Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Heat and Mass Transfer 47 (2004) 3927–3941
International Journal of HEAT and MASS TRANSFER
www.elsevier.com/locate/ijhmt

A Bayesian inference approach to the inverse heat conduction problem

Jingbo Wang, Nicholas Zabaras *

Materials Process Design and Control Laboratory, Sibley School of Mechanical and Aerospace Engineering, 188 Frank HT Rhodes Hall, Cornell University, Ithaca, NY 14853-3801, USA

Received 28 July 2003; received in revised form 10 February 2004
Available online 6 May 2004

Abstract

A Bayesian inference approach is presented for the solution of the inverse heat conduction problem. The posterior probability density function (PPDF) of the boundary heat flux is computed given temperature measurements within a conducting solid. Uncertainty in temperature measurements is modeled as stationary zero-mean white noise. The inverse solution is obtained by computing the expectation of the PPDF. The posterior state space is exploited using Markov chain Monte Carlo (MCMC) algorithms in order to obtain estimates of the statistics of the unknown heat flux. The MCMC sampling strategy enables the extension of the Bayesian inference approach to inverse problems having high-dimensional, non-standard distribution, and/or complex PPDFs. The ill-posedness (un-identifiability) of the inverse problem is cured through prior distribution modeling (Bayesian prior regularization) of the unknown heat flux. A special model of Markov random field (MRF) is adopted for prior distribution modeling of the unknown heat flux. An augmented Bayesian model is proposed for estimating the statistics of the measurement noise as well as the unknown heat flux. Two inverse heat conduction examples are presented to demonstrate the potential of the MCMC-based Bayesian approach. The simulation results indicate that MRF provides an effective prior regularization, the estimates using MCMC samples are accurate and the Bayesian approach captures very well the probability distribution of the unknown heat flux.

1. Introduction

In the inverse heat conduction problem (IHCP), one is interested in identifying the heat flux on part of the boundary given sufficient conditions on the remaining of the boundary and temperature measurements within the domain of a conducting solid. Inverse heat conduction is of interest in a wide range of scientific and engineering areas including manufacturing process control, metallurgy, chemical, aerospace and nuclear engineering, food science, medical diagnostics, etc. The main characteristic of the IHCP versus a well-posed direct heat conduction problem is that it leads to solutions that generally are not unique or stable to small changes in the given data (Alifanov [1]).

A number of solution methods has been developed for the IHCP. No attempt is given here to review these techniques and the interested reader can consult Alifanov [1] and Beck et al. [2]. The majority of the approaches restate the problem as a least-squares minimization problem over the whole-time domain or in sequential time intervals. The cost functional is usually stated as an appropriate norm of the deviation between computed temperature at the sensor locations for a guessed heat flux and the given temperature measurements. Gradient optimization techniques are introduced for performing the optimization process using either a finite or an infinite dimensional representation of the unknown heat flux. In addition to the direct model,

* Corresponding author. Tel.: +1-607-255-9104; fax: +1-607-255-9410.
E-mail address: zabaras@cornell.edu (N. Zabaras).

0017-9310/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijheatmasstransfer.2004.02.028

Nomenclature

c	normalizing constant of PDF
C_p	thermal capacity
d	distance of thermocouple to boundary
E	expectation
f	arbitrary function
F	direct system solver
H	sensitivity matrix
I	identity matrix
$\mathscr{J}$	objective function
k	thermal conductivity
l	likelihood function
L	length scale of domain, length of Markov chain in Section 4
L_2	square-integrable function space
m	dimension of θ
M	number of thermocouples
n	total number of measurements
$\mathbf{n}$	unit normal vector
N	number of measurement steps
p	probability density function (PDF)
q	heat flux
q_h	known boundary heat flux
q_0	unknown boundary heat flux
$\bar{q}$	deterministic inverse solution
r	relative error of estimation
t	time
t_{max}	time scale
$\hat{t}$	time of measurement
Δt	thermocouple sampling interval
$\mathrm{d}t$	time interval in discretization of the inverse solution
T	temperature
T_g	known boundary temperature
T_H	homogeneous part of direct solution
T_I	inhomogeneous part of direct solution
T_0	known initial temperature
u	random number
U	uniform distribution
w	basis function of inverse solution
W	covariance matrix of MRF
$\hat{\mathbf{x}}$	location of thermocouple
Y	temperature measurement vector

Greek symbols

α	regularization parameter
β	parameter of exponential distribution
γ	scaling constant of general MRF
$\delta(\cdot)$	Kronecker delta function
θ	parameter form of unknown heat flux
$\hat{\theta}$	estimate of θ
λ	scaling constant of Gauss MRF
ρ	material density
σ	standard deviation
ω_m	measurement noise
Γ	boundary
Γ_g	Dirichlet boundary
Γ_h	Neuman boundary
Γ_0	boundary with unknown heat flux
Φ	kernel function of MRF
Ω	physical domain

Superscripts

T	transpose
(i)	ith iteration or ith time step
*	candidate
+	non-dimensional quantity

Subscripts

aug	augmented model
i	ith component
$-i$	full conditional set
$i \sim j$	site neighborhood
max	maximum
min	minimum
med	median
LS	least squares
MLE	maximum likelihood estimator
MAP	maximum a posteriori

appropriate continuum or discrete sensitivity and/or adjoint problems are usually required in such approaches. The ill-posedness of the resulting formulations is addressed using combination of appropriate regularization techniques including regularization by Tikhonov [3], the future information method (function specification method) by Beck et al. [2], Zabaras and Liu [4] or the iterative regularization techniques by Alifanov [1] and Sampath and Zabaras [5].

The above deterministic inverse techniques based on exact matching or least-squares optimization, lead to point estimates of unknowns without rigorously considering the statistical nature of system uncertainties and without providing quantification of the uncertainty in the inverse solution. The existing methods for uncertainty quantification that can be applied in heat transfer analysis can be divided into deterministic and probabilistic methods as in Narayanan and Zabaras [6]. The deterministic methods, including interval analysis and propagation using sensitivity derivatives, are easy to implement but not sufficiently accurate as they are not able to provide higher order statistics of the estimates

[6]. The reliability method, which is a probabilistic technique, is able to provide characterization of uncertainties up to the second statistical moment, however, it is not sufficient for complex problems. Two other more promising probabilistic methods are the spectral stochastic finite element method (SSFEM) [6] and Monte Carlo simulation based methods as discussed by Liu [7]. In this paper, Bayesian inference is explored for the solution of the IHCP and for uncertainty quantification using an implementation based on advanced Monte Carlo techniques.

A frequentist approach to the IHCP based on the maximum likelihood estimator (MLE) is often used in parameter estimation problems. In MLE, the probability of observations (temperature measurements) given the inverse solution is maximized. This probability is called likelihood and is also incorporated here in the Bayesian formulation to update the prior distribution. Recently, Fadale et al. [8] and Emery et al. [9] presented an extended MLE approach to investigate the system uncertainties for parameter identification in heat transfer problems. MLE takes into account the statistics of uncertainties but excludes the prior knowledge of unknowns, hence, it often leads to an ill-posed problem and fails to provide smooth solutions when it is extended to function estimation.

The Bayesian inference approach can provide a solution to the IHCP by formulating a complete probabilistic description of the unknowns and uncertainties given temperature data. The Bayesian approach incorporates the known information regarding the unknown heat flux and system uncertainties into a prior distribution model that is then combined with the likelihood to formulate the posterior probability density function (PPDF) [10,11].

It regularizes the ill-posed IHCP through prior distribution modeling (Emery [12], Vogel [13]), and in addition provides means to estimate the statistics of uncertainties. In this work, emphasis is given on the prior distribution modeling techniques based on a special model of Markov random field (MRF) described by Besag et al. in [11] and Besag and Kooperberg [14]. Although the Bayesian inference technique has been developed over a long period, only with the recent propagation of efficient sampling methods, such as Markov chain Monte Carlo (MCMC), its application to engineering inverse problems becomes feasible.

The plan of the remaining of this paper is as follows. Section 2 provides a definition of the IHCP. Section 3 introduces the minimum required background on the generic Bayesian inference approach, discussing the roles of likelihood and prior distribution model. Section 3 also provides a formulation of the Bayesian model for the IHCP with a review of the relations between Bayesian prior distribution modeling and Tikhonov regularization theory. Section 4 discusses in details the numerical implementation of Bayesian estimation by MCMC. Finally, Section 5 demonstrates the potential of the developed Bayesian approach through the solution of one- and two-dimensional inverse heat conduction problems.

2. The inverse heat conduction problem

The inverse heat conduction problem of interest is defined with the following equations (see Fig. 1):

$$\rho C_p \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T), \quad \text{in } \Omega, \ \ t \in [0, t_{\max}] \tag{1}$$

$$T(\mathbf{x}, t) = T_{\mathrm{g}}, \quad \text{on } \Gamma_{\mathrm{g}}, \ \ t \in [0, t_{\max}] \tag{2}$$

$$k \frac{\partial T(\mathbf{x}, t)}{\partial \mathbf{n}} = q_{\mathrm{h}}, \quad \text{on } \Gamma_{\mathrm{h}}, \ \ t \in [0, t_{\max}] \tag{3}$$

$$T(\mathbf{x}, 0) = T_0(\mathbf{x}), \quad \text{in } \Omega \tag{4}$$

where ρ, C_p, k denote the density, heat capacity and thermal conductivity, respectively. The known boundary conditions include the heat flux q_{h} and the temperature T_{g} on the subsets Γ_{h} and Γ_{g}, respectively, of Γ. Finally, T_0 is the known initial temperature field.

The main objective of the IHCP is the calculation of the heat flux q_0 on $\Gamma_0 \times [0, t_{\max}]$ ($\Gamma_{\mathrm{g}} \cup \Gamma_{\mathrm{h}} \cup \Gamma_0 = \Gamma$):

$$k \frac{\partial T(\mathbf{x}, t)}{\partial \mathbf{n}} = q_0, \quad (\text{unknown}) \text{ on } \Gamma_0, \ \ t \in [0, t_{\max}] \tag{5}$$

In addition to the conditions of Eqs. (1)–(4), approximations of the temperature field are available at M points (temperature sensor locations) within the domain:

$$Y_i^{(j)} \simeq T(\hat{\mathbf{x}}_i, \hat{t}_j), \quad i = 1, \ldots, M, \ \ j = 1, \ldots, N \tag{6}$$

Eqs. (1)–(5) 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

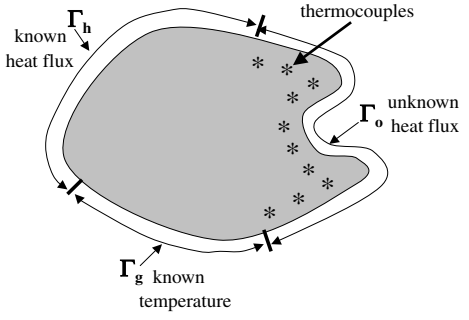

Fig. 1. Schematic of the inverse heat conduction problem (IHCP). The objective is to compute the boundary heat flux q_0 on Γ_0 given appropriate boundary conditions on the remaining of the boundary and temperature measurements at a number of points within the domain.

problem as $T(\mathbf{x},t;q_0)$. In the IHCP, one is interested to compute 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. (6)). Let us denote with $F(q_0)$ the vector of computed temperatures at the sensor locations $\hat{\mathbf{x}}_i$, $i=1,\ldots,M$, at times $\hat{t}_j$, $j=1,\ldots,N$ ($\hat{t}_N=t_{\max}$). 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. 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 is approximated as follows:

$$q_0(\mathbf{x},t)=\sum_{i=1}^{m}\theta_i w_i(\mathbf{x},t) \tag{7}$$

where w_i's are pre-defined basis functions. The IHCP is then transformed to the estimation of the weights θ_i's. These weights are considered to be represented by an unknown random vector θ of length m.

Let us denote with ω_{m} the sensor uncertainty (sensor noise). Then one looks for the vector θ such that:

$$Y\simeq F(\theta)+\omega_{\mathrm{m}} \tag{8}$$

The statistics of noise may be known a priori or unknown as in most real-world engineering systems. In both cases, either in order to improve knowledge of noise distribution or to capture this (unknown) information, the parameters controlling noise distribution can be incorporated into the problem unknowns. 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 ω_{m} is determined by a single unknown parameter σ, which is the standard deviation (std). An augmented unknown vector $\theta_{\mathrm{aug}}=[\,\theta\quad\sigma\,]^{\mathrm{T}}$ can be formed to be estimated from the observation vector Y. In principle, other system uncertainties can be treated as unknowns and be computed from the given data as well.

Direct inversion of Eq. (8) to compute θ 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 [3]. In particular, one looks for a flux $\bar{q}_0(\mathbf{x},t)\in L_2(\Gamma_0\times[0,t_{\max}])$ such that:

$$\mathscr{J}(\bar{q}_0)\leqslant\mathscr{J}(q_0),\quad\forall q_0\in L_2(\Gamma_0\times[0,t_{\max}]) \tag{9}$$

where, $L_2(\Gamma_0\times[0,t_{\max}])$ is the space of all square integrable functions defined over the spatial and temporal domains Γ_0 and $[0,t_{\max}]$, respectively. The objective function $\mathscr{J}(q_0)\equiv\mathscr{J}(\theta)$ to be minimized is usually chosen as the L_2 norm of the error between the estimated and measured temperatures along the sensor locations:

$$\begin{aligned}\mathscr{J}(q_0)&=\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{N}\{T(\hat{\mathbf{x}}_i,\hat{t}_j;q_0)-Y(\hat{\mathbf{x}}_i,\hat{t}_j)\}^2\\&=\frac{1}{2}\|F(\theta)-Y\|_{L_2}^2\end{aligned} \tag{10}$$

The discrete L_2 norm is introduced above to simplify in the following analysis the notation of the cost functional.

3. A Bayesian inference approach to the IHCP

3.1. Bayesian inference fundamentals

Bayesian statistics study the probability of a hypothesis from both current achieved information and previous knowledge. The basis of Bayesian inference is the Bayes' formula:

$$p(\theta|Y)=\frac{p(Y|\theta)p(\theta)}{p(Y)}=\frac{1}{c}p(Y|\theta)p(\theta) \tag{11}$$

θ is used here to represent a hypothesis and Y stands for observation related to this hypothesis. $p(\theta|Y)$, $p(Y|\theta)$ and $p(\theta)$ are called posterior probability density function (PPDF), likelihood function and prior probability density function (PDF), respectively. Based on Eq. (11), the posterior probability of a hypothesis given some observations (evidence) is proportional to the product of its likelihood and prior (unconditional) probability. Bayesian inference provides powerful techniques for estimation, hypothesis testing, model selection and decision problems [10].

A modern outlook on Bayesian estimation that is adopted here is that of a supervised regression method of statistical learning. In this sense, the vector parameter θ representing the unknown heat flux in the IHCP can be treated as a random process and Bayesian estimation can be used to compute its joint PPDF.

From Eq. (11), it is seen that all the information available about θ is contained in the posterior PDF $p(\theta|Y)$. Once $p(\theta|Y)$ is formulated, one can compute the estimates of θ (inverse solution) and corresponding probability bounds. The most common estimates based on this posterior PDF are the 'posterior mean estimate'

$$\hat{\theta}_{\mathrm{postmean}}=E\theta|Y \tag{12}$$

and the 'maximum a posteriori' (MAP) estimate:

$$\hat{\theta}_{\mathrm{MAP}}=\mathrm{augmax}_{\theta}\,p(\theta|Y) \tag{13}$$

Note that it is not necessary to compute the normalizing constant c in Eq. (11) when exploiting the posterior state space. This greatly simplifies the analysis as it is not easy to find the marginal distribution of Y.

3.2. Prior distribution modeling and Markov random field

In this work, the only uncertainty considered is the measurement error, which is modeled as a Gauss random variable. Thus, the prior is selected from the exponential conjugate family. Also since the main unknown in the IHCP is the heat flux q_0 that is generally a function of space and time, Markov random field (MRF) is considered here as a suitable choice for prior distribution modeling. The model described in [14,15] is used that originated in spatial statistics [16,17].

MRF has been successfully used for prior distribution modeling in numerous applications [11,14,15,18]. In this work, MRF is introduced for simultaneous regularization in space and time by treating time simply as another spatial dimension. Therefore, in discussion of the MRF, the unknown θ is treated as spatially distributed quantity.

MRF is a point-pair spatial model of θ. It takes the general form:

$$p(\theta) \propto \exp\left\{ -\sum_{i\sim j} W_{ij}\Phi(\gamma(\theta_i - \theta_j)) \right\} \tag{14}$$

in which θ_i is the unknown variable at spatial site i, γ is a scaling parameter and Φ is an even function that determines the specific form of the MRF. The summation in Eq. (14) is over all pairs of sites $i \sim j$ that are neighbors and W_{ij}'s are specified nonzero weights [11]. In general, the neighbors to a particular nodal heat flux at a given time refer to nodal heat fluxes at adjacent nodal points and time steps. A wide range of different MRFs can be found in [11,19].

The Φ used in the current study is in the form $\Phi(u) = \frac{1}{2}u^2$, which is a widely used model in spatial problems. The MRF then can be rewritten as:

$$p(\theta) \propto \lambda^{m/2} \exp\left(-\frac{1}{2}\lambda\theta^{\mathrm{T}} W \theta \right) \tag{15}$$

In the one-parameter model of Eq. (15), 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. W determines the dependence between components of θ and λ controls the scale on which the random vector is distributed. This simple form of MRF has been reported to be effective in a number of applications [11,14,20].

The prior introduced by the above MRF model is invariant under space shift, therefore, its highest density region (HDR) is able to capture the entire state space and it will not over-constrain the likelihood. In addition, it is able to model various dependence relations between variables by changing the form of W. The single impropriety in this prior is removed from the corresponding posterior distribution by the presence of any informative data [11].

The scaling parameter λ is also of great importance. As it will be discussed later in this paper, λ controls the strength of spatial dependence and regularization to the inverse problem.

3.3. Regularization in Bayesian inverse approach

Tikhonov regularization modifies the original least-squares problem (Eq. (10)) as follows:

$$\hat{\theta}_{\mathrm{LS}} = \mathrm{augmin}_\theta \left\{ \frac{1}{2}\|F(\theta) - Y\|_{L_2}^2 + \alpha\|\theta\|_{\mathrm{p}}^2 \right\} \tag{16}$$

where $\hat{\theta}_{\mathrm{LS}}$ is the deterministic estimate of θ, α is a regularization parameter and $\|\cdot\|_{\mathrm{p}}$ represents different norms in the θ-parameter space.

To clarify the relation between the above Tikhonov regularization and the Bayesian prior regularization induced by Eq. (15), it is assumed that ω_{m} is Gauss white noise with known standard deviation σ. Then the likelihood function is the following:

$$l(\theta|Y) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{ -\frac{(F(\theta) - Y)^{\mathrm{T}}(F(\theta) - Y)}{2\sigma^2} \right\} \tag{17}$$

where $n = M \times N$ is the length of Y. The posterior PDF of θ can then be written as follows:

$$p(\theta|Y) \propto \exp\left\{ -\frac{(F(\theta) - Y)^{\mathrm{T}}(F(\theta) - Y)}{2\sigma^2} \right\} \cdot \lambda^{m/2} \cdot \exp\left(-\frac{1}{2}\lambda\theta^{\mathrm{T}} W \theta \right) \tag{18}$$

The MAP estimate of θ can then be derived as:

$$\hat{\theta}_{\mathrm{MAP}} = \mathrm{augmin}_\theta \left\{ \frac{1}{2}(F(\theta) - Y)^{\mathrm{T}}(F(\theta) - Y) + \frac{\lambda\sigma^2}{2}\theta^{\mathrm{T}} W \theta \right\} \tag{19}$$

By comparing Eqs. (16) and (19), it is seen that the least-squares estimator with Tikhonov regularization and the MAP estimator have similar mathematical forms. For example, by choosing $\lambda = \frac{2\alpha}{\sigma^2}$ and W as an identity matrix, Eq. (19) becomes identical to the zeroth-order Tikhonov regularization approach.

It is now clear that in Bayesian formulation, λ plays the role of a regularization constant when σ is known. Different W's and Φ's can be proposed for different problems. One can in principle derive a MRF model to emulate various norms in the parameter space used in Tikhonov regularization. In either approach, selection of the regularization parameter, α or equivalently of $\frac{1}{2}\lambda\sigma^2$, is important. There are in general three approaches for determining an optimal value of regularization parameter. One is the so called 'unbiased predictive risk estimator' (UPRE) method [13]. Another approach is

the use of a hierarchical Bayesian model [15] when σ is known. In this case, the problem is modeled in a more flexible way and one is able to investigate uncertainty in spatial dependence as well as to find the optimal regularization parameter. The third way, as used in this work, is the heuristic Tikhonov method. Here, the inverse problem is solved with a set of regularization parameters. Within a certain range of the regularization parameter, the obtained inverse solution remains practically unchanged. The regularization parameter is then chosen from this range.

In summary, there are two formulations that will be used for the solution of the IHCP.

Formulation I—known σ and user pre-determined α:

$$p(\theta|Y) \propto \exp\left\{-\frac{1}{2\sigma^2}[F(\theta)-Y]^{\mathrm{T}}[F(\theta)-Y]\right\} \cdot \exp\left(-\frac{\alpha}{\sigma^2}\theta^{\mathrm{T}}W\theta\right) \tag{20}$$

Formulation II—unknown σ:

$$p(\theta,\sigma^2|Y) \propto (\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sigma^2}[F(\theta)-Y]^{\mathrm{T}}[F(\theta)-Y]\right\} \cdot (\sigma^2)^{-m/2}\exp\left(-\frac{\alpha}{\sigma^2}\theta^{\mathrm{T}}W\theta\right)p(\sigma^2) \tag{21}$$

where $p(\sigma^2)$ is the prior distribution of σ^2.

To explore the posterior state space numerically, MCMC is employed next.

4. Markov chain Monte Carlo

Numerical sampling methods are needed because the PPDF may have a non-standard form, the problem may be nonlinear or has an implicit likelihood. In addition, the posterior state space in general has high dimension and one needs to obtain the marginal distributions of individual components. The most widely adopted numerical method for exploring the posterior state space is Monte Carlo simulation (MCS), which approximates the true expectation of a function of θ by the sample mean. MCS is based upon large sample set from the target PDF (here the posterior PDF $p(\theta|Y)$). Various sampling strategies have been proposed [21]. Among these techniques, Markov chain Monte Carlo (MCMC) is the most powerful one [21,22]. In the remaining of this section, θ denotes the unknown random vector, $p(\theta)$ denotes any probability density function of θ (not only the prior) and f denotes any function of θ.

4.1. Monte Carlo principle

The idea of Monte Carlo simulation is to draw an identical independently distributed (iid) set of samples $\{\theta^{(i)}\}_{i=1}^{L}$ from a target PDF $p(\theta)$ defined on a high dimensional space R^m [21]. These L samples can be used to approximate the target density with the following empirical point-mass function:

$$p_L(\theta) = \frac{1}{L}\sum_{i=1}^{L}\delta_{\theta^{(i)}}(\theta) \tag{22}$$

where $\delta_{\theta^{(i)}}(\theta)$ denotes the delta-Dirac mass located at $\theta^{(i)}$. Consequently, one can approximate the expectation of any function f of θ by its mean as follows:

$$E_L(f) = \frac{1}{L}\sum_{i=1}^{L} f(\theta^{(i)}) \mapsto_{L\to\infty} E(f) = \int f(\theta)p(\theta)\,\mathrm{d}\theta \tag{23}$$

By the strong law of large numbers, $E_L(f)$ converges to $E(f)$. When $f(\theta)=\theta$, one is able to use Eq. (23) to compute the posterior mean estimate of θ. The L samples can also be used to obtain the MAP estimate of θ as follows:

$$\hat{\theta}_{\mathrm{MAP}} = \mathrm{argmax}_{\theta^{(i)}}\, p(\theta^{(i)}) \tag{24}$$

However, it is possible to construct simulated annealing algorithms [23] that allow us to sample approximately from the global maximum of the target PDF.

4.2. MCMC—the Gibbs sampler

MCMC is a strategy for generating samples $\theta^{(i)}$ while exploring the state space of θ using a Markov chain mechanism [21,24]. 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 he cannot 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 a m-dimensional random vector θ, the full conditional PDF is defined as $p(\theta_i|\theta_{-i})$, where θ_{-i} stands for $\{\theta_1,\ldots,\theta_{i-1},\theta_{i+1},\ldots,\theta_m\}$. 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 paper, Gibbs sampler is used to speedup the chain convergence. With Nmcmc 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 $\theta_1^{(i+1)} \sim p(\theta_1|\theta_2^{(i)}, \theta_3^{(i)}, \ldots, \theta_m^{(i)})$
   - sample $\theta_2^{(i+1)} \sim p(\theta_2|\theta_1^{(i+1)}, \theta_3^{(i)}, \ldots, \theta_m^{(i)})$
   - $\vdots$
   - sample $\theta_m^{(i+1)} \sim p(\theta_m|\theta_1^{(i+1)}, \theta_2^{(i+1)}, \ldots, \theta_{m-1}^{(i+1)})$

For the two Bayesian formulations in Eqs. (20) and (21), different Gibbs samplers have been designed for the linear IHCP. Using Eq. (7), the relation between the unknown vector θ and observation Y can be written as follows:

$$Y = H\theta + T_{\mathrm{I}} + \omega_{\mathrm{m}} \tag{25}$$

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

$$H((i-1)N + j, k) = T_{\mathrm{H}}(\hat{\mathbf{x}}_i, \hat{t}_j; w_k) \tag{26}$$

In above equation, T_{H} denotes the direct solution with zero initial temperature condition, and zero temperature and flux boundary conditions on Γ_{g} and Γ_{h}, respectively, and w_k as the applied heat flux on Γ_0. T_{I} denotes the direct solution with zero heat flux condition on Γ_0 and the prescribed known initial conditions in Ω and boundary conditions on $\Gamma_{\mathrm{g}} \cup \Gamma_{\mathrm{h}}$. Hence, $F(\theta)$ in Eqs. (20) and (21) is replaced by $H\theta + T_{\mathrm{I}}$. When σ 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:

$$p(\theta_i|\theta_{-i}) \sim N(\mu_i, \sigma_i^2), \quad \mu_i = \frac{b_i}{2a_i}, \quad \sigma_i = \sqrt{\frac{1}{a_i}} \tag{27}$$

$$a_i = \sum_{s=1}^{n} \frac{H_{si}^2}{\sigma^2} + \lambda W_{ii}, \quad b_i = 2\sum_{s=1}^{n} \frac{\mu_s H_{si}}{\sigma^2} - \lambda \mu_p \tag{28}$$

$$\begin{aligned} \mu_s &= Y_s - T_{\mathrm{I}} - \sum_{t \neq i} H_{st}\theta_t, \\ \mu_p &= \sum_{j \neq i} W_{ji}\theta_j + \sum_{k \neq i} W_{ik}\theta_k \end{aligned} \tag{29}$$

The standard form of the full conditional PDF enables us to use the Gibbs sampler in this problem. In this formulation note that the MAP estimate is the same as the posterior mean estimate. In all examples reported later in this paper, the posterior mean is used to refer to the inverse solution. If one is only interested in the point estimate, gradient methods can be used to solve the optimization problem of Eq. (19) for a linear IHCP. This approach is referred to as 'direct optimization' in the remaining of the paper. However, implementation of a Gibbs sampler is indispensable for exploiting the posterior state space of θ to achieve marginal PDFs, to quantify uncertainty (probability bounds) in posterior mean estimate, or to compute the expectation of an arbitrary function of θ.

When σ is not known, a modified Gibbs sampler is designed for the PPDF (Eq. (21)). In the current work, a negative exponential distribution is assigned for the prior PDF of σ^2. This is to constrain the non-negativity of σ^2 while exploring the fact that the probability for σ^2 to take small values is high. Then Eq. (21) can be written as:

$$\begin{aligned} p(\theta, \sigma^2|Y) \propto (\sigma^2)^{-n/2} \exp\left\{ -\frac{1}{2\sigma^2}[H\theta + T_{\mathrm{I}} - Y]^{\mathrm{T}}[H\theta + T_{\mathrm{I}} - Y] \right\} \\ (\sigma^2)^{-m/2} \exp\left\{ -\frac{\alpha}{\sigma^2}\theta^{\mathrm{T}} W \theta \right\} \exp\left\{ -\frac{\sigma^2}{\beta} \right\} \end{aligned} \tag{30}$$

where β defines the exponential distribution.

With any specified σ, the full conditional PDFs of θ_i's remain the same as in Eqs. (27)–(29). Therefore, the Gibbs sampler is used to update θ. After each update of all θ_i's, the full conditional PDF of σ^2 is maximized to obtain the update sample of σ^2:

$$p(\sigma^2|\theta, Y) \propto (\sigma^2)^{-\frac{n+m}{2}} \exp\left(-\frac{C}{2\sigma^2} - \frac{\sigma^2}{\beta} \right) \tag{31}$$

where $C = [H\theta + T_{\mathrm{I}} - Y]^{\mathrm{T}}[H\theta + T_{\mathrm{I}} - Y] + 2\alpha\theta^{\mathrm{T}} W \theta$.

Statisticians have developed a large number of techniques for convergence assessment [25]. The convergence of MCMC is determined here by monitoring the histogram and marginal density of accepted samples.

5. Examples

In this section, two IHCP examples are considered using the introduced Bayesian inference techniques. The first example is a detailed study of a classical one-dimensional IHCP. The second two-dimensional example provides estimation of the statistics of a space and time varying heat flux given several temperature measurements in the domain. Dimensionless quantities are used in both examples.

5.1. Example I: boundary heat flux identification in 1D heat conduction

A 1D transient linear heat conduction problem is considered (Fig. 3). The governing equations are as follows:

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2}, \quad 0 < t < 1, \ 0 < x < 1 \tag{32}$$

$$T(x, 0) = 0, \quad 0 \leqslant x \leqslant 1 \tag{33}$$

$$\left.\frac{\partial T}{\partial x}\right|_{x=1} = 0, \quad 0 < t < 1 \tag{34}$$

$$\left.\frac{\partial T}{\partial x}\right|_{x=0} = q(t) \text{ (to be computed)}, \quad 0 < t < 1 \tag{35}$$

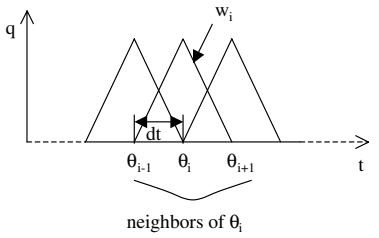

Fig. 2. Basis functions for the discretization of the 1D heat flux.

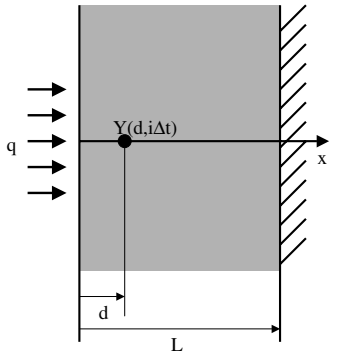

Fig. 3. One-dimensional inverse heat conduction example. One is interested to compute the temporal history of the boundary heat flux q given temperature measurements at the sensor placed at $x = d$.

The direct problem considered represents the solution of the system of equations above with the heat flux $q(t)$ defined in Fig. 4. A FEM solver was implemented for the solution of this problem with 100 linear elements in space and 1000 time steps. It is considered here that this discretization can accurately reflect the direct heat conduction process. Simulated measurement data at $x = d$ (considered as the thermocouple location) are generated by using the obtained direct solution $T(d,t)$. The measurement vector Y is obtained by adding random noise to the vector $[T(d,\Delta t) \quad T(d,2\Delta t) \quad \ldots \quad T(d,1)]^{\mathrm{T}}$, where Δt is the time interval between two consecutive measurements. Therefore, the size of Y is $n = N = \frac{1}{\Delta t}$. The inverse problem of interest is thus the reconstruction of the heat flux of Fig. 4 given Eqs. (32)–(34) and the sensor data Y.

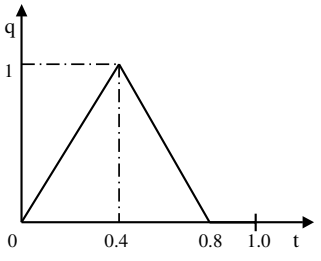

Fig. 4. True boundary heat flux in example I.

The sensitivity matrix H is given as follows:

$$H(i,j) = T(d, i\Delta t; w_j), \quad i = 1:N, \;\; j = 1:m \tag{36}$$

The 1D linear finite element shape functions are used as the basis functions in Eq. (7). By doing so, θ_i is nothing but the heat flux value at $t = i\,\mathrm{d}t$, where $\mathrm{d}t = \frac{1.0}{m-1}$. Fig. 2 shows the basis functions and neighborhood (as used in the MRF definition) for the 1D heat flux. For this problem, T_1 is a zero vector. Substitution of H into the first Bayesian formulation of Eq. (20), leads to the following posterior PDF of θ:

$$p(\theta|Y) \propto \exp\left\{ -\frac{1}{2\sigma^2}[H\theta - Y]^{\mathrm{T}}[H\theta - Y] \right\} \cdot \exp\left(-\frac{1}{2}\lambda\theta^{\mathrm{T}} W\theta \right) \tag{37}$$

Using the Gibbs sampler discussed earlier, several cases were studied for this problem. In the first three cases, it is taken that $d = 0.3$, $\Delta t = 0.02$ ($n = 50$) and $\mathrm{d}t = 0.04$ ($m = 26$). The measurement errors are white noises, which are generated independently from zero mean normal distributions. The standard deviation σ for these three cases is 0.001, 0.005 and 0.01, respectively. The maximum, median and minimum of the recorded (here simulated) temperatures are approximately 0.4, 0.2 and 2.0e−4, respectively. Therefore, the values of the ratio of noise to signal for these cases are approximately as follows:

Case	σ	$\frac{\sigma}{T_{\max}}$ (%)	$\frac{\sigma}{T_{\min}}$ (%)	$\frac{\sigma}{T_{\mathrm{med}}}$ (%)
I	0.001	0.25	500	0.5
II	0.005	1.25	2500	2.5
III	0.010	2.5	5000	5

Figs. 5 and 6 show the plots of the MLE and posterior mean estimates of the boundary heat flux for these three cases. The relative estimation error is defined as $r = \frac{\|\hat{\theta} - \theta_{\mathrm{true}}\|_{L_2}}{\|\theta_{\mathrm{true}}\|_{L_2}}$. Since the standard deviation of measurement noise is stationary, the MLE estimate is the same as the LS estimate. For these three cases, the MLE estimates are obtained by solving the optimization problem in Eq. (16) without regularization. This is for the purpose of showing the smoothing effect of the MRF. It can be seen from comparison of Figs. 5 and 6 that the MRF prior model greatly smoothes the inverse solutions. The posterior mean estimates shown are computed by direct optimization. The regularization parameter α for the three cases was selected as 5.0e−4, 5.0e−3 and 2.5e−2, respectively, using the heuristic Tikhonov method.

It is noticed that the posterior mean estimate is not sensitive to the knowledge of σ once the regularization

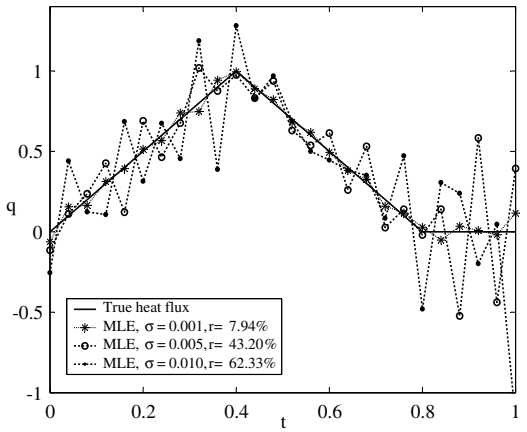

Fig. 5. MLE estimates of heat flux when $d = 0.3$, $\Delta t = 0.02$ ($n = 50$), and $dt = 0.04$ ($m = 26$).

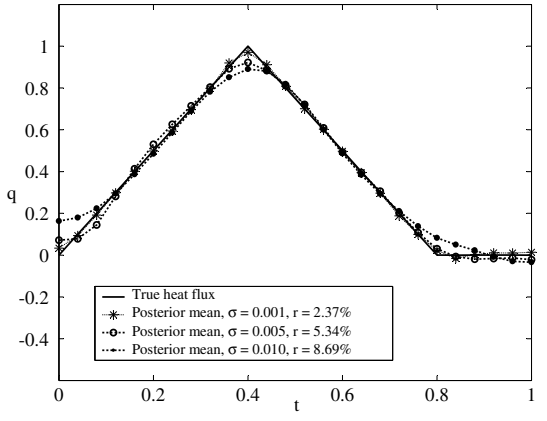

Fig. 6. Posterior mean estimates of the heat flux when $d = 0.3$, $\Delta t = 0.02$ ($n = 50$) and $dt = 0.04$ ($m = 26$).

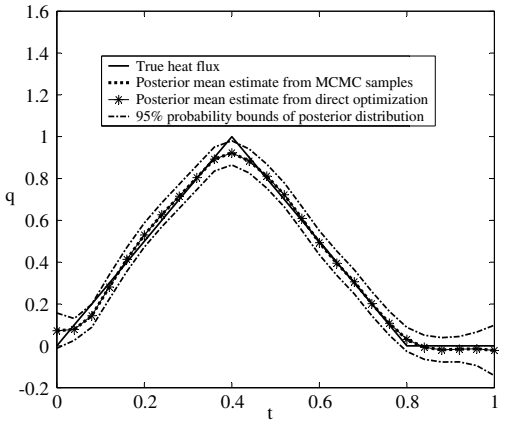

Fig. 7. Posterior mean estimate and associated two-side 95% probability bounds of case II from MCMC samples using the true σ in the chain.

constant is fixed. This implies that statistical information of white noise is not necessary for the point estimate. However, to quantify the uncertainty of point estimate, the knowledge of σ is crucial because it affects the PPDF. To investigate the effect of σ on PPDF, two MCMC chains are run for case II using the Gibbs sampler. In Fig. 7, the posterior mean estimate is plotted with associated two-side 95% probability bounds for the first chain, where the true σ is used. The probability bound is an indication of the range of the highest density region of the posterior state space. The posterior mean estimates obtained from direct optimization and MCMC samples are not distinguishable, which is a verification of the accuracy of the MCMC.

Fig. 8 provides plots of three marginal distributions obtained from the MCMC samples of the first chain. Figs. 9 and 10 plot the same quantities as Figs. 7 and 8 from the second chain, where a σ value is used that is twice the true σ. In both runs of the Gibbs sampler,

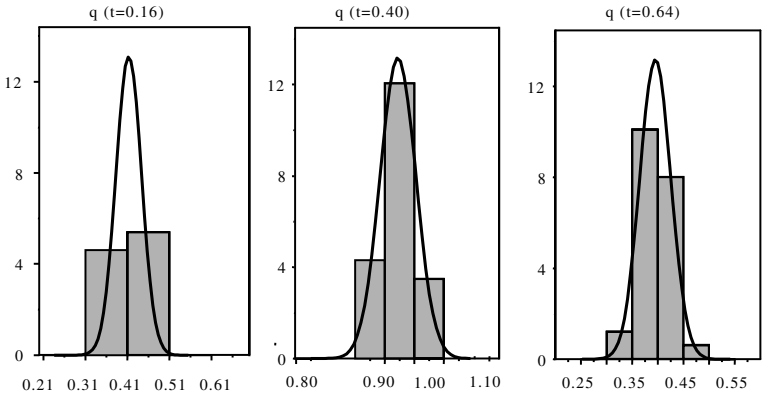

Fig. 8. Marginal PDF examples of case II from MCMC samples using the true σ in the chain.

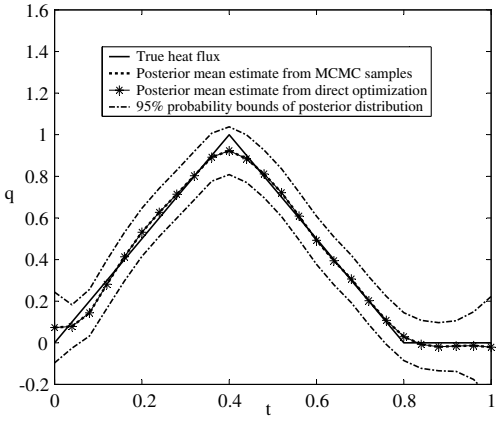

Fig. 9. Posterior mean estimate and associated two-side 95% probability bounds of case II from MCMC samples using twice of the true σ in the chain.

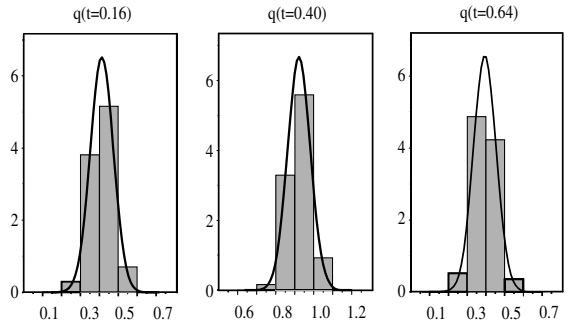

Fig. 10. Marginal PDF examples of case II from MCMC samples using twice of the true σ in the chain.

10,0000 samples of θ are recorded and the last 50,000 are used in the above plots. It can be seen that the posterior mean estimates are exactly the same using these two chains. The width of 95% probability bounds in the second chain, however, is almost twice the one in the first chain.

In practice, an estimate of σ is achieved through either previous knowledge of the measurement equipment or by repeating the experiments and collecting a sequence of data. The more the data available at each site, the higher the accuracy of the PPDF. Herein, the augmented Bayesian formula is used to update σ as well as the distribution of the heat flux for case II. The selection of β is based on a rough estimation of the possible range of σ^2. Here β is taken as 1.0e−4 to demonstrate the algorithm. For comparison purposes, the modified Gibbs sampler is run with $d = 0.3$, $\Delta t = 0.02$ ($n = 50$), $\mathrm{d}t = 0.04$ ($m = 26$) and $\alpha = 0.01$. The same data set is used as in case II and every condition is the same as in Figs. 7 and 9 except that there is no prior information on σ. The same length of Markov chain is run as in the previous two chains. The posterior mean of estimated σ is 0.008. This is indeed the best estimate a single run experimental data can provide for σ without knowing any characteristic of the measurement equipment. The estimate of σ will of course vary according to specific noise feature in the single experimental data set or with the value of β. It is seen from Fig. 11 that the posterior mean estimate of the heat flux is exactly the same as in Figs. 7 and 9. The width of the posterior distribution is between these two previous cases.

To check the effect of the temperature measurement location on the inverse solution, a fourth case is considered in which $d = 0.1$ with all other conditions remaining the same as case II. It is seen from Fig. 12 that the accuracy of the posterior mean estimate is slightly improved compared to that of case II. Finally, in Fig. 13, the temperature prediction based on posterior mean

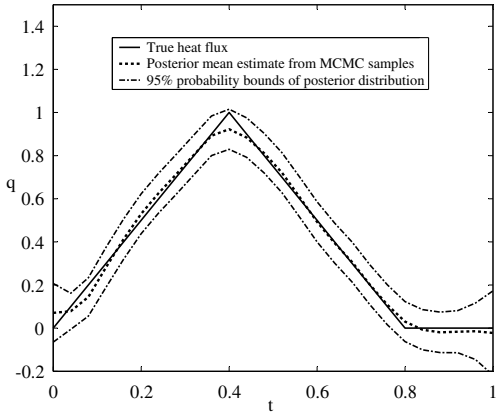

Fig. 11. Posterior mean estimate and associated two-side 95% probability bounds for data case II from MCMC samples using the augmented Bayesian model.

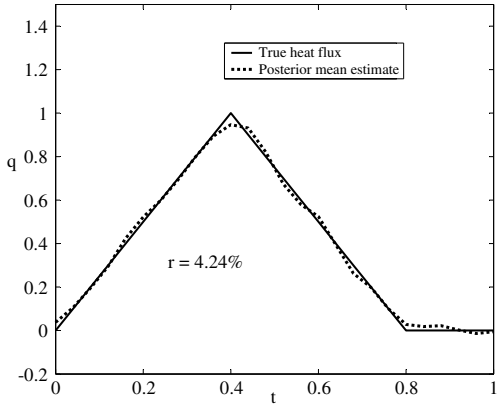

Fig. 12. Posterior mean estimate of heat flux when $d = 0.1$, $\Delta t = 0.02$ ($n = 50$) and $\mathrm{d}t = 0.04$ ($m = 26$).

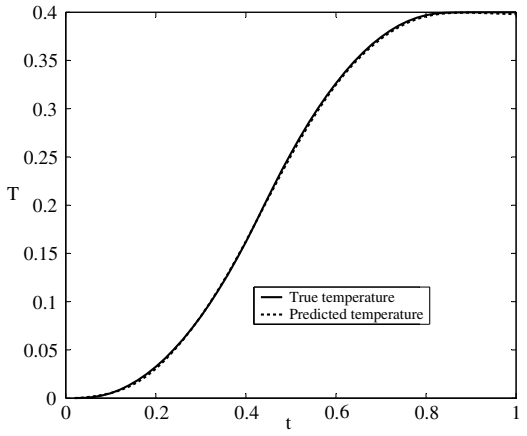

Fig. 13. Temperature prediction at $d = 0.5$ for case II.

estimate in case II is compared with the true temperature at $x = 0.5$. The predicted temperature is computed as sample mean of $H\theta$ from MCMC samples. The prediction agrees very well with the true temperature.

The sampling frequency in the above cases was selected to emphasize that the Bayesian analysis works well with relatively few data. Even though other methods of analysis of the IHCP (e.g. the sequential version of the function specification method) are quite sensitive to sampling rates, we have repeated the above calculations for smaller Δt to show that the Bayesian analysis provides better estimates of the heat flux for each of the cases reported above. This is expected from a statistical point of view since more data implies better understanding of the noise distribution and better estimation. The effect of discretization of the parameter space (see Fig. 2) was also investigated. For example, considering $\Delta t = 0.004$ and all other conditions as in case II, we have shown that the computed heat flux estimate remains quite accurate. Finer discretization of the parameter space for a fixed set of measured data leads as expected to a progressive loss of accuracy.

5.2. Example II: boundary heat flux identification in 2D heat conduction

In this section, a solution is presented to a 2D transient heat flux identification problem. The direct problem considered is defined in dimensionless form as follows:

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}, \quad 0 < t < 1, \ 0 < x, y < 1 \tag{38}$$

$$T(x, y, 0) = 2\sin(\pi x)\sin(\pi y), \quad 0 \leqslant x, y \leqslant 1 \tag{39}$$

$$T|_{x=1} = T|_{y=1} = 0, \quad 0 < t < 1 \tag{40}$$

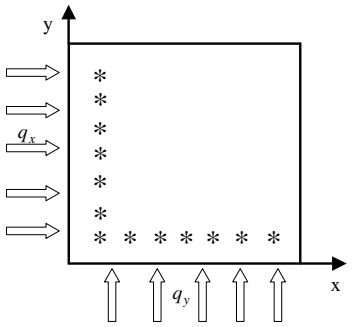

Fig. 14. Thermocouple locations in example II.

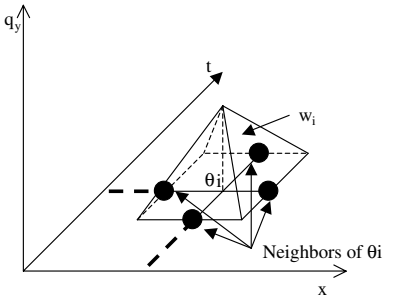

Fig. 15. Basis functions for 2D heat flux and graphical representation of the neighborhood used in the MRF model.

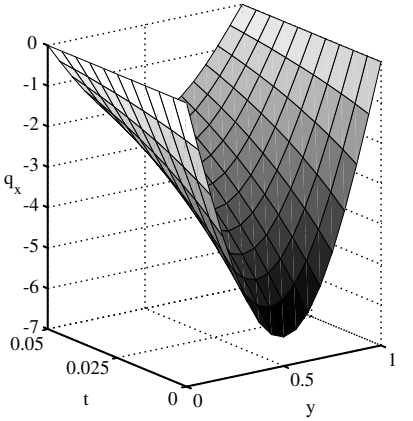

Fig. 16. True heat flux at $x = 0$.

$$\left.\frac{\partial T}{\partial x}\right|_{x=0} = q_x, \quad \left.\frac{\partial T}{\partial y}\right|_{y=0} = q_y, \quad 0 < t < 1 \tag{41}$$

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

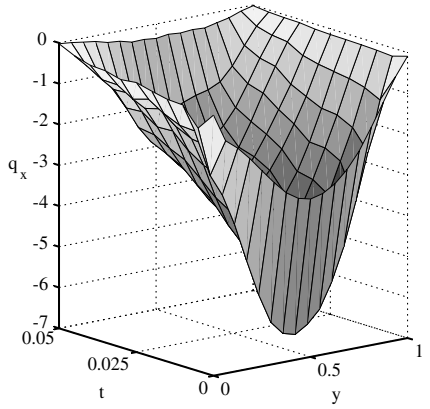

Fig. 17. MLE estimate of heat flux at $x = 0$ when $\sigma = 0.005$.

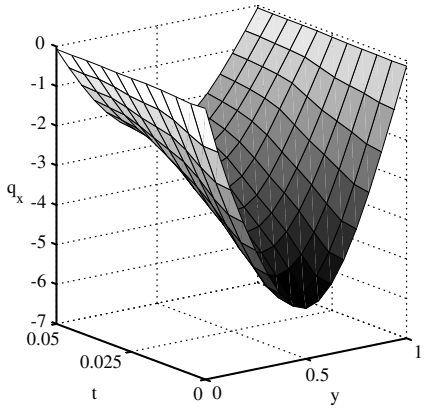

Fig. 20. Posterior mean estimate of heat flux at $x = 0$ when $\sigma = 0.01$.

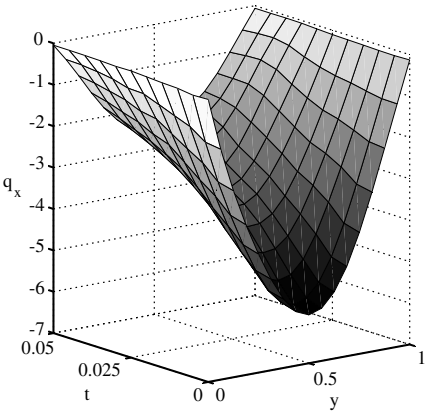

Fig. 18. Posterior mean estimate of heat flux at $x = 0$ when $\sigma = 0.005$.

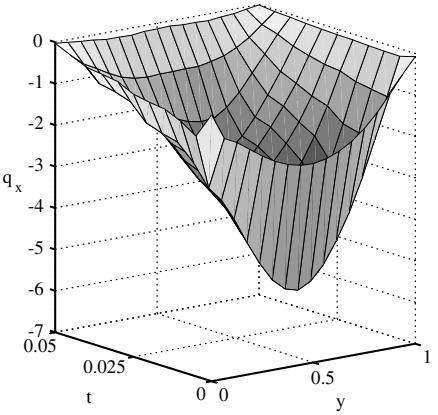

Fig. 21. MLE estimate of heat flux at $x = 0$ when $\sigma = 0.02$.

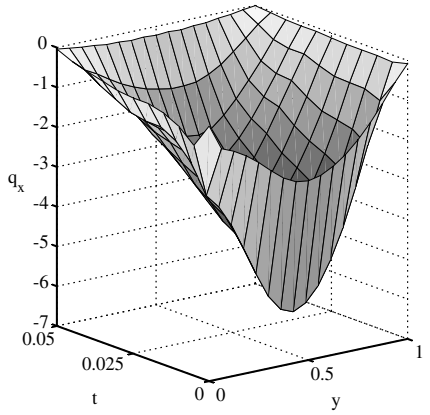

Fig. 19. MLE estimate of heat flux at $x = 0$ when $\sigma = 0.01$.

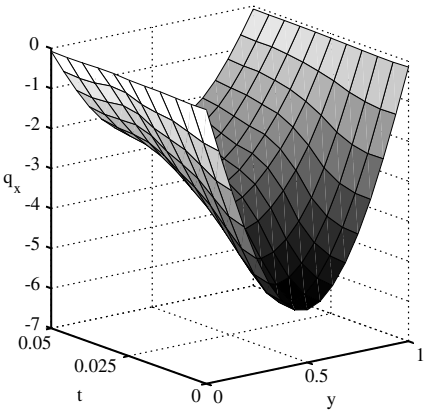

Fig. 22. Posterior mean estimate of heat flux at $x = 0$ when $\sigma = 0.02$.

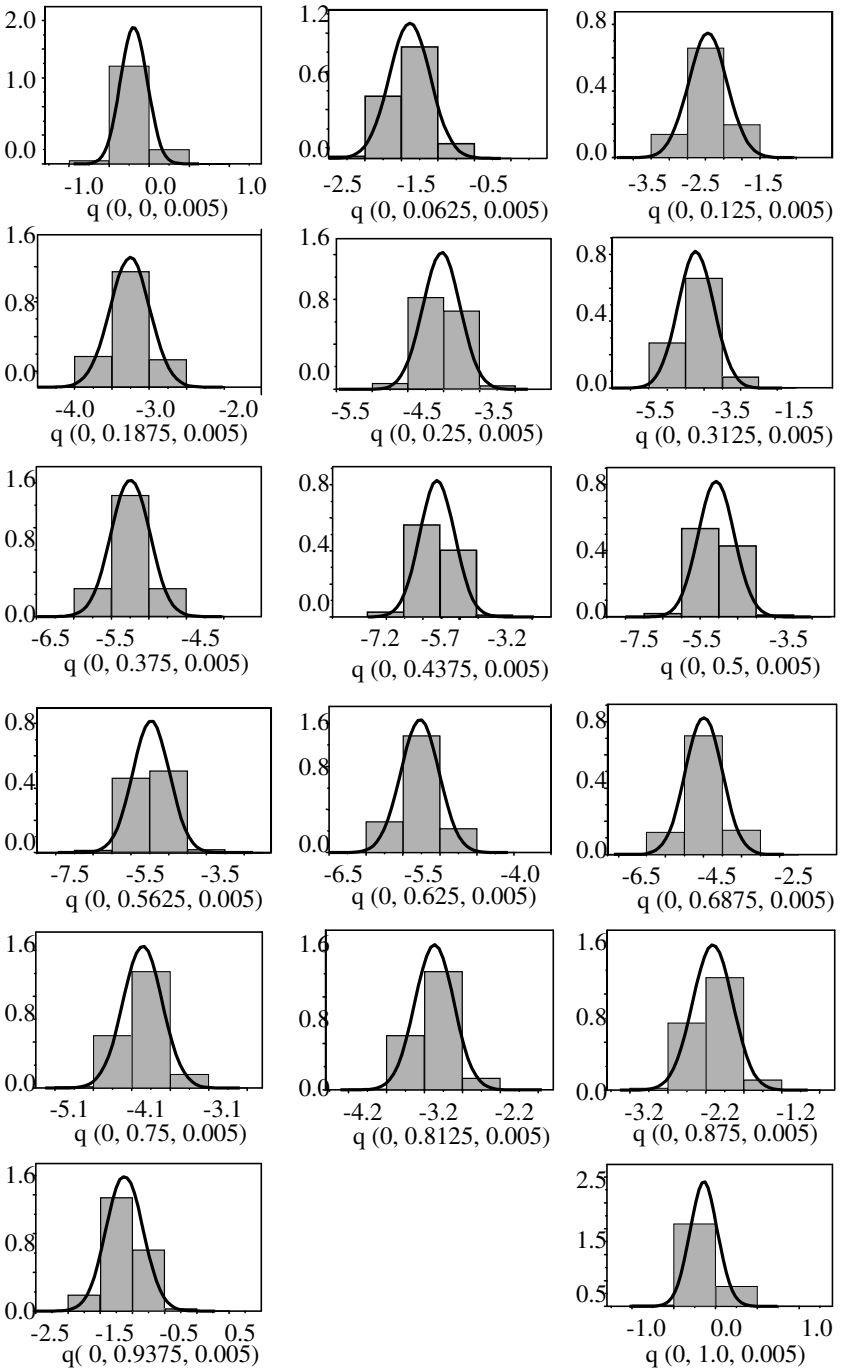

Fig. 23. Marginal PDF of q_x at $t = 0.005$ when $\sigma = 0.02$.

$$q_x = -2\pi \sin(\pi y) \exp(-2\pi^2 t) \tag{42}$$

$$q_y = -2\pi \sin(\pi x) \exp(-2\pi^2 t) \tag{43}$$

and is given as follows:

$$T(x, y, t) = 2 \sin(\pi x) \sin(\pi y) \exp(-2\pi^2 t) \tag{44}$$

The inverse problem is to reconstruct q_x and q_y. The locations of simulated thermocouples are shown in Fig. 14. Thirteen evenly distributed thermocouples are

considered with space interval $d = 0.125$ and the distance to the boundary is also 0.125. The sampling time interval is taken as $\Delta t = 0.002$. The heat flux history was reconstructed for the time range $t \in [0, 0.05]$, $N = 25$, hence, there are 325 observations. 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 11 basis functions are used in the temporal direction ($\mathrm{d}t = 0.005$), thus $m = 374$. Fig. 15 shows the basis functions and the neighborhood to each nodal component of the 2D heat flux. The direct solver of the previous example is extended to compute the sensitivity matrix H. The Gibbs sampler is still applicable. It is necessary to point out that the same MRF regularization model is used in both time and space. Also note that the temperature T_I is not zero in this example.

Three cases were studied with standard deviation of measurements taken as 0.005, 0.01 and 0.02 (4% of $\Delta T_{\max}$), respectively. Fig. 16 plots the true heat flux q_x whereas Fig. 17 shows the MLE estimates of q_x for case I. Note that zeroth order Tikhonov regularization was used in the MLE estimation. There is no unique solution to the optimization problem of Eq. (16) without regularization in this problem because the number of thermocouples is limited. The regularization parameter used for case I is 5.0e−5.

Fig. 18 plots the posterior mean estimate of q_x using the same regularization parameter for case I. It is seen by comparing with Fig. 17 that the posterior mean estimate is much more smooth than the MLE estimate. In fact, the relative estimation error is 28.76% for the MLE estimate and is reduced to 4.62% by using MRF regularization. One point to be addressed is that by using the primal form of the MRF as introduced earlier in this paper, the posterior mean estimate has obvious 'edge effects'. That is, for example, the estimate of q_x has large deviation from the true solution at $y = 0$ and $y = 1$. This is due to the fact that MRF models the interaction between all adjacent sites but the sites on the boundary have only neighbors on one side of the boundary. To bypass this problem, the covariance matrix W is modified such that when θ_i is on the edge, n_i is assigned a large number to diminish the dependence from inner sites. Similar results are observed for case II and III, for which the MLE estimates are plotted in Fig. 19 (case II) and Fig. 21 (case III). The posterior estimates are plotted in Fig. 20 (case II) and Fig. 22 (case III). The regularization parameters for these two cases are 2.5e−4 and 5.0e−4, respectively. The relative estimation errors of MLE estimates in cases II and III are 33.19% and 35.92% respectively, which are reduced to 5.45% and 5.73%, respectively, for the posterior mean estimates.

For demonstration of the computed posterior distribution of heat fluxes, the marginal PDFs when $\sigma = 0.02$ are plotted for q_x at $t = 0.005$ in Fig. 23.

6. Conclusions

A Bayesian inference approach was introduced for the solution of the inverse heat conduction problem. A MCMC-based numerical sampling strategy was adopted to exploit the posterior state space. The proposed techniques were shown through a number of examples to provide satisfactory solutions to the IHCP. They lead to not only a point estimate of the unknown heat flux but also an estimate of its statistical information as well as quantification of the system uncertainties. In addition, the inverse problem is regularized statistically through the modeling of prior distribution. A Bayesian approach provides a complete investigation of the IHCP, hence, it is expected to provide more robust estimates, especially when fewer sensor data are available.

Acknowledgements

This work was partially supported from the Advanced Mechanical Technologies Program at General Electric Global Research Center (GE-GRC), by NASA, Office of Biological and Physical Sciences Research (98-HEDS-05), the Air Force Office of Scientific Research (grant F49620-00-1-0373) and by the National Science Foundation (grant DMI-0113295). This research was conducted using the resources of the Cornell Theory Center.

References

[1] O.M. Alifanov, Inverse Heat Transfer Problems, Springer-Verlag, Berlin, 1994.
[2] J.V. Beck, B. Blackwell, C.R. St-Clair Jr., Inverse Heat Conduction: ill-Posed Problems, Wiley-Interscience, New York, 1985.
[3] A.N. Tikhonov, Solution of ill-Posed Problems, Halsted Press, Washington, 1977.
[4] N. Zabaras, J.C. Liu, An analysis of two-dimensional linear inverse heat-transfer problems using an integral method, Num. Heat Transfer 13 (4) (1988) 527–533.
[5] R. Sampath, N. Zabaras, A functional optimization approach to an inverse magneto-convection problem, Comput. Meth. Appl. Mech. Eng. 190 (15–17) (2001) 2063–2097.
[6] V.A. Badri Narayanan, N. Zabaras, Stochastic inverse heat conduction using a spectral approach, Int. J. Numer. Meth. Eng. 60 (7) (2004).
[7] J.S. Liu, Monte Carlo Strategies in Scientific Computing, Springer-Verlag, Berlin, 2001.
[8] T.D. Fadale, A.V. Nenarokomov, A.F. Emery, Uncertainties in parameter estimation: the inverse problem, Int. J. Heat Mass Transfer 38 (3) (1995) 511–518.
[9] A.F. Emery, A.V. Nenarokomov, T.D. Fadale, Uncertainties in parameter estimation: the optimal experiment design, Int. J. Heat Mass Transfer 43 (2000) 3331–3339.

[10] C.P. Robert, The Bayesian Choice, From Decision-Theoretic Foundations to Computational Implementation, second ed., Springer, 2001.
[11] J. Besag, P. Green, D. Higdon, K. Mengersen, Bayesian computation and stochastic systems, Stat. Sci. 10 (1) (1995) 3–41.
[12] A.F. Emery, Stochastic regularization for thermal problems with uncertain parameters, Inverse Probl. Eng. 9 (2001) 109–125.
[13] C. Vogel, An applied mathematician's perpective 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, September 2002.
[14] J. Besag, C. Kooperberg, On the conditional and intrinsic autoregressions, Biometrika 82 (1995) 733–746.
[15] H.K. Lee, D.M. Higdon, Z. Bi, M.A.R. Ferreira, M. West, Markov random field models for high-dimensional parameters in simulations of fluid flow in porous media, Technometrics 44 (3) (2002) 230–241.
[16] J. Møller (Ed.), Spatial Statistics and Computational Methods, Springer-Verlag, New York, 2003.
[17] J. Mateu, F. Montes (Eds.), Spatial Statistics Through Applications, International Series on Advances in Ecological Sciences, WIT Press, Boston, 2002.
[18] R. Chellappa, A. Jain (Eds.), Markov Random Fields Theory and Application, Academic Press Inc., 1991.
[19] J. Besag, R.A. Kempton, Statistical analysis of field experiments, Biometrics 78 (1986) 301–304.
[20] D. Higdon, H.K. Lee, Z. Bi, A Bayesian approach to characterizing uncertainty in inverse problems using coarse and fine-scale information, IEEE Trans. Signal Process. 50 (2) (2002) 389–399.
[21] C. Andrieu, N. Freitas, A. Doucet, M.I. Gordan, An introduction to MCMC for machine learning, Mach. Learn. 50 (2003) 5–43.
[22] P. Brémaud, Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues, Springer-Verlag, New York, 1999.
[23] P.J. Van Laarhoven, E.H.L. Arts, Simulated Annealing: Theory and Applications, Reidel Publishers, Amsterdam, 1987.
[24] L. Tierney, Markov chains for exploring posterior distributions, The Ann. Stat. 22 (4) (1994) 1701–1762.
[25] S.P. Brooks, G.O. Roberts, Convergence assessment techniques for Markov chain Monte Carlo, Stat. Comput. 8 (1998) 319–335.