it is often advantageous to use it as the proposal PDF, which is used to generate a new sample. The important feature of this sampler is that the acceptance probability is always 1. This means that the candidate sample  $\theta^{(*)}$  generated in this way will always be accepted. It is interesting to note that, in the IHCP, it is very costly to calculate the likelihood since for each candidate sample one has to perform a direct numerical simulation. Therefore, if a Gibbs sampler can be designed for this kind of problem, it will avoid the computation of the likelihood. For all linear IHCP examined in this chapter, the Gibbs sampler is used to speedup the chain convergence. With  $N_{mcmc}$  denoting the number of the MCMC steps, the algorithm can be summarized as follows:

1. Initialize  $\theta^{(0)}$

2. For  $i = 1 : \text{Nmcmc} - 1$

— sample  $\theta_1^{(i+1)} \sim p(\theta_1 | \theta_2^{(i)}, \theta_3^{(i)}, \dots, \theta_m^{(i)})$

— sample  $\theta_2^{(i+1)} \sim p(\theta_2 | \theta_1^{(i+1)}, \theta_3^{(i)}, \dots, \theta_m^{(i)})$

—  $\vdots$

— sample  $\theta_m^{(i+1)} \sim p(\theta_m | \theta_1^{(i+1)}, \theta_2^{(i+1)}, \dots, \theta_{m-1}^{(i+1)})$

From Eq. (17.70), the relation between the unknown vector  $\theta$  and observation  $Y$  can be written as

$$Y = H\theta + T_I + \omega_m \quad (17.77)$$

where  $H$  is the sensitivity matrix determined as follows ( $i = 1 : M$ ,  $j = 1 : N$ ,  $k = 1 : m$ ):

$$H[(i-1)N + j, k] = T_H(\hat{\mathbf{x}}_i, \hat{t}_j; w_k) \quad (17.78)$$

In the equation above,  $T_H$  denotes the direct solution with a zero initial temperature condition, zero temperature, and flux boundary conditions on  $\Gamma_g$  and  $\Gamma_h$ , respectively, and  $w_k$  as the applied heat-flux on  $\Gamma_0$ .  $T_I$  denotes the direct solution with zero heat-flux condition on  $\Gamma_0$  and the prescribed known initial conditions in  $\mathcal{D}$  and boundary conditions on  $\Gamma_g \cup \Gamma_h$ . Hence,  $F(\theta)$  in Eqs. (17.70) is replaced by  $H\theta + T_I$ . When  $\sigma$  is known, it can be shown that the posterior distribution follows a multivariate Gaussian; hence, the full conditional distribution of each random variable is in standard form, which can be derived as follows:

$$p(\theta_i | \theta_{-i}) \sim N(\mu_i, \sigma_i^2), \quad \mu_i = \frac{b_i}{2a_i}, \quad \sigma_i = \sqrt{\frac{1}{a_i}} \quad (17.79)$$

$$a_i = \sum_{s=1}^N \frac{H_{si}^2}{\sigma^2} + \lambda W_{ii}, \quad b_i = 2 \sum_{s=1}^N \frac{\mu_s H_{si}}{\sigma^2} - \lambda \mu_p \quad (17.80)$$

$$\mu_s = Y_s - T_I - \sum_{t \neq i} H_{st} \theta_t, \quad \mu_p = \sum_{j \neq i} W_{ji} \theta_j + \sum_{k \neq i} W_{ik} \theta_k \quad (17.81)$$

The standard form of the full conditional PDF enables us to use the Gibbs sampler in this problem. With this formulation note that the MAP estimate is the same as the posterior mean estimate. If one is interested only in the point estimate, then gradient methods can be used to solve the optimization problem of Eq. (17.67) for a linear IHCP. This approach is referred here to as “direct optimization” approach. However, implementation of a Gibbs sampler is indispensable for exploiting the posterior state space of  $\theta$  to achieve marginal PDFs, to quantify uncertainty (probability bounds) in the posterior mean estimate, or to compute the expectation of an arbitrary function of  $\theta$ .

When  $\sigma$  is not known, a modified posterior PDF and a corresponding Gibbs sampler can be employed with details provided in [48]. Statisticians have developed a large number of techniques for convergence assessment [54]. The convergence of MCMC is determined here by monitoring the histogram and marginal density of accepted samples.

#### 17.3.4 Numerical results with a Bayesian approach

One- and two-dimensional inverse heat-conduction examples are considered here. Applications to non-linear problems can be found in [55], where identification of the time history of a heat source in radiation is discussed.

##### Example 1: Identification of a triangular heat-flux profile

A one-dimensional transient linear heat-conduction problem is considered as shown in Fig. 17.4. The governing equations are as follows:

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2}, \quad 0 < t < 1, \quad 0 < x < 1 \quad (17.82)$$

$$T(x, 0) = 0, \quad 0 \leq x \leq 1 \quad (17.83)$$

$$\frac{\partial T}{\partial x} \Big|_{x=1} = 0, \quad 0 < t < 1 \quad (17.84)$$

$$\frac{\partial T}{\partial x} \Big|_{x=0} = q(t) \text{ (to be computed)}, \quad 0 < t < 1 \quad (17.85)$$

The direct problem considered represents the solution of the system of equations above with the heat-flux  $q(t)$  defined in Fig. 17.4. A FEM solver was implemented for the solution of this problem with 40 linear elements in space and 1000 time steps. As simulated measurement data at  $x = d$  (considered as the thermocouple location) we use one of the 15 data sets generated at the spectral approach discussed earlier with the assumption of uncorrelated noise. The inverse problem of interest is thus the reconstruction of the triangular heat-flux in Fig. 17.4, given Eqs. (17.82)–(17.84) and the sensor data  $Y$ .

The 1D linear finite-element shape functions are used as the basis functions in inverse solution discretization. Component of inverse vector  $\theta_i$  is the heat-flux value at  $t = idt$ , where  $dt = \frac{1.0}{m-1}$  and  $m$  is the dimension of  $\theta$ . In all the cases examined, 50,000 MCMC samples are used to generate the plots shown. The studied cases are as follows:

Case I:  $d = 0.5$ ,  $\sigma = 0.01$ ,  $\Delta t = 0.01$ ,  $dt = 0.02$ , regularization constant is  $2.5e-3$ . The computed heat-flux is shown in Fig. 17.11.

Case II:  $d = 0.5$ ,  $\sigma = 0.001$ ,  $\Delta t = 0.01$ ,  $dt = 0.02$ , regularization constant is  $2.5e-5$ . The computed heat-flux is shown in Fig. 17.12.

Case III:  $d = 0.1$ ,  $\sigma = 0.001$ ,  $\Delta t = 0.01$ ,  $dt = 0.02$ , regularization constant is  $1.0e-4$ . The computed heat-flux is shown in Fig. 17.13.

Case IV:  $d = 0.5$ ,  $\sigma = 0.001$ ,  $\Delta t = 0.01$ ,  $dt = 0.01$ , regularization constant is  $1.0e-4$ . The computed heat-flux is shown in Fig. 17.14.

To demonstrate the uncertainty quantification in the inverse solutions, besides posterior mean estimates, both confidence intervals and probability bounds are plotted. The probability bounds, which are indication of the highest density region (HDR) of the posterior state space, are relatively wide in all cases. This is due to the fact that the generated data are highly correlated while an uncorrelated assumption is used in the likelihood. The bounds are supposed to be narrower when the data are uncorrelated as seen in [48] or when more than one set of measurements are used. On the contrary, the confidence interval, which is an indication of goodness of the estimate (amount of bias of the point estimate), is extremely small such that the interval bounds are not distinguishable from the posterior mean estimate. Both probability bounds and confidence intervals are calculated using statistical analysis software S-plus after the large sample data sets (50,000 samples for each case) are obtained.

#### Example 2: Boundary heat-flux identification in 2D heat-conduction

The direct problem considered is defined in dimensionless form as follows:

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}, \quad 0 < t < 1, \quad 0 < x, y < 1 \quad (17.86)$$

$$T(x, y, 0) = 2\sin(\pi x)\sin(\pi y), \quad 0 \leq x, y \leq 1 \quad (17.87)$$

$$T|_{x=1} = T|_{y=1} = 0, \quad t > 0 \quad (17.88)$$

$$\frac{\partial T}{\partial x}|_{x=0} = q_x, \quad \frac{\partial T}{\partial y}|_{y=0} = q_y \quad t > 0 \quad (17.89)$$

An analytical solution to this problem can be obtained for the case:

$$q_x = -2\pi\sin(\pi y) \exp(-2\pi^2 t) \quad (17.90)$$

$$q_y = -2\pi\sin(\pi x) \exp(-2\pi^2 t) \quad (17.91)$$

and is given as

$$T(x, y, t) = 2\sin(\pi x)\sin(\pi y) \exp(-2\pi^2 t) \quad (17.92)$$

The inverse problem is to reconstruct  $q_x$  and  $q_y$ . The locations of simulated thermocouples are shown in Fig. 17.15. Thirteen evenly distributed thermocouples are considered with space interval of 0.125 and with the distance to the boundary also 0.125. A modified MRF is used to eliminate the edge effect [48].

The particular details of the example considered are as follows:  $\Delta t = 0.002$ , heat-flux was reconstructed for the time range  $t = 0$  to  $t = 0.1$ ,  $N = 50$ , total 650 measurements,  $\sigma = 0.01$ . For the parametric representation of  $q_x$  and  $q_y$ , 17 linear basis functions are used in both directions of  $x$  ( $q_y$ ) and  $y$  ( $q_x$ ), and 21 basis functions in temporal direction ( $dt = 0.005$ ); thus,  $m = 714$ . The regularization constant is  $1.0e - 3$ . The computed heat-flux is shown in Figs. 17.16 and 17.17. The marginal PDFs obtained by MCMC samples are plotted in Fig. 17.18. 25,000 samples are used in generating these PDFs.

### 17.3.5 Further work on using Bayesian inference for inverse thermal problems

The inverse problems examined here are linear in nature. That does not imply limitations of the method. On the contrary, in [55] a non-linear heat source identification problem in radiation was addressed using model reduction [56] for computing the likelihood, thus accelerating the computation. At present, there is limited or no work on applying Bayesian estimation to complex thermal transport systems and addressing inverse problems in that context. Treating sensor data with techniques that remove outliers remains essential in stochastic optimization. Both Bayesian and spectral approximations can be used in this context. We feel that these techniques provide an exciting new way of looking at problems where uncertainty plays a key role.

## **ACKNOWLEDGEMENTS**

The author thanks J. Wang and B. V. Asokan, respectively, for assisting in preparing the examples in the Bayesian and spectral sections of this chapter. They both have made significant contributions to the research topics addressed here in their forthcoming doctoral theses. This work was partially supported by NASA, Office of Biological and Physical Sciences Research (grant 98-HEDS-05), the Computational Mathematics program of the Air Force Office of Scientific Research (grant F49620-00-1-0373), and by the National Science Foundation (grant DMI-0113295). This work was prepared while the author was supported during his sabbatical leave from Cornell by the Statistical and Applied Mathematics Science Institute (SAMSI) at the Research Triangle Park, NC. The results reported in this chapter were obtained using the resources of the Cornell Theory Center.

## NOMENCLATURE

$A$	assembly symbol
$C$	specific heat
$d$	distance of thermocouple to the boundary
$\mathcal{D}$	space domain
$E$	expectation
$F$	direct numerical solver
$\mathcal{F}$	minimal $\sigma$ -algebra of $\Omega$
$H$	sensitivity matrix
$H_1$	$H_1$ norm
$J$	objective functional
$k$	thermal conductivity
$L$	length of Markov Chain
$L_2$	$L_2$ norm
$\mathbf{n}$	normal vector
$\mathcal{N}$	normal distribution
$p$	probability density function
$\mathcal{P}$	probability measure
$q$	heat-flux
$q_h$	known stochastic boundary heat-flux

$\hat{f}$	known stochastic boundary heat-flux
$q_0$	unknown stochastic boundary heat-flux
$R$	real number line
$t$	time
$dt$	discretization time step
$\Delta t$	sampling interval of thermocouple
$T$	temperature
$\mathcal{T}$	time domain
$w$	stochastic process in Section 2
	finite-element basis functions in Section 3
$W$	covariance matrix of MRF
$\mathbf{x}$	space vector
$X$	real-valued random variable
$Y$	thermocouple measurement
$\mathcal{J}(\theta)$	cost functional
$\mathcal{J}'(\theta)$	gradient of cost functional
$w_n(\mathbf{x}, \theta)$	tensor-product SSFEM weighting function

### *Greek Symbols*

$\alpha$	regularization constant
$\theta$	sample space variable in Section 2
	parametric form of $q_0$ in Section 3

$\lambda$	eigenvalue of KL expansion in Section 2
	scaling parameter of MRF in Section 3
$\sigma$	standard deviation
$\phi$	adjoint temperature
$\omega_m$	measurement noise
$\Gamma$	boundary of space domain
$\Gamma_0$	boundary with unknown heat-flux
$\Gamma_g$	boundary with known temperature
$\Gamma_h$	boundary with known heat-flux
$\Gamma_I$	internal surface
$\Theta$	temperature sensitivity field
$\Phi$	kernel function of MRF
$\Omega$	probability sample space
$\psi_i(\theta)$	Wiener-Askey orthogonal polynomials

## References

- [1] D.L. Powers, *Boundary Value Problems*, 4th ed., Academic Press, London, 1999.
- [2] M.N. Ozisik, *Boundary Value Problems of Heat Conduction*, Dover, Mineola, NY, 1989.
- [3] S. Kaka and Y. Yener, *Heat Conduction*, 3rd ed., Taylor & Francis, Washington DC, 1993.
- [4] L.C. Evans, *Partial Differential Equations*, American Mathematical Society, Providence, RI, 1998.
- [5] J. Hadamard, *Lectures on Cauchy's Problem in Linear Differential Equations*, Yale University Press, New Haven, CT, 1923.
- [6] N. Zabararas, Y. Ruan and O. Richmond, On the Design of Two-dimensional Stefan Processes with Desired Freezing Front Motions, *Numerical Heat Transfer, Part B: Fundamentals*, 21: 307–325 (1992).
- [7] J.V. Beck, B. Blackwell and C. St. Clair, *Inverse Heat Conduction: Ill-posed Problems*, Wiley, New York, 1985.
- [8] J.V. Beck and K.J. Arnold, *Parameter Estimation in Engineering and Science*, Wiley, New York, 1977.
- [9] R. Siegel and J.R. Howell, *Thermal Radiation Heat Transfer*, 4th ed., Taylor & Francis, New York, 1992.
- [10] N.J. McCormick, Inverse Radiative Transfer Problems: A Review, *Nuclear Science and Engineering*, 112: 185–198 (1992).

- [11] S.M.H. Sarvari, J.R. Howell and S.H. Mansouri, Inverse Boundary Design Conduction-Radiation Problem in Irregular Two-dimensional Domains, *Numerical Heat Transfer, Part B: Fundamentals*, 44: 209–224 (2003).
- [12] O.M. Alifanov, *Inverse Heat Transfer Problems*, Springer-Verlag, Berlin, 1994.
- [13] D.A. Murio, *The Mollification Method and the Numerical Solution of Ill-posed Problem*, Wiley, New York, 1993.
- [14] N. Zabararas and S. Kang, On the Solution of an Ill-posed Inverse Design Solidification Problem Using Minimization Techniques in Finite and Infinite Dimensional Spaces, *International Journal for Numerical Methods in Engineering*, 36: 3973–3990 (1993).
- [15] R. Sampath and N. Zabararas, A Functional Optimization Approach to An Inverse Magneto-convection Problem, *Computer Methods in Applied Mechanics and Engineering*, 190: 2063–2097 (1990).
- [16] A.N. Tikhonov, *Solution of Ill-posed Problems*, Halsted, Washington, DC, 1977.
- [17] V.A. Morozov, *Regularization Methods for Ill-posed Problems*, CRC Press, Boca Raton, FL, 1992.
- [18] N. Zabararas and J. C. Liu, An Analysis of Two-dimensional Linear Inverse Heat Transfer Problems Using an Integral Method, *Numerical Heat Transfer*, 13: 527–533 (1988).
- [19] J.P. Kaipio and E. Somersalo, *Computational and Statistical Methods for Inverse Problems*, Springer-Verlag, Berlin, 2004.
- [20] A.F. Emery, A.V. Nenarokomov, and T.D. Fadale, Uncertainties in Parameter Estimation: The Optimal Experimental Design, *International Journal of Heat and Mass Transfer*, 43: 3331–3339 (2002).

- [21] A.F. Emery, B.F. Blackwell, and K.J. Dowling, The Relationship between Information, Sampling Rates, and Parameter Estimation Models, *ASME Journal of Heat Transfer*, 124: 1192–1199 (2002).
- [22] K.A. Woodbury (Editor), *Inverse Engineering Handbook*, CRC Press, Boca Raton, FL, 2002.
- [23] J.V. Beck and B. Blackwell, Inverse problems, In W.J. Minkowycz, E.M. Sparrow, G.E. Schneider, and R.H. Pletcher (Editors), *Handbook of Numerical Heat Transfer*, Chapter 19, Wiley, New York, 1988.
- [24] M.N. Ozisik and H.R.B. Orlande, *Inverse Heat Transfer: Fundamentals and Applications*, Taylor & Francis, New York, 2000.
- [25] T. Hisada and S. Nakagiri, Stochastic Finite Element Method Developed for Structure, Safety and Reliability, *Proc. 3rd Int. Conf. on Struct. Safety and Reliability*, pp. 395–408, Wiley, New York, 1989.
- [26] T. Hisada and S. Nakagiri, Role of Stochastic Finite Element Methods in Structural Safety and Reliability, *Proc. 4th Int. Conf. on Struct Safety and Reliability*, pp. 385–394, 1985.
- [27] M. Shinozuka and G. Deodatis, Response Variability of Stochastic Finite Element Systems, *Journal of Engineering Mechanics*, ASCE, 114: 499–519 (1988).
- [28] D. Xiu, D. Lucor, C-H Su, and G.E. Karniadakis, Stochastic Modeling of Flow–structure Interactions Using Generalized Polynomial Chaos, *J. Fluids Engineering*, 124: 51–59 (2002).
- [29] M. Jardak, C-H Su, and G.E. Karniadakis, Spectral Polynomial Chaos Solutions of the Stochastic Advection Equation, *J. Sci. Comp.*, 17: 319–338 (2002).

- [30] D. Xiu and G.E. Karniadakis, Modeling Uncertainty in Flow Simulations via Generalized Polynomial Chaos, *Journal of Computational Physics*, 187: 137–167 (2003).
- [31] V.A. Badri Narayanan and N. Zabaras, Stochastic Inverse Heat Conduction Using a Spectral Approach, *International Journal for Numerical Methods in Engineering*, 60: 1569–1593 (2004).
- [32] V.A. Badri Narayanan and N. Zabaras, Variational Multiscale Stabilized FEM Formulations for Transport Equations: Stochastic Advection-diffusion and Incompressible Stochastic Navier-Stokes Equations, *Journal of Computational Physics*, 202: 94–133 (2005).
- [33] R.G. Ghanem and P.D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991.
- [34] R. Ghanem, Probabilistic Characterization of Transport in Heterogenous Media, *Computer Methods in Applied Mechanics and Engineering*, 158: 199–220 (1998).
- [35] R. Ghanem, Ingredients for a General Purpose Stochastic Finite Elements Implementation, *Computer Methods in Applied Mechanics and Engineering*, 168: 19–34 (1999).
- [36] D.W. Heermann and K. Binder, *Monte Carlo Simulation in Statistical Physics*, 4th ed., Springer-Verlag, New York, 2002.
- [37] C.Z. Mooney, *Monte Carlo simulation*, Sage, California, Thousand Oaks, 1997.
- [38] P. Bremaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues*, Springer-Verlag, New York, 1999.
- [39] C. Andrieu, N. Freitas, A. Doucet and M. Gordan, An Introduction to MCMC for Machine Learning, *Mach. Learning*, 50: 5–43 (2003).

- [40] C.P. Robert, *The Bayesian Choice, From Decision-Theoretic Foundations to Computational Implementation*, 2nd ed., Springer-Verlag, New York, 2001.
- [41] J. Besag, P. Green, D. Higdon, and K. Mengersen, Bayesian computation and stochastic systems, *Statistical Science*, 10: 3–41 (1995).
- [42] J. Moler (ed), *Spatial Statistics and Computational Methods*, Springer-Verlag, New York, 2003.
- [43] M. Loève, *Probability Theory*, 4th ed., Springer-Verlag, Berlin, 1977.
- [44] N. Wiener, The Homogenous Chaos, *American Journal of Mathematics*, 60: 897–936 (1938).
- [45] R. Cameron and W. Martin, The Orthogonal Development of Nonlinear Functionals in Series of Fourier-Hermite Functionals, *Annals of Mathematics*, 48: 385–392 (1947).
- [46] W. Schoutens, *Stochastic Processes and Orthogonal Polynomials*, Springer-Verlag, Berlin, 2000.
- [47] C. Ferrero and K. Gallagher, Stochastic Thermal History Modelling, 1. Constraining Heat Flow Histories and Their Uncertainty, *Marine and Petroleum Geology*, 19: 633–648 (2002).
- [48] J. Wang and N. Zabarar, A Bayesian Inference Approach to The Stochastic Inverse Heat Conduction Problem, *International Journal of Heat and Mass Transfer*, 47: 3927–3941 (2004).
- [49] A.F. Emery, Stochastic Regularization for Thermal Problems with Uncertain Parameters, *Inverse Problems in Engineering*, 9: 109–125 (2001).
- [50] C. Vogel, An Applied Mathematician’s Prospective on Regularization Methods, *Lecture in Opening Workshop for Inverse problem, Methodology in Complex*

*Stochastic Models, Session of Parameter Estimation and Inverse Problems Statistics Perspective*, Statistical and Applied Mathematical Sciences Institute at Duke University, 2002.

- [51] H.K.H. Lee, D.M. Higdon, Z. Bi, M.A.R. Ferreira, and M. West, Markov Random Field Models for High-dimensional Parameters in Simulations of Fluid Flow in Porous Media, *Technometrics*, 44: 230–241 (2002).
- [52] P. Brémaud, *Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues*, Springer–Verlag, New York, 1999.
- [53] I. Beichl and F. Sullivan, The Metropolis algorithm, *Computing in Science and Engineering*, 2: 65–69 (2000).
- [54] S.P. Brooks and G.O. Roberts, Convergence assessment techniques for Markov Chain Monte Carlo, *Statistics and Computing*, 8: 319–335 (1998).
- [55] J. Wang and N. Zabaras, Using Bayesian statistics in the estimation of heat source in radiation, *International Journal of Heat and Mass Transfer*, 48: 15–29 (2005).
- [56] P. Holmes, J.L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge, 1998.

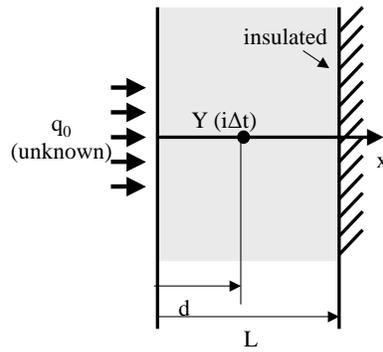


Figure 17.1 One-dimensional inverse heat conduction example. Temperature data  $Y$  are collected in the location  $d$  and one is interested in estimating the temporal history of the boundary heat flux  $q_0$ . For simplicity of the presentation, the right wall is assumed to be insulated.

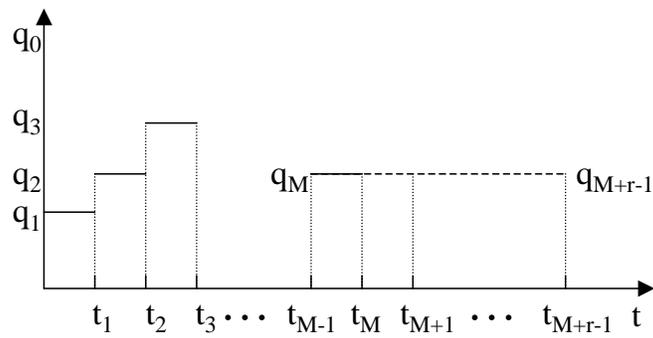


Figure 17.2 The essence of the sequential function specification of Beck. Piecewise heat-flux variations in time are taken and the calculation of  $q_M$  proceeds sequentially using future collected temperature data as indicated in Eq. (17.3).

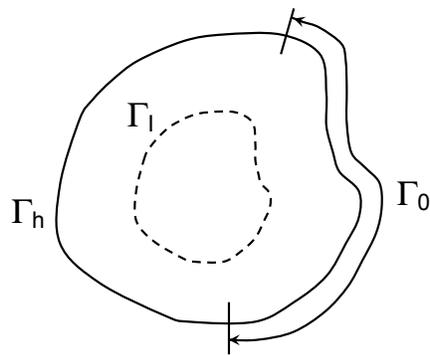


Figure 17.3 Schematic of the stochastic inverse-design heat-conduction problem (SIHCP). The stochastic boundary heat-flux  $q_0$  is not known on  $\Gamma_0$  but the stochastic measured/desired temperature is known on the surface  $\Gamma_l$  within the domain.

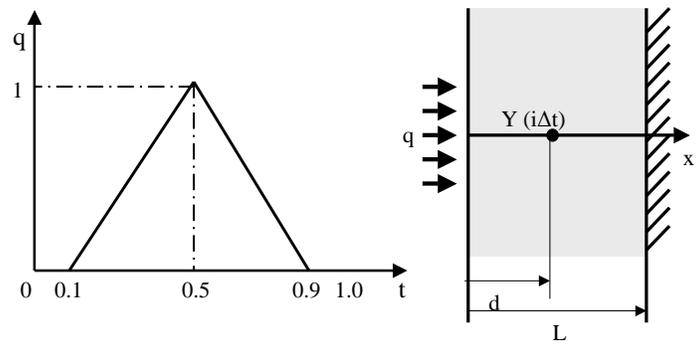


Figure 17.4 One-dimensional inverse heat-conduction example. One is interested in computing the temporal history of the boundary heat flux given temperature measurements at the sensor placed at  $x = d$ .

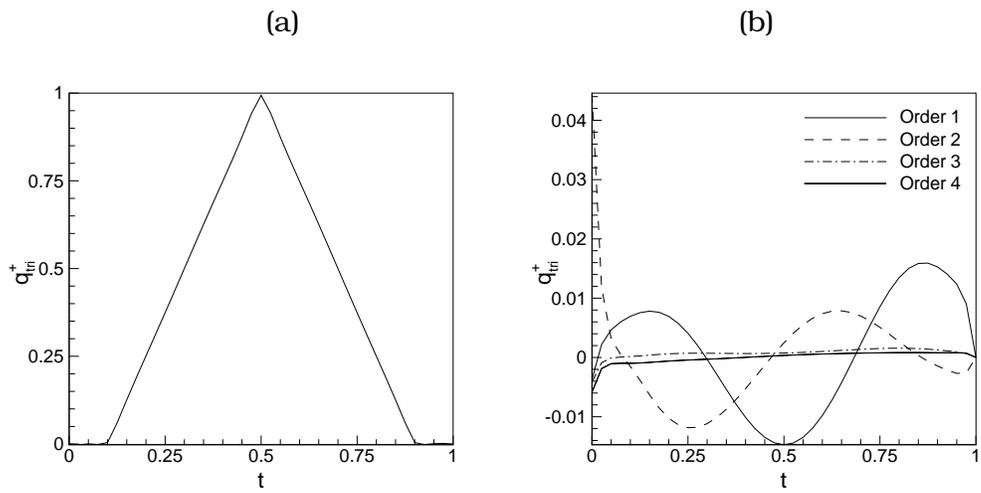


Figure 17.5 Computed (a) heat-flux mean and (b) higher-order polynomial chaos terms for case I.

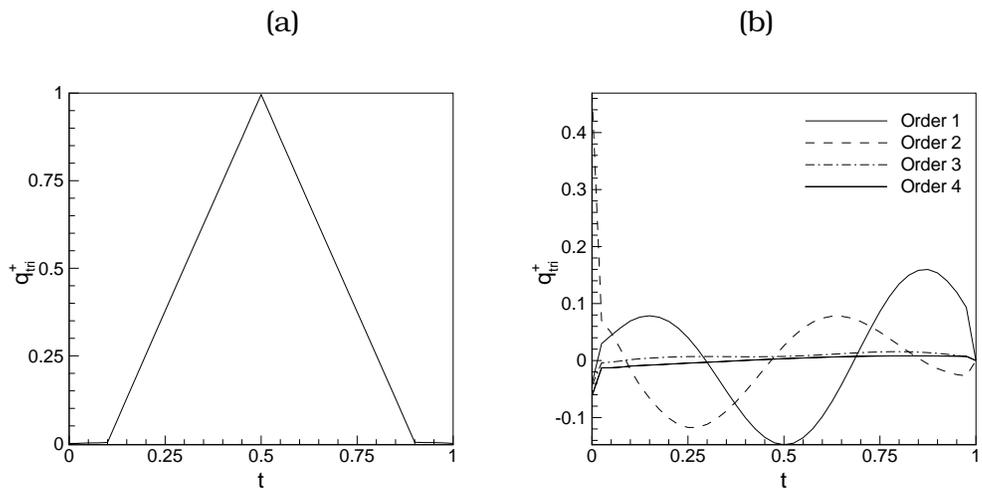


Figure 17.6 Computed (a) heat-flux mean and (b) higher-order polynomial chaos terms for case II.

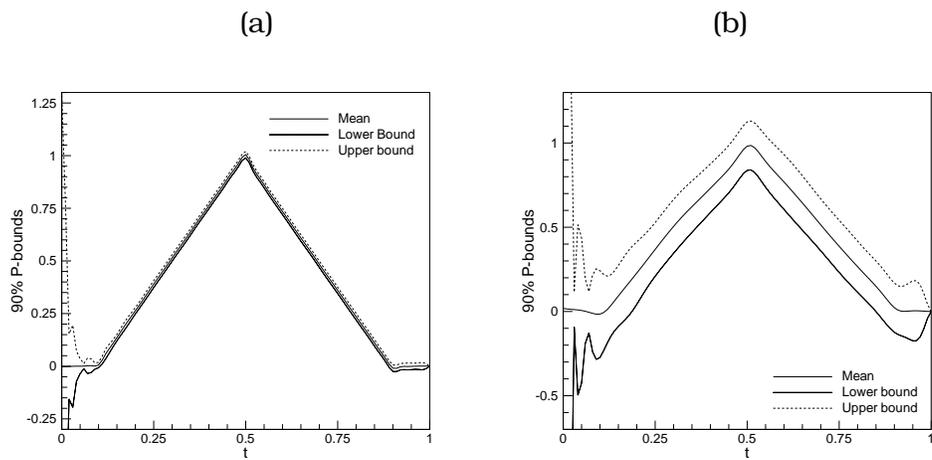


Figure 17.7 Computed probability bounds for triangular heat-flux estimation for (a) case I (b) case II.

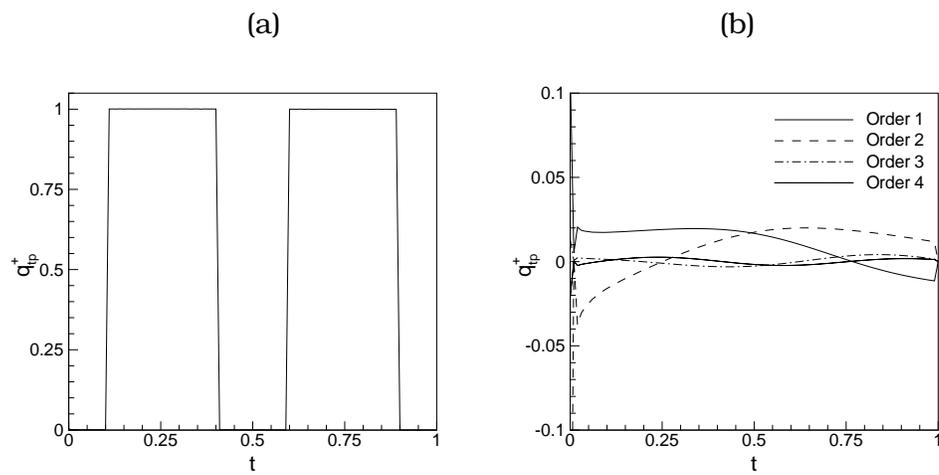


Figure 17.8 Computed (a) heat-flux mean and (b) higher-order polynomial chaos terms for the impulse flux problem for case I (low noise).

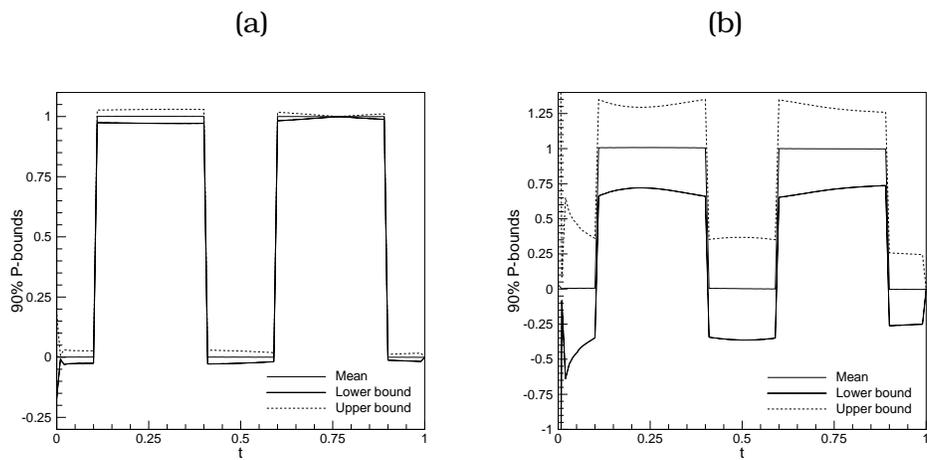


Figure 17.9 Computed probability bounds for square pulse heat-flux estimation: (a) case I, (b) case II.

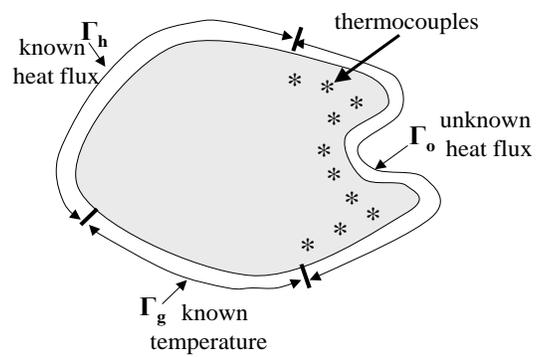


Figure 17.10 Schematic of the inverse heat-conduction problem (IHCP) with discrete temperature measurements.

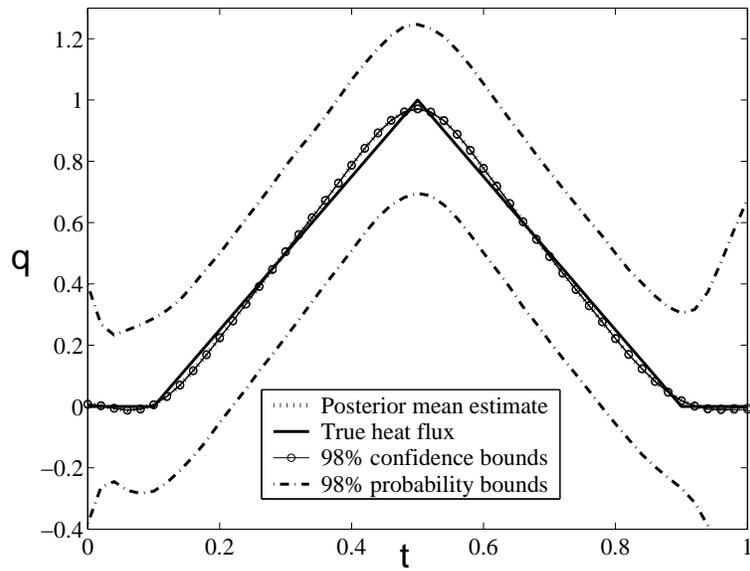


Figure 17.11 Posterior mean estimate and associated statistical bounds of  $q$  when  $d = 0.5$ ,  $\sigma = 0.01$ ,  $\Delta t = 0.01$ ,  $dt = 0.02$ , regularization constant is  $2.5e-3$ .

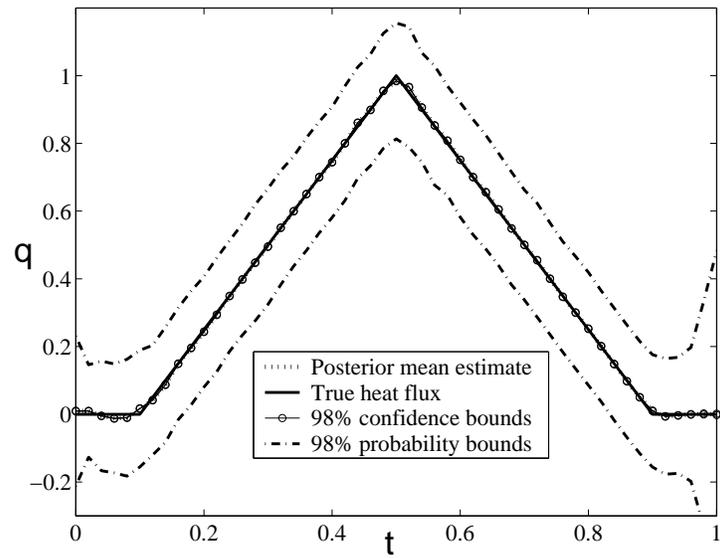


Figure 17.12 Posterior mean estimate and associated statistical bounds of  $q$  when  $d = 0.5$ ,  $\sigma = 0.001$ ,  $\Delta t = 0.01$ ,  $dt = 0.02$ , regularization constant is  $2.5e-5$ .

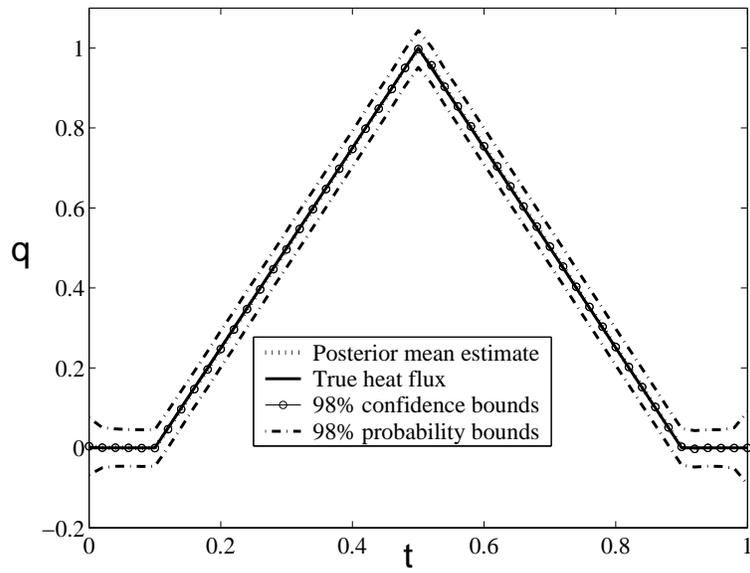


Figure 17.13 Posterior mean estimate and associated statistical bounds of  $q$  when  $d = 0.1$ ,  $\sigma = 0.001$ ,  $\Delta t = 0.01$ ,  $dt = 0.02$ , regularization constant is  $1.0e-4$ .

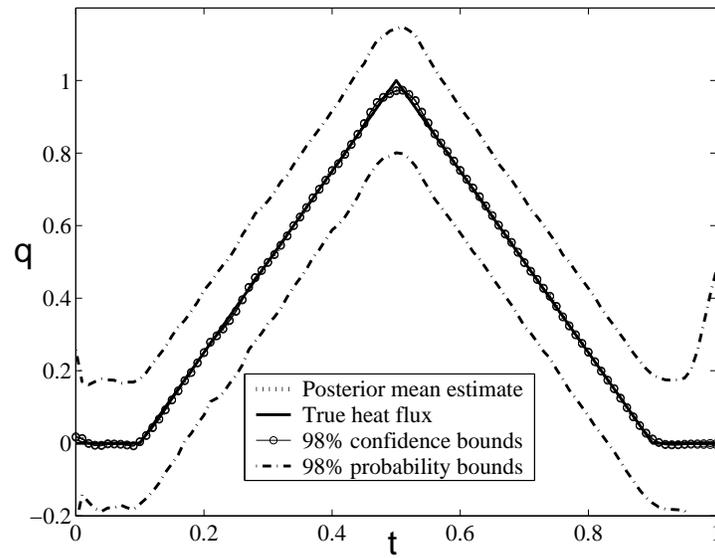


Figure 17.14 Posterior mean estimate and associated statistical bounds of  $q$  when  $d = 0.5$ ,  $\sigma = 0.001$ ,  $\Delta t = 0.01$ ,  $dt = 0.01$ , regularization constant is  $1.0e-4$ .

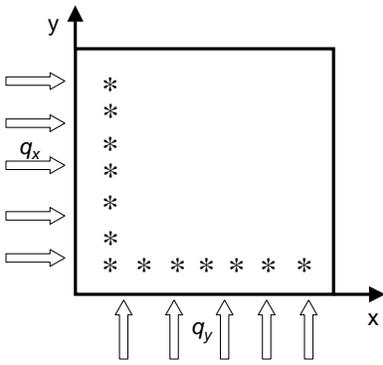


Figure 17.15 Thermocouple location in example 2.

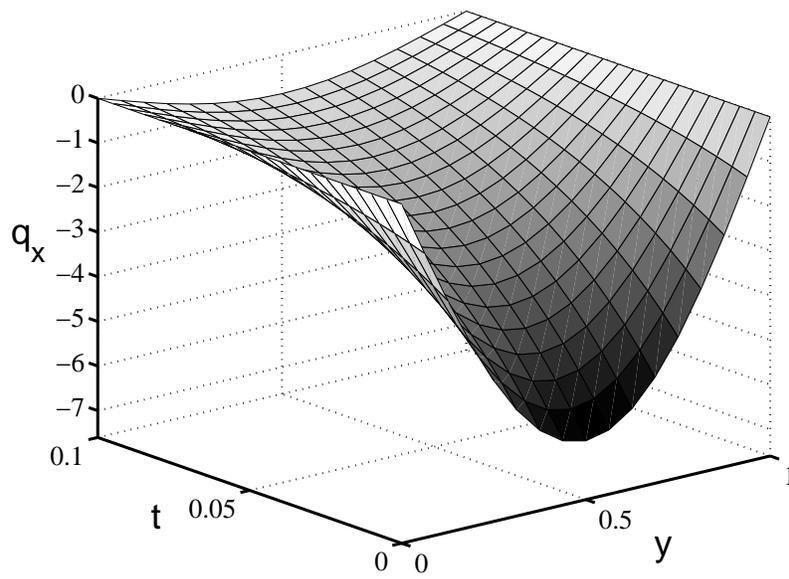


Figure 17.16 True heat flux along the  $y$  axis.

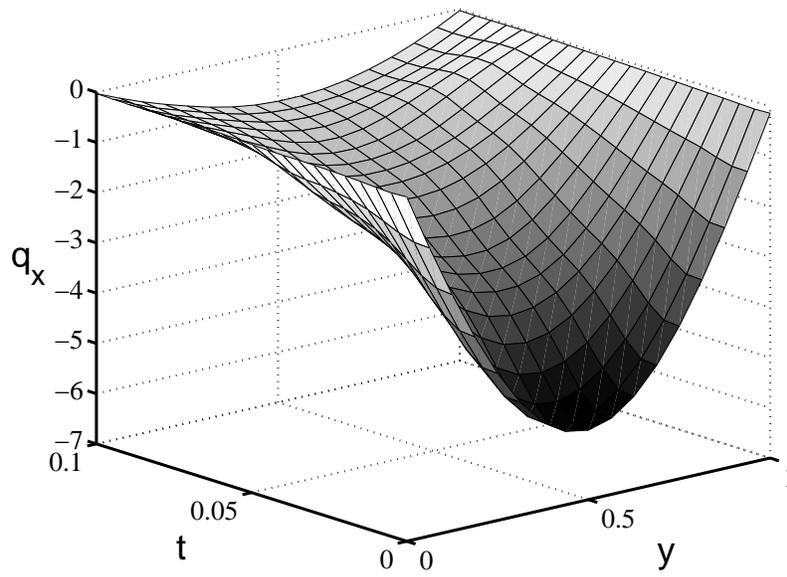


Figure 17.17 Posterior mean estimate of heat flux along the  $y$  axis.

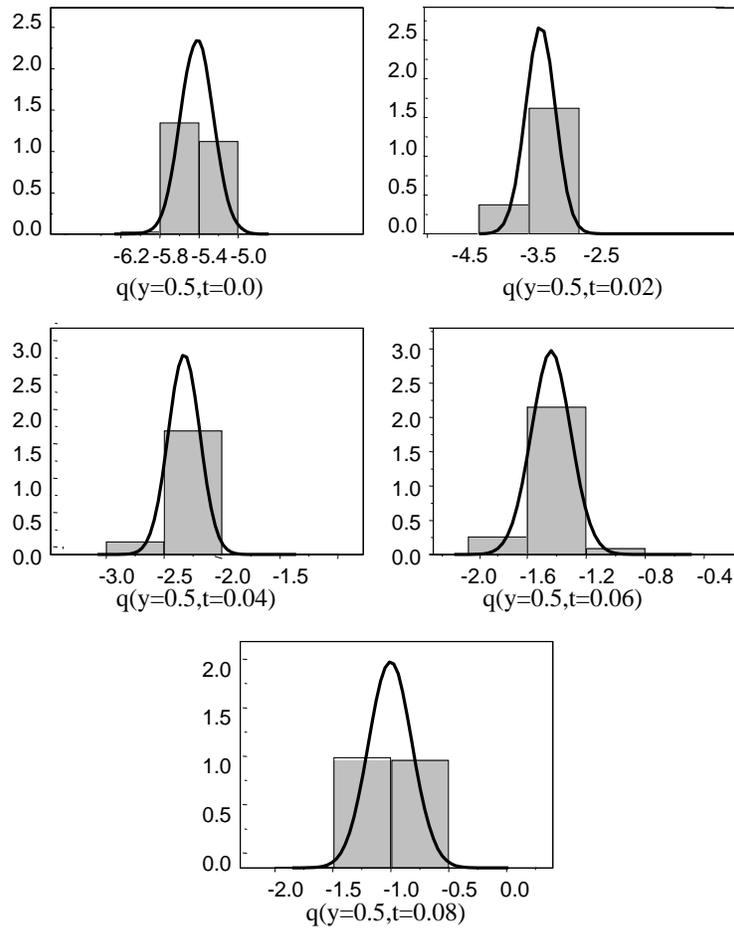


Figure 17.18 Marginal PDFs of heat flux at  $y = 0.5$  at different times.