COMPUTATIONAL MECHANICS
WCCM VI in conjunction with APCOM'04, *Sept. 5-10, 2004, Beijing, China*

A Computational Statistics Approach to Stochastic Inverse Problems and Uncertainty Quantification in Heat Transfer

J. Wang[1], N. Zabaras[1*]

[1] *Materials Process Design and Control Laboratory, Sibley School of Mechanical and Aerospace Engineering, Cornell University, 188 Frank H.T. Rhodes Hall, Ithaca, NY, 14853-3801, USA.*
e-mail: jw332@cornell.edu, zabaras@cornell.edu

Abstract As most engineering systems and processes operate in an uncertain environment, it becomes increasingly important to address their analysis and inverse design in a stochastic manner using statistical data-driven methods. Recent advances in computational Bayesian and spatial statistics enable complete and efficient solution procedures to such problems. Herein, a novel framework based on Bayesian inference is presented for the solution of stochastic inverse problems in heat transfer. The posterior probability density function (PPDF) of unknowns (modeled as random variables or stochastic processes), such as material thermal properties and boundary heat flux, is computed given finite set of thermocouple temperature measurements. Markov Chain Monte Carlo (MCMC) algorithms are exploited to obtain estimates of statistics of random unknowns. A parameter estimation problem is first solved using simple, hierarchical and augmented Bayesian models. Boundary heat flux reconstruction in heat conduction is then studied. Simulation results demonstrate the great potential of applying a Bayesian approach to stochastic estimation and design problems. Although discussed in the context of thermal systems, the methodology presented is general and applicable to design and estimation problems in diverse areas of engineering.

Key words: inverse problems, heat transfer, uncertainty, Bayesian inference, MCMC

INTRODUCTION

Two types of inverse problems are often raised in heat transfer processes, estimation of material thermal properties such as thermal conductivity and reconstruction of the history of the boundary heat flux. A number of deterministic optimization theories and algorithms have been developed toward the solution of these inverse problems [1] - [3]. The general approach is to pose the inverse problem as a least-squares optimization problem. To address the ill-posed nature of these inverse problems, it is common to augment the objective function with Tikhonov regularization terms [2, 4]. Gradient methods are then applied to minimize the objective function and compute a point estimate of the inverse solution.

Due to the existence of uncertainties in any real-world system, even small perturbations in the given data or variabilities in the system parameters propagate to and get amplified in the inverse solution. Therefore, a complete inverse solution should provide uncertainty quantification and not only a point estimate of unknowns. With the rapid explosion of computational power and critical demands on engineering system robustness and reliability, optimization under uncertainty is receiving a growing attention [5, 6]. Recently, a sequence of methods have been proposed to solve stochastic inverse heat transfer problems including sensitivity analysis by Norris [7] and Blackwell and Dowding [8], extended Maximum Likelihood Estimator (MLE) approach by Emery et al. [9, 10], spectral stochastic method by Narayanan and Zabaras [11], and Bayesian inference method by Ferrero and Gallagher [12], Leoni and Amon [13] and Wang and Zabaras [14, 15].

Compared with other methods, the Bayesian statistical inference method [16, 17] has some significant advantages. Firstly, it provides conditional probability density function (PDF) of the unknown quantities on all available and missing data. The incorporation of likelihood and prior distribution enables a complete probabilistic description of the unknowns under all uncertainties [18]. Secondly, it explores the distribution information in polluted data, which is rather critical since the inverse solutions are extremely sensitive to noise. In addition, the Bayesian method provides inherent regularization to the ill-posed inverse problem [14, 19]. Furthermore, a Bayesian approach to data-driven inverse problems has no critical requirement on the number of needed measurements. Another advantage is that under the Bayesian framework, the forward problems are solved deterministically with the uncertainties taken care by statistical inference. Finally, the associated rich family of sampling strategies overcome the optimization difficulties encountered when dealing with nonlinear problems.

In this paper, the Bayesian inference method is used to address inverse problems in heat transfer. The posterior probability density functions (PPDF) of unknowns conditional on temperature measurements are computed by multiplying the likelihoods and assumed prior distribution functions. The ill-posed inverse problems are regularized through Markov Random Field (MRF) prior distribution modeling [19]. A sequence of numerical samplers based on Markov Chain Monte Carlo (MCMC) simulation algorithms [20, 21] are designed to explore the posterior state spaces. Bayesian point estimates and associated probability bounds are then computed from the sample marginal density functions. The purpose of this study is to outline a novel approach for solving stochastic inverse problems in continuum engineering systems.

The rest of this paper is organized in the following way. An introduction to Bayesian statistics and MRF is first given to make the information in this paper self-contained. Then the parameter estimation problem in heat conduction is studied using Bayesian inference techniques. In addition to a simple Bayesian formulation, both hierarchical and augmented Bayesian models are discussed to relax the prior assumptions on uncertainties. This is followed by a discussion of the IHCP where boundary heat flux is reconstructed by updating the MRF-based prior. The following section is devoted to an introduction of MCMC algorithms. Numerical experiments are then conducted to demonstrate the methodologies presented. Finally, a discussion on the numerical studies is provided.

FUNDAMENTALS OF BAYESIAN STATISTICS

Bayesian statistics study the probability of a random variable from both current achieved information and previous knowledge [16]. The basis of Bayesian inference is the Bayes' formula:

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

where θ and Y denote the unknown random parameter and related data, respectively. 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 defined such as the 'Maximum A Posteriori' (MAP) estimator:

$$\hat{\theta}_{\text{MAP}} = augmax_{\theta} \;\; p(\theta|Y) \tag{2}$$

and the posterior mean estimator:

$$\hat{\theta}_{\text{postmean}} = E \;\; \theta|Y \tag{3}$$

In general, the probability $p(Y)$ is not explicitly known and is rather difficult to compute. However, 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,

$$p(\theta|Y) \propto p(Y|\theta)p(\theta) \tag{4}$$

The likelihood function can be obtained from the following relationship,

$$Y = F(\theta) + \omega \tag{5}$$

where F is the mathematical model for the forward problem, and ω is the measurement noise, which is usually assumed as Gauss random noise with mean 0 and variance υ_T. Subsequently, the likelihood can be written as,

$$p(Y|\theta) = \frac{1}{(2\pi)^{n/2}\upsilon_T^{n/2}} \exp\{-\frac{(Y - F(\theta))^T(Y - F(\theta))}{2\upsilon_T}\}, \tag{6}$$

where n is the number of measurements. The prior distribution reflects the knowledge, if there is any, of the unknowns, before Y is gathered. For instance, it can be the estimate of $p(\theta)$ resulting from previous experiments or simulations. From an inverse point of view, the prior distribution model provides regularization to the ill-posed inverse problem of estimating θ using the data Y [14]. In the current study, conjugate priors [16] are used for scalar random variables, and Markov Random Field (MRF) [19] is adopted for the prior modeling of random processes such as heat flux. The MRF used herein has the form of

$$p(\theta) \propto \lambda^{m/2} \exp(-\frac{1}{2}\lambda\theta^T W\theta), \tag{7}$$

where m is the dimension of θ. As an example, in the case of estimating the heat flux on the boundary from temperature measurements within the domain, each component θ_i of θ represents the value of the heat flux at a particular spatial and temporal site. In the one-parameter model of Eq. (7), the entries of the $m \times m$ matrix W are determined as, $W_{ij} = n_i$ if i=j, $W_{ij} = -1$ if sites i and j are adjacent, and as 0 otherwise. n_i is the number of neighbors adjacent to site i. λ is a constant that controls the scaling of distribution of θ. This MRF model is equivalent to Tikhonov regularization provided the measurement noise is Gaussian and the objective is to maximize the posterior probability (MAP) [14].

THERMAL PARAMETER ESTIMATION

In many practical situations, the thermal properties of conducting solids are not directly measurable. Therefore, experiments are designed to measure close-related quantities such as temperature. Inverse problem is then solved to obtain the optimal estimate of the unknown property. Bayesian inference is applicable to this type of inverse problem since temperature is recognized as a sufficient statistic of the thermal properties. Herein, the case to estimating the thermal conductivity is analyzed.
Considering the model problem where temperature variation in a conducting solid is governed by,

$$\rho C_p \frac{\partial T}{\partial t} = \bigtriangledown \cdot (k \bigtriangledown T), \quad \text{in domain} \quad \Omega, \quad t \in [0, t_{\max}] \tag{8}$$

with prescribed initial and boundary conditions. Material density ρ and heat capacity C_p are known, and the thermal conductivity k is unknown. However, temperature can be measured at some location $\vec{d}$ within the domain. Let Y denotes the $n \times 1$ vector of temperature measurements. The aim of Bayesian estimation is to derive the conditional PDF $p(k|Y)$ and compute an optimal estimate of k. From the previous section, $p(k|Y)$ can be evaluated as,

$$p(k|Y) \propto p(Y|k)p(k), \tag{9}$$

where the likelihood is given by Eq. (6). F in this case denotes a direct numerical solver of Eq. (8) that computes temperatures at thermocouple location at a given value of k. In the current study, it is assumed that the magnitude of the numerical error is much less than that of the measurement noise. p(k) herein is assumed to take the form,

$$p(k) \propto \exp\{-\frac{(k - \bar{k})^2}{2\upsilon_k}\}, \quad \text{when} \quad k > 0, \quad \text{and} \quad 0 \quad \text{otherwise} \tag{10}$$

where $\bar{k}$ and v_k are the mean and variance of a normal distribution. This is in fact a renormalized normal distribution to enforce the non-negativity of k. Then the PPDF can be evaluated as,

$$p(k|Y) \propto \exp\{-\frac{(Y-F(k))^T(Y-F(k))}{2v_T}\}\exp\{-\frac{(k-\bar{k})^2}{2v_k}\}, \quad \text{when} \quad k>0, \quad \text{and} \quad 0 \quad \text{otherwise} \tag{11}$$

Eq. (11) can be interpreted as a balance between the prior belief of unknown parameter and the information contained in data (likelihood). More precise prior model or more accurate measurements can both lead to a better posterior estimate. Hence, the advantages of the above formulation over likelihood inference are i.) when the number of measurements is limited, accurate posterior estimate is still possible by updating the prior PDF, and ii.) the prior belief of parameter is able to correct the effects of biased data.

It is not possible to have strong prior belief as Eq. (10) in all circumstances. A more reasonable approach is to relax the prior assumption on k by treating $\bar{k}$ and v_k as random variables. A hierarchical Bayesian PPDF is hence used,

$$p(k,\bar{k},v_k|Y) \propto p(Y|k)p(k|\bar{k},v_k)p(\bar{k})p(v_k). \tag{12}$$

A natural way to select priors for hyper-parameters $\bar{k}$ and v_k is to use conjugate priors [16]. In the current study, local uniform distribution and inverse Gamma distribution are assigned to $\bar{k}$ and v_k, respectively. Eq. (12) can then be evaluated as,

$$p(k,\bar{k},v_k|Y) \propto \exp\{-\frac{(Y-F(k))^T(Y-F(k))}{2v_T}\}v_k^{-1/2}\exp\{-\frac{(k-\bar{k})^2}{2v_k}\}v_k^{-(1+\alpha)}\exp\{-\beta v_k^{-1}\},$$

$$\text{when} \quad k\in(0,\infty)\cap\bar{k}\in(0,k_{max})\cap v_k\in(0,\infty), \quad \text{and} \quad 0 \quad \text{otherwise} \tag{13}$$

where k_{max} is the possible maximum value of $\bar{k}$, which can be an arbitrary large number, and α and β are the parameters of the inverse Gamma distribution. For a similar reason, v_T can be treated as unknown as well since it is rather difficult to quantify the magnitude of measurement noise directly from data. In this case, the augmented Bayesian PPDF is used as,

$$p(k,\bar{k},v_k,v_T|Y) \propto v_T^{-n/2}\exp\{-\frac{(Y-F(k))^T(Y-F(k))}{2v_T}\}v_k^{-1/2}\exp\{-\frac{(k-\bar{k})^2}{2v_k}\}v_k^{-(1+\alpha)}$$

$$\exp\{-\beta v_k^{-1}\}v_T^{-(1+\alpha_1)}\exp\{-\beta_1 v_T^{-1}\}$$

$$\text{when} \quad k\in(0,\infty)\cap\bar{k}\in(0,k_{max})\cap v_k\in(0,\infty)\cap v_T\in(0,\infty), \quad \text{and} \quad 0 \quad \text{otherwise} \tag{14}$$

The above PPDFs are implicit due to the presence of numerical solver F, hence can only be evaluated up to some normalizing constants. Numerical sampling strategies are introduced later in this paper to explore the posterior state spaces. As the end of discussion of Bayesian parameter estimation, it is necessary to point out that all distributions chosen in above formulations are general but not unique. The selection of distributions for measurement noise and hence the priors can be varied according to the nature of the uncertainties in different scenarios.

INVERSE HEAT CONDUCTION PROBLEMS (IHCP)

A schematic of the inverse heat conduction problem (IHCP) is shown in Fig. 1. Given Eq. (8) and the following conditions,

$$T(\mathbf{x},t) = T_g, \quad \text{on} \quad \Gamma_g, \quad t\in[0,t_{\max}] \tag{15}$$

$$k\frac{\partial T(\mathbf{x},t)}{\partial \mathbf{n}} = q_h, \quad \text{on} \quad \Gamma_h, \quad t\in[0,t_{\max}] \tag{16}$$

$$T(\mathbf{x},0) = T_0(\mathbf{x}), \quad \text{in} \quad \Omega \tag{17}$$

the main objective of the IHCP is the calculation of the heat flux q_0 on $\Gamma_0\times[0,t_{\max}]$ from temperature measurements Y at a given number of points (temperature sensor locations) within the domain.

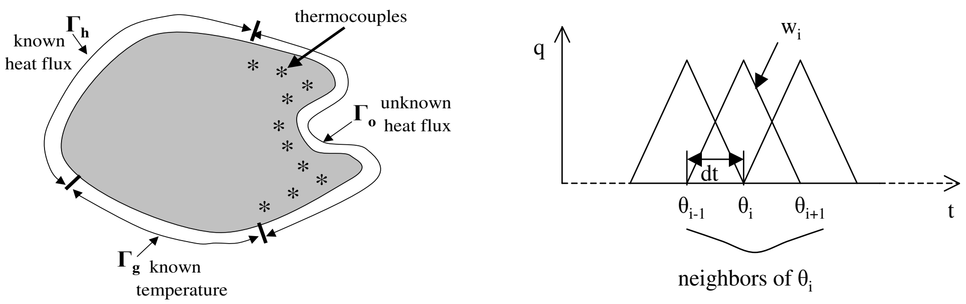

Figure 1: Inverse heat conduction problem (IHCP). The left figure shows schematic of IHCP and the right figure shows the basis function for discretization of q_0 and neighborhood definition for θ in 1D IHCP.

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. Thus the unknown q_0 can be written as:

$$q_0(\mathbf{x},t) = \sum_{i=1}^{m} \theta_i w_i(\mathbf{x},t) \tag{18}$$

where $w_i's$ are the pre-defined basis functions. The problem is then transformed to the estimation of the weights $\theta_i's$. These weights are considered to be represented by an unknown random vector θ of length m (number of basis functions).
The likelihood of IHCP has the form of Eq. (6) provided the measurement noise is Gaussian. In this case, F denotes the direct numerical solver that computes temperatures at thermocouple locations at given value of θ. It is again assumed that the numerical errors are much less than the measurement noise. MRF (Eq. 7) is used to model the prior distribution of θ. Hence, the joint PPDF is,

$$p(\theta|Y) \propto \exp\{-\frac{1}{2v_T}[F(\theta)-Y]^T[F(\theta)-Y]\} \cdot \exp\{-\frac{1}{2}\lambda\theta^T W\theta\}. \tag{19}$$

It has to be point out that the hierarchical and augmented Bayesian formulations presented in the parameter estimation section can be applied to the IHCP as well.
Having the implicit PPDF functions, one will need to explore the posterior state space to compute the optimal estimates of the unknowns. A Monte Carlo simulation approach towards this goal is taken in the current study.

MARKOV CHAIN MONTE CARLO (MCMC) SIMULATION

The idea of Monte Carlo simulation is to draw an i.i.d set of samples $\{\theta^{(i)}\}_{i=1}^{L}$ from a target PDF $p(\theta)$ ($p(\theta|Y)$ in current study) defined on a high dimensional space R^m or having an implicit form [20]. These L samples can be used to approximate the target density with the following empirical point-mass function:

$$p_L(\theta) = \frac{1}{L}\sum_{i=1}^{L} \delta_{\theta^{(i)}}(\theta) \tag{20}$$

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:

$$E_L(f) = \frac{1}{L}\sum_{i=1}^{L} f(\theta^{(i)}) \longmapsto_{L\rightarrow\infty} E(f) = \int f(\theta)p(\theta)d\theta \tag{21}$$

The L samples can also be used to obtain the MAP estimate of θ as follows:

$$\hat{\theta}_{MAP} = argmax_{\theta^{(i)}} \;\; p(\theta^{(i)}) \tag{22}$$

MCMC is a strategy for generating samples $\theta^{(i)}$ while exploring the state space of θ using a Markov chain mechanism [20]. 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 samples from $p(\theta)$ directly, but can evaluate $p(\theta)$ up to a normalizing constant.
The Metropolis-Hastings (MH) algorithm is the most basic form of an MCMC algorithm [21]. For a target PDF $p(\theta)$, the MH algorithm samples a candidate $\theta^{(*)}$ from a proposal PDF $q(\theta^{(*)}|\theta^{(i)})$ at each iteration, where $\theta^{(i)}$ is the sample at the previous iteration. The chain moves to the next state $(\theta^{(*)})$ with an acceptance probability A defined as $A(\theta^{(*)}, \theta^{(i)}) = min\{1, \frac{p(\theta^{(*)})q(\theta^{(i)}|\theta^{(*)})}{p(\theta^{(i)})q(\theta^{(*)}|\theta^{(i)})}\}$. The pseudo-code is the following:

1. Initialize $\theta^{(0)}$
2. For $i = 1 : \text{Nmcmc} - 1$

   — sample $u \sim U(0, 1)$

   — sample $\theta^{(*)} \sim q(\theta^{(*)}|\theta^{(i)})$

   — if $u < A(\theta^{(*)}, \theta^{(i)}) = \min\{1, \frac{p(\theta^{(*)})q(\theta^{(i)}|\theta^{(*)})}{p(\theta^{(i)})q(\theta^{(*)}|\theta^{(i)})}\}$

   $\theta^{(i+1)} = \theta^{(*)}$

   — else

   $\theta^{(i+1)} = \theta^{(i)}$

By its construction, the MH algorithm guarantees that the chain converges to the target PDF $p(\theta)$ for any proposal PDF, however, careful design of the proposal PDF will accelerate the convergence speed.
The symmetric sampler, which is intensively used in the current study, is a special case of the MH algorithm. It uses the proposal PDF of $q(\theta^{(*)}|\theta^{(i)}) = q(\theta^{(i)}|\theta^{(*)})$ and, hence, the acceptance ratio simplifies to $A(\theta^{(*)}, \theta^{(i)}) = min\{1, \frac{p(\theta^{(*)})}{p(\theta^{(i)})}\}$.
The Gibbs sampler [22] is the most widely used MCMC algorithm. It emphasizes the spatial ingredient of MCMC algorithms [23]. For a m-dimensional random vector θ, the full conditional PDF is defined as $p(\theta_i|\theta_{-i})$, where θ_{-i} stands for $\{\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_m\}$. When the full conditional PDF is known, it is often advantageous to use it as the proposal PDF. The algorithm is summarized as,

1. Initialize $\theta^{(0)}$
2. For $i = 1 : \text{Nmcmc} - 1$

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

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

   — $\vdots$

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

The important feature of this sampler is that the acceptance probability A is always 1. It is thus adopted in the current study of the IHCP. For the Bayesian formulation Eq. (19), the relation between the unknown vector θ and observation Y can be written as,

$$Y = H\theta + T_I + \omega, \tag{23}$$

where H is the sensitivity matrix determined as follows:

$$H((i-1)N + j, k) = T_H(\hat{\mathbf{x}}_i,\ \hat{t}_j; w_k),\ i = 1 : M,\ j = 1 : N,\ k = 1 : m. \tag{24}$$

In the above equation, T_H denotes the direct solution with zero initial temperature condition, zero boundary temperature condition on Γ_g, zero heat flux on Γ_h and heat flux w_k on Γ_0. T_I denotes the direct solution with zero heat flux on Γ_o, the known initial conditions and the known boundary conditions on Γ_g and Γ_h. M and N are the number of thermocouples and number of measurements at each sensor location, respectively. 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:

$$p(\theta_i|\theta_{-i}) \sim N(\mu_i, \sigma_i^2), \quad \mu_i = \frac{b_i}{2a_i}, \quad \sigma_i = \sqrt{\frac{1}{a_i}} \tag{25}$$

$$a_i = \sum_{s=1}^{N} \frac{H_{si}^2}{\sigma_T^2} + \lambda W_{ii}, \quad b_i = 2\sum_{s=1}^{N} \frac{\mu_s H_{si}}{\sigma_T^2} - \lambda\mu_p, \quad \mu_s = Y_s - \sum_{t\neq i} H_{st}\theta_t, \quad \mu_p = \sum_{j\neq i} W_{ji}\theta_j + \sum_{k\neq i} W_{ik}\theta_k \tag{26}$$

EXAMPLES

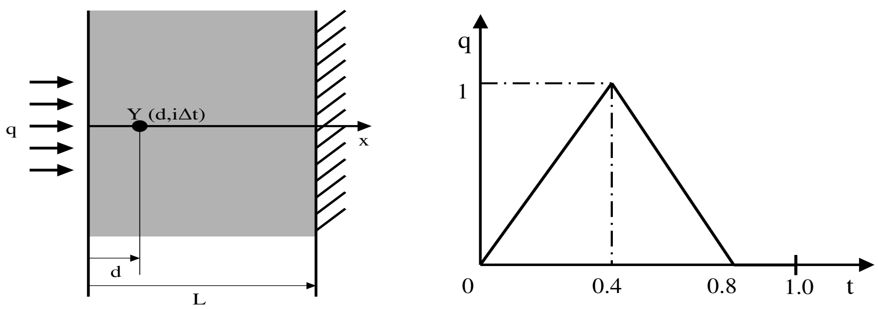

Figure 2: 1D heat conduction. The left figure is the schematic of 1D heat conduction. The right one is the profile of boundary heat flux.

1. Example I The first example studied is the estimation of the thermal conductivity k of a conducting solid. Let us consider the experiment in Fig. 2. The body has zero initial temperature and is insulated at the right end (x=L). A triangular profile heat flux q(t) is applied at the left end (x=0). Temperatures are recorded at x=d. To simplify the discussion, the numerical study is conducted in a dimensionless manner as,

$$\frac{\partial T}{\partial t} = k\frac{\partial^2 T}{\partial x^2}, \quad 0 < t < 1, \quad 0 < x < 1 \tag{27}$$

$$T(x,0) = 0, \quad 0 \leq x \leq 1, \quad \frac{\partial T}{\partial x}|_{x=1} = 0, \quad \frac{\partial T}{\partial x}|_{x=0} = q(t), \quad 0 < t < 1 \tag{28}$$

The simulation data are generated by adding Gauss random noise with mean 0 and variance v_T (standard deviation σ_T), to the computed temperatures at a true value of k, which is randomly generated from a normal distribution with mean $\bar{k}$ and variance v_k. The Metropolis-Hastings algorithm is used in this example. To explore the PPDF Eq. (11), a symmetric proposal PDF $q(\cdot|k^{(i)}) \sim N(k^{(i)}, v_{kq})$ is used, where v_{kq} is specified as 10% of the mean value. For Eqs. (13) and (14), the proposal PDFs for all random variables have the same structure. However, for PPDFs in Eqs. (13) and (14), a strategy to update one variable each time is taken to increase the acceptance probability. α, β, α_1 and β_1 all have values of $1.0e-3$.
In this example, $\bar{k}$ and v_k are taken as 1.0 and 0.15, respectively, and a value of 1.2146 is generated as true k. In Table 1, the Bayesian estimates using different formulations at different data are listed. $\hat{k}_{postmean}$ is the posterior mean estimate and $\sigma_{\hat{k}}$ is the standard deviation of the posterior distribution. σ_T in Table 1 refers to the standard deviation used to generate the temperature data. The posterior densities of k in all listed cases are plotted in Fig. 3. 20000 runs of the MH sampler are conducted for each case and the last 15000 samples are used to compute the estimates. It is clear that the posterior mean estimates are rather accurate and stable at different thermocouple locations and measurement noise.

Table 1. Bayesian estimates of k using different models

case#	Bayesian model	prior of k	# of data	σ_T	$\hat{k}_{postmean}$	$\sigma_{\hat{k}}$
1	simple (Eq. 11)	normal	50	0.005	1.2094	0.0263
2	simple (Eq. 11)	normal	50	0.001	1.2209	0.0055
3	simple (Eq. 11)	normal	100	0.005	1.2268	0.0196
4	simple (likelihood)	uniform	50	0.005	1.2164	0.0269
5	hierarchical ($\bar{k}$ unknown) (Eq. 13)	normal	50	0.005	1.2140	0.0271
6	hierarchical ($\bar{k}$ and v_k unknown) (Eq. 13)	normal	50	0.005	1.2157	0.0269
7	hierarchical + augmented (Eq. 14)	normal	50	0.005	1.2182	0.0446

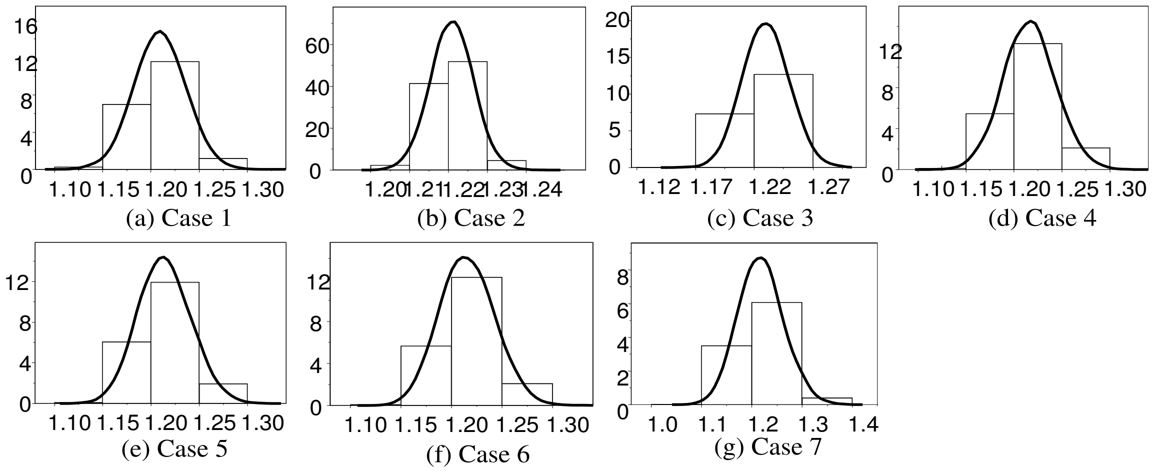

Figure 3: Posterior density of k.

However, increasing the number of measurements or decreasing the magnitude of the measurement error can both reduce the standard deviation of the posterior distribution. In the first 4 cases, the uniform prior leads to the best estimate. This can be explained by the fact that the normal prior (mean 1.0 and standard deviation 0.15) is rather biased in representing the true value of k whereas the data contains more accurate information of k. In addition, case 7 enforces almost no assumption on the uncertainties, however, it still provides rather accurate estimate. Meanwhile, the posterior mean estimate of σ_T is 0.0083 in case 7.

2. Example II Considering the same problem as in example I except that k is known to be 1.0 and $q(t)$ is unknown. Four cases are studied in this example. In each case, 100 measurements are simulated. The results are plotted in Fig. 4. The results are obtained by running 50000 iterations of Gibbs sampler and the last 25000 samples are used in the estimation. λ in each case is chosen using heuristic Tikhonov method [14]. It is observed that the posterior mean estimates (same as MAP estimates in this linear problem) are rather accurate in all cases and stable to the sensor location, magnitude of noise and discretization of $q(t)$.

CONCLUSIONS

A Bayesian inference approach was introduced for the solution of inverse problems in heat transfer processes. Both parameter estimation and IHCP problems are studied using various Bayesian formulations. MCMC-based numerical sampling strategy was adopted to exploit the posterior state space. The proposed techniques were shown through numerical examples. They lead to not only a point estimate of unknown quantities but also to a full estimate of their statistical information and quantification of the system uncertainties. In addition, the inverse problem is regularized statistically through the modeling of prior distribution. The methodology presented is applicable to diverse inverse prob-

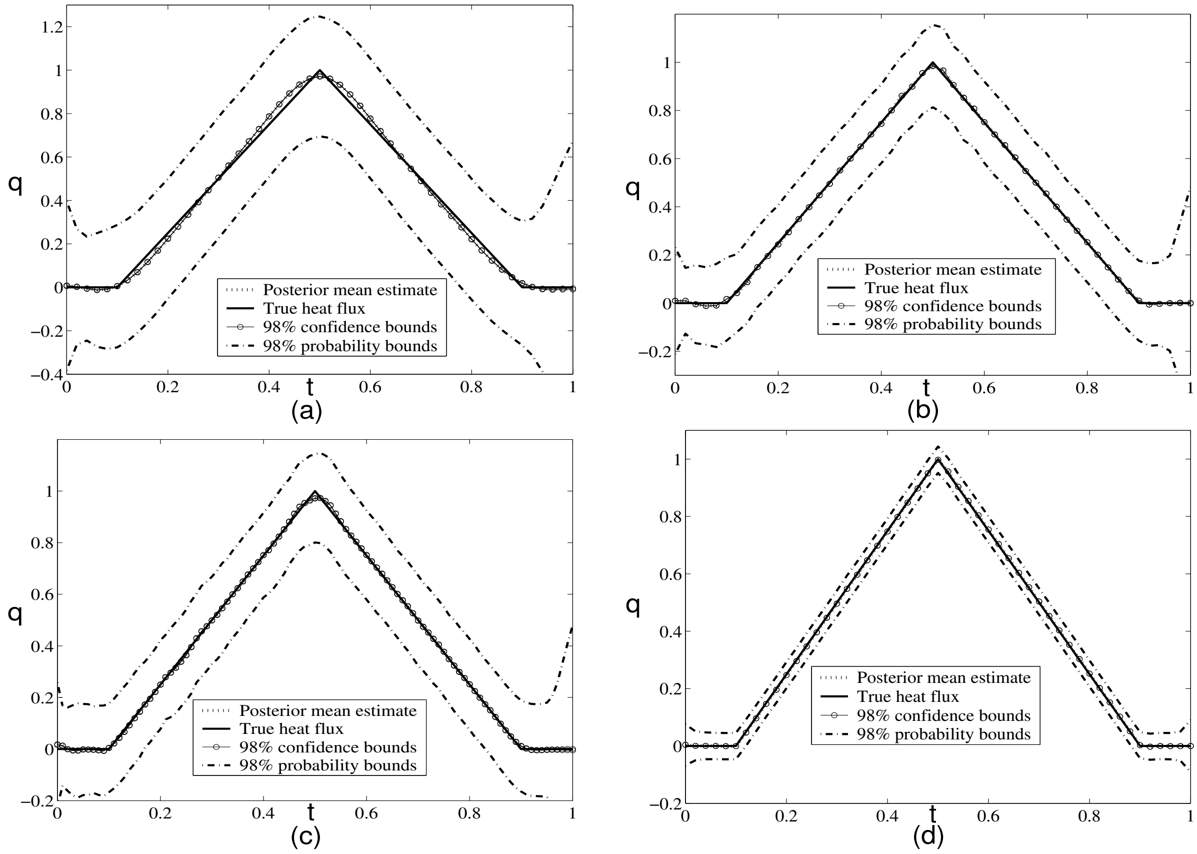

Figure 4: Posterior mean estimates of q(t) and associated 98% probability and confidence bounds. (a) d=0.5, $\sigma_T = 0.01$, m=51 (b) d=0.5, $\sigma_T = 0.001$, m=51. (c) d=0.5, $\sigma_T = 0.001$, m=101. (d) d=0.1, $\sigma_T = 0.001$, m=51.

lems and provides a potential approach to address stochastic design/control problems in continuum systems.

ACKNOWLEDGEMENTS

Partial support for the work reported here was provided by the U.S. Air Force Office of Scientific Research (grant F49620-00-1-0373) and by the U.S. National Science Foundation (grant DMI-0113295). This research was conducted using the resources of the Cornell Theory Center.

References

[1] J.V. Beck, K.J. Arnold, *Parameter Estimation in Engineering and Science*, New York : Wiley (1977).

[2] O.M. Alifanov, *Inverse Heat Transfer Problems*, Springer-Verlag, Berlin (1994).

[3] K.A. Woodbury (Editor), *Inverse Engineering Handbook*, CRC Press, 1st edition (2002).

[4] A.N. Tikhonov, *Solution of Ill-posed Problems*, Halsted Press, Washington (1977).

[5] T.A. Zang, M.J. Hemsch, M.W. Hilburger, S.P. Kenny, J.M. Luckring, P. Maghami, S.L. Padula, W.J. Stroud, *Needs and pportunities for uncertainty-based multidisciplinary design methods for aerospace vehicles*, NASA/TM-2002-211462 technical report series, Langley Research Center, Hampton, Virginia (2002).

[6] L. Huyse, *Solving problems of optimization under uncertainty as statistical decision problems*, AIAA paper 2001-1519 in 42^{nd} *AIAA/ASME/ASCE/AHS/ASC structures, Structural Dynamics and Materials Conference and Exhibit*, Seattle, WA (2001).

[7] P.M. Norris, *Application of experimental design methods to assess the efect of uncertain boundary conditions in inverse heat transfer problems*, Int. J. Heat Mass Transfer, 41(2), (1998), 313-322.

[8] B.F. Blackwell, K.J. Dowding, *Sensitivity and Uncertainty Analysis for Thermal Problems*, *Preceedings of 4th International Conference on Inverse Problems in Engineering*, Rio de Janeiro, Brazil (2002).

[9] T.D. Fadale, A.V. Nenarokomov, A.F. Emery, *Uncertainties in parameter estimation: the inverse problem*, Int. J. of Heat Mass Transfer, 38, (1995), 511-518.

[10] A.F. Emery, A.V. Nenarokomov, T.D. Fadale, *Uncertainties in parameter estimation: the optimal experiment design*, Int. J. of Heat Mass Transfer, 43, (2000), 3331-3339.

[11] V.A. Badri Narayanan, N. Zabaras, *Stochastic inverse heat conduction using a spectral approach*, Int. J. Numer. Meth. Eng., in press.

[12] C. Ferrero, K. Gallagher, *Stochastic thermal history modeling, 1. constraining heat flow histories and their uncertainty*, Marine and Petroleum Geology, 19, (2002), 633-648.

[13] N. Leoni, C.H. Amon, *Bayesian surrogates for integrating numerical, analytical and experimental data: application to inverse heat transfer in wearable computers*, IEEE Transactions Comps. Pack. Manuf. Technology, 23, (2000), 23-33.

[14] J. Wang, N. Zabaras, A Bayesian inference approach to the stochastic inverse heat conduction problem, *International Journal of Heat and Mass transfer*, accepted for publication.

[15] J. Wang, N. Zabaras, *Using Bayesian statistics in the estimation of heat sources in radiation*, Int. J. of Heat Mass transfer, submitted for publication.

[16] C.P. Robert, *The Bayesian Choice, From Decision-Theoretic Foundations to Computational Implementation*, Second Edition, Springer (2001).

[17] G.E.P. Box, G.C. Tiao, *Bayesian Inference in Statistical Analysis*, New York : Wiley (1992).

[18] P. Congdon, *Bayesian Statistical Modeling*, John Wiley & Sons : New York (2001).

[19] J. Besag, P. Green, D. Higdon, K. Mengersen, *Bayesian Computation and Stochastic Systems*, Statistical Science, 10, (1995), 3-41.

[20] C. Andrieu, N.D. Freitas, A. Doucet, M.I. Gordan, *An introduction to MCMC for machine learning*, Machine Learning, 50, (2003), 5-43.

[21] W.R. Gilks (ed), S. Richardson (ed), D.J. Spiegelhalter (ed), *Markov Chain Monte Carlo in Practice*, Chapman & Hall Ltd, New York (1996).

[22] S. Geman, D. Geman, *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*, Transactions on Pattern Analysis and Machine Intelligence, 6(6), (1984), 721741.

[23] J. Besag, P.J. Green, *Spatial statistics and Bayesian computation*, Journal of the Royal Statistical Society, Series B, Methodological, 55, (1993), 25-37.