

An efficient Markov chain Monte Carlo simulation of a stochastic inverse radiation problem

Jingbo Wang^a and Nicholas Zabaras^b

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

^ajw332@cornell.edu, ^bzabaras@cornell.edu

Abstract

A novel methodology that combines recent advances in computational statistics and reduced-order modeling is presented to explore the application of Bayesian statistical inference to a stochastic inverse problem in radiative heat transfer. The underlying objective of this work is to reveal the potential of using statistical approaches, mainly Bayesian computational statistics and spatial statistics, to solve data-driven stochastic optimization and uncertainty quantification problems raised in various complex continuum system design and control applications when the robustness and reliability requirements are critical. In this work, an unknown transient heat source in a three-dimensional participating medium is reconstructed from the temperature measurements. The heat source is modeled as a stochastic process, of which the joint posterior probability density function (PPDF) is computed using the Bayes' formula. Random errors in thermocouple readings are modeled as Gauss random variables. 'Maximum A Posteriori' (MAP) and posterior mean estimates of the heat source are then computed. A Markov chain Monte Carlo (MCMC) sampler composed of a cycle of symmetric MCMC kernels is designed to explore the posterior state space numerically. To expedite the sampling speed, a proper orthogonal decomposition (POD) based reduced order modeling technique is used in the likelihood computation. Typical heat source profiles are reconstructed using the simulated data to demonstrate the presented methodology. The results indicate that the Bayesian inference method can provide accurate point estimates as well as uncertainty quantification to the solution of the inverse radiation problem.

1 Introduction

Robustness and reliability have become the focus of modern design and control technology, towards which the development of efficient stochastic optimization and uncertainty quantification methodologies and algorithms is critical [1]. The majority of these optimization problems are data-driven in nature. Hence, the key steps in obtaining robust solutions are statistical data reduction, probabilistic modeling of uncertainties and quantification of uncertainty propagation in continuum systems [2]. Among various techniques being developed towards these goals, statistical learning is a rigorous approach for the reasons that it is non-deterministic in nature, it is able to generate complete probabilistic description of uncertainties and design solutions, it

dynamically interacts with data, and it only requires deterministic solvers to compute the system response [3, 4]. In particular, most recent advances in Bayesian computational statistics [5, 6] and spatial statistics [7] provide more robust formulations for stochastic inverse problems, rich family of data kringing models and efficient numerical algorithms.

Despite all of the advantages of statistical learning, the application to stochastic inverse problems in complex continuum systems is restricted by the extensive computation cost. The degrees of freedom (DOF) in the direct simulation of a complex continuum system can easily reach hundreds of thousands or even higher. Usually the direct simulation involves solving coupled nonlinear partial differential equations (PDEs). Furthermore, the underlying multi-physical processes may happen over a range of length scales, therefore, multi-scale simulation algorithms have to be implemented. On the other hand, a computational statistics approach to optimization requires multiple runs of such direct simulation. To overcome this computation obstacle, reduced order-modeling techniques have to be introduced to reduce the direct simulation cost [8, 9].

This work aims at introducing Bayesian computation and spatial statistics to the multi-disciplinary engineering design/control community and demonstrating the feasibility of extending these advanced stochastic optimization techniques to complex design/control problems in continuum systems. The methodology is discussed within the context of inverse heat radiation problem, where heat sources are reconstructed from the temperature measurements within participating media [10]. However, it is rather general and applicable to other inverse problems as well.

Study of inverse thermal radiation problems 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 engine, furnace technology and other [11, 12]. In participating media, radiation is generally accompanied by heat conduction and convection. In a direct simulation, a system of nonlinearly coupled PDEs governing temperature and radiation intensity evolution needs to be solved iteratively. Difficulties arise in the solution of such systems because the heat flux contributed by radiation varies nonlinearly with temperature, the radiation intensity varies in space and direction, and the radiation intensity equation is an integro-differential equation [13]. The direct radiation problem, in which the temperature distribution is computed with prescribed thermal properties, source generation and initial/boundary conditions, is often solved using a combination of spatial discretization methods such as finite volume or finite element methods (FEM) and ordinate approximation such as P_N and S_N methods [13]. Distinctly different from the well-posed direct radiation problem, the inverse problem is in general ill-posed, i.e. its solution is not unique and is unstable to small errors in the given data [14].

The usual solution approaches restate the inverse problem as a least-squares minimization problem [14, 15]. Gradient optimization techniques are introduced and appropriate continuum or discrete sensitivity and/or adjoint problems are required [10, 16]. The ill-posedness of the inverse problem can be addressed using combination of appropriate regularization techniques including Tikhonov regularization [17], the function specification method by Beck et al. [15], Zabaras and Liu [18] and the iterative regularization technique by Alifanov [14].

A new stochastic outlook to inverse thermal problems has recently been introduced using Bayesian inference [19, 20], which can account for uncertainties in the measurement data and is able to provide probabilistic specification to the inverse solutions. Bayesian inference incorporates a priori information regarding the unknowns into a prior distribution model that is then combined with the likelihood to formulate the posterior probability density function (PPDF) [5, 6]. The

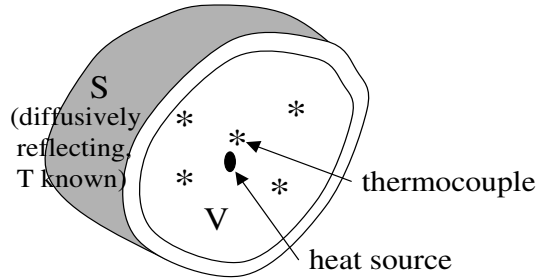


Figure 1: Schematic of the inverse radiation problem. The objective is to compute the point heat source $g(t)$ given initial temperature conditions, temperature boundary conditions on the surface and temperature measurements at a number of points within the domain.

method regularizes the ill-posed inverse problem through prior distribution modeling [21, 22] and in addition provides means to estimate the statistics of uncertainties [20].

With the recent propagation of efficient sampling methods, such as Markov chain Monte Carlo (MCMC) [23], the application of Bayesian inference to engineering inverse problems becomes feasible. MCMC provides 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. Running a Markov chain usually involves repetitive solution of the direct problem and this is the place where reduced-order models can play a critical role [24]. One widely used approach of model reduction is the computation of the proper orthogonal decomposition (POD) basis using the method of snapshots [25, 26].

In this paper, the joint PPDF of the strength of a transient heat source at consecutive time points is computed through temperature measurements in 3D participating media. A MCMC sampler is designed to explore the posterior state space. The kernel of the MCMC sampler is composed of a cycle of simple symmetric MCMC kernels. In each computation of the likelihood, the direct problem is solved using model reduction. The remaining of this paper is organized in the following sequence. Section 2 introduces the inverse radiation problem with an outline of the direct simulation. The formulation of the likelihood is presented in Section 3 together with the prior distribution model and the PPDF under a Bayesian inference framework. The design of the MCMC sampler is discussed in Section 4 including the exploration of the posterior state space. Section 5 discusses the reduced-order modeling implementation. In Section 6, two examples of reconstruction of step and triangular heat source profiles are provided. Finally, Section 7 summarizes the observations of this numerical study.

2 Heat source reconstruction in 3D participating media

The schematic of the problem of interest is given in Fig. 1. Inside the 3D domain V , heat conduction takes place simultaneously with absorption, scattering and emission of the electromagnetic waves. On the boundary surface S , the temperature is known and the electromagnetic waves are diffusively reflected. The governing equations for the temperature and radiation intensity evolution in the domain V are as follows:

$$\rho C_p \frac{\partial T}{\partial t} = k \nabla^2 T - \nabla \cdot \vec{q}_r + g(t) G(x - x^*, y - y^*, z - z^*) \quad (1)$$

$$\vec{s} \cdot \nabla I + (\kappa + \sigma)I - \frac{\sigma}{4\pi} \int_{4\pi} I(\vec{r}, \vec{s}') d\Omega' = \kappa I_b \quad (2)$$

where I_b is the black body radiation intensity determined by Planck function:

$$I_b = \frac{\sigma_b T^4}{\pi} \quad (3)$$

and \vec{q}_r is the heat flux contributed by radiation:

$$\nabla \cdot \vec{q}_r = 4\pi\kappa(I_b - \frac{1}{4\pi} \int_{4\pi} I(\vec{r}, \vec{s}) d\Omega) \quad (4)$$

On the surface S , the following holds:

$$I(\vec{r}, \vec{s}) = \epsilon I_b + \frac{1 - \epsilon}{\pi} \int_{\vec{n} \cdot \vec{s}' < 0} |\vec{n} \cdot \vec{s}'| I(\vec{r}, \vec{s}') d\Omega' \quad \vec{n} \cdot \vec{s} > 0 \quad (5)$$

$$T = T_w \quad (6)$$

In the above equations, T and I denote the temperature and radiation intensity, respectively, \vec{r} is the position vector and \vec{s} is the direction vector. $G(x - x^*, y - y^*, z - z^*)$ is the spatial approximation of a point heat source located at (x^*, y^*, z^*) . In this work, a 3D normal density function is used for G . Ω stands for the solid angle over the entire space. ρ is the density of the medium, C_p is the thermal capacity, k is the thermal conductivity, and κ, σ, ϵ are the absorption coefficient, scattering coefficient and boundary wall emissivity, respectively. Finally, σ_b is the Stefan-Boltzmann constant and \vec{n} is the unit normal vector on S and pointing into the domain.

In the inverse problem of interest, the heat source $g(t)$ is the main unknown. Its calculation becomes feasible by providing the values of the temperature at a given number of locations within the domain as shown in Fig. 1. Let Y denote the measured temperature data, i.e. $Y = [Y_1^{(1)}, Y_2^{(1)}, \dots, Y_M^{(1)}, Y_1^{(2)}, Y_2^{(2)}, \dots, Y_M^{(2)}, \dots, Y_1^{(N)}, Y_2^{(N)}, \dots, Y_M^{(N)}]^T$, where

$$Y_i^{(j)} = T(\vec{r}_i, \hat{t}_j) + \omega \quad (7)$$

where $i = 1, \dots, M, j = 1, \dots, N$ and $\hat{t}_N = t_{\max}$. M and N are the number of thermocouples and number of measurements at each site, respectively. ω is the random measurement noise. The inverse problem is then stated as follows: find an estimate $\hat{g}(t)$ of the real heat source $g(t)$ such that the computed temperatures given $\hat{g}(t)$ can match Y in some optimal sense.

The direct problem must be solved before inverse calculation. The direct problem can be solved using a combination of the finite element method (FEM) in space discretization and the S_4 method in ordinate discretization. The iterative process at each time step to solve the coupled Eqs. (1) and (2) is summarized as follows:

1. Set $T_{guess}^{(i)} = T^{(i-1)}$;
2. Substitute $T_{guess}^{(i)}$ into Eq. (3) to compute I_b ;
3. Solve Eq. (2) for $I^{(i)}$;
4. Use Eq. (4) to compute $\nabla \cdot \vec{q}_r$;
5. Solve Eq. (1) and update $T_{guess}^{(i)}$ with the solution;

6. If the solutions converged, set $T_{guess}^{(i)}$ as $T^{(i)}$ and save $I^{(i)}$; otherwise, go to step 2.
7. Go to the next time step.

Here $T^{(i)}$ denotes the temperature solution at the i^{th} time step (note that $T^{(0)}$ is a known initial temperature) and $T_{guess}^{(i)}$ is the guessed temperature solution. In each iteration of the above procedure, the integro-differential Eq. (2) is solved using the S_4 method [13]. In this approach, the intensity I at each point is discretized into 24 directions. The integration over solid angles (directions) is approximated as weighted sums in these 24 directions. The direction vectors and the associated weights of the S_4 method are specified in [13]. In each direction, the governing equation for I can be written as follows:

$$\vec{s}_i \cdot \nabla I_i + (\kappa + \sigma)I_i - \frac{\sigma}{4\pi} \sum_{j=1}^{24} I_j(\vec{r})w_j = \kappa I_b \quad (8)$$

The associated boundary condition takes the following form:

$$I_i = \epsilon I_b + \frac{1 - \epsilon}{\pi} \sum_{\vec{n} \cdot \vec{s}_j < 0} |\vec{n} \cdot \vec{s}_j| w_j I_j, \quad \vec{n} \cdot \vec{s}_i > 0 \quad (9)$$

where I_i is the intensity in the i^{th} direction and w_j is the weight of the j^{th} direction. For any given temperature field, 24 equations as Eq. (8) with fixed direction vector \vec{s}_i need to be solved iteratively. It is noticed that Eq. (8) has a convection term $\vec{s}_i \cdot \nabla I_i$, hence streamline-upwind/Petrov-Galerkin (SUPG) formulation [27] is used to stabilize the FEM equations. In summary, the weak formulations of temperature Eq. (1) and intensity Eq. (8) can be written as follows:

$$\int_V \rho C_p T^{(i)} W dv + \Delta t \int_V k \nabla T^{(i)} \cdot \nabla W dv = \Delta t \int_V (-\nabla \cdot \vec{q}_r + g(t)G(x - x^*, y - y^*, z - z^*)) W dv + \int_V \rho C_p T^{(i-1)} W dv, \quad (10)$$

and

$$\int_V \vec{s}_i \cdot \nabla I_i \tilde{W} dv + \int_V (\kappa + \sigma) I_i \tilde{W} dv = \int_V \kappa I_b \tilde{W} dv + \int_V \frac{\sigma}{4\pi} \sum_{j=1}^{24} I_j w_j \tilde{W} dv, \quad (11)$$

where W and \tilde{W} are the test functions of classical Galerkin and SUPG formulations, respectively.

Using the above direct simulation framework, the total number of degrees of freedom for the system is $N_n^3 \times 25$, where N_n is the number of nodes in each coordinate. Also note that there are two iteration loops in each time step. Thus, it is expected that a single run of the above full-order direct model solver will take a significant computation time.

3 Bayesian inverse formulation

To introduce the Bayesian formulation, the unknown heat source function is first discretized using linear finite element basis functions in time as follows:

$$\hat{g}(t) = \sum_{i=1}^m h_i(t)\theta_i \quad (12)$$

where h_i 's are the linear FEM basis functions, θ_i 's are the corresponding nodal values of \hat{g} and m is the number of basis functions used.

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

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{p(Y)} \quad (13)$$

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

$$\hat{\theta}_{\text{MAP}} = \underset{\theta}{\text{argmax}} p(\theta|Y) \quad (14)$$

and the posterior mean estimator:

$$\hat{\theta}_{\text{postmean}} = E \theta|Y \quad (15)$$

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

$$p(\theta|Y) \propto p(Y|\theta)p(\theta) \quad (16)$$

The likelihood function can be obtained from the following relationship,

$$Y = F(\theta) + \omega \quad (17)$$

where F is a direct solver that computes temperature at sensor locations given θ (Eq. (12)). Here, ω is a vector that contains identical independently distributed (i.i.d.) random errors generated from a normal distribution with mean 0 and standard deviation (std) σ_T . Subsequently, the likelihood can be written as,

$$p(Y|\theta) = \frac{1}{(2\pi)^{n/2}\sigma_T^n} \exp\left\{-\frac{(Y - F(\theta))^T(Y - F(\theta))}{2\sigma_T^2}\right\} \quad (18)$$

The prior distribution reflects the knowledge, if there is any, of the heat source, before Y is gathered. From an inverse point of view, the prior distribution model provides regularization to the ill-posed inverse problem [20]. In the current study, a specific form of Markov random fields (MRF) [28] is adopted for the prior distribution modeling of θ by treating the temporal direction as another (additional) 'spatial' coordinate. In general, the MRF can be mathematically expressed as follows:

$$p(\theta) \propto \exp\left\{-\sum_{i \sim j} W_{ij}\Phi(\gamma(\theta_i - \theta_j))\right\} \quad (19)$$

where γ is a scaling parameter, Φ 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 on the finite

element discretization of θ , and W'_{ij} s are specified non-zero weights [6]. Let $\Phi(u) = \frac{1}{2}u^2$, the MRF can then be rewritten as:

$$p(\theta) \propto \lambda^{m/2} \exp\left(-\frac{1}{2}\lambda\theta^T W\theta\right) \quad (20)$$

In the one-parameter model of Eq. (20), 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 sites adjacent to site i . λ is a constant that controls the scaling of distribution of θ .

With the specified likelihood function in Eq. (18) and prior pdf in Eq. (20), the PPDF for the inverse problem can then be formulated as,

$$p(\theta|Y) \propto \exp\left(-\frac{1}{2\sigma_T^2}[F(\theta) - Y]^T[F(\theta) - Y]\right) \cdot \exp\left(-\frac{1}{2}\lambda\theta^T W\theta\right) \quad (21)$$

Both point estimates of MAP (Eq. (14)) and posterior mean (Eq. (15)) and probability bounds of the posterior distribution are computed based on this formulation.

4 MCMC sampler

The above introduced PPDF has an implicit form (likelihood has a numerical solver), hence, can only be explored numerically 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 θ by the sample mean and sample statistics from a large set of i.i.d. samples $\{\theta^{(i)}, i = 1 : L\}$ drawn from the target distribution $p(\theta)$,

$$E_L f(\theta) = \frac{1}{L} \sum_{i=1}^L f(\theta^{(i)}) \xrightarrow{L \rightarrow \infty} E f(\theta) = \int f(\theta) p(\theta) d\theta \quad (22)$$

The convergence is guaranteed by the strong law of large number. Therefore, the posterior mean estimate of Eq. (21) can be obtained through the above approximation, and the MAP estimate can be approximated as:

$$\hat{\theta}_{MAP} = \underset{\theta^{(i)}}{\operatorname{argmax}} p(\theta^{(i)}) \quad (23)$$

Notice that the target distribution 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) algorithms provide such sampling strategy using the Markov chain mechanism [23] [29]. The basic form of MCMC, Metropolis-Hastings (MH) algorithm [30], is reviewed here:

1. Initialize $\theta^{(0)}$
2. For $i = 0 : \text{Nmcmc} - 1$
 - sample $u \sim U(0, 1)$
 - sample $\theta^{(*)} \sim q(\theta^{(*)}|\theta^{(i)})$
 - if $u < A(\theta^{(*)}, \theta^{(i)}) = \min\left\{1, \frac{p(\theta^{(*)})q(\theta^{(i)}|\theta^{(*)})}{p(\theta^{(i)})q(\theta^{(*)}|\theta^{(i)})}\right\}$
 - $\theta^{(i+1)} = \theta^{(*)}$
 - else
 - $\theta^{(i+1)} = \theta^{(i)}$

In the above algorithm, Nmcmc is the total number of runs, u is a random number sampled from the standard uniform distribution $U(0, 1)$, $p(\theta)$ is the target distribution (PPDF here) and $q(*|\theta^{(i)})$ is a proposal pdf that has standard form and generates candidate sample based upon the previous sample. By its design, the algorithm guarantees that the samples will converge to the target distribution for any proposal distribution.

In this study, a modified MH sampler is designed to update the vector θ one component at each time. The following notation is introduced:

$$\theta_{-j}^{(i+1)} = \{\theta_1^{(i+1)}, \theta_2^{(i+1)}, \dots, \theta_{j-1}^{(i+1)}, \theta_{j+1}^{(i)}, \dots, \theta_m^{(i)}\}$$

in which, superscript (i) here refers to the i^{th} sample and the subscript j refers to the j^{th} component. The sampler is designed as follows:

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

For $j = 1 : m$

- sample $u \sim U(0, 1)$
- sample $\theta_j^{(*)} \sim q_j(\theta_j^{(*)} | \theta_{-j}^{(i+1)}, \theta_j^{(i)})$
- if $u < A(\theta_j^{(*)}, \theta_j^{(i)})$
 - $\theta_j^{(i+1)} = \theta_j^{(*)}$
- else
 - $\theta_j^{(i+1)} = \theta_j^{(i)}$,

where, $A(\theta_j^{(*)}, \theta_j^{(i)}) = \min\left\{1, \frac{p(\theta_j^{(*)} | \theta_{-j}^{(i+1)}) q(\theta_j^{(i)} | \theta_j^{(*)}, \theta_{-j}^{(i+1)})}{p(\theta_j^{(i)} | \theta_{-j}^{(i+1)}) q(\theta_j^{(*)} | \theta_j^{(i)}, \theta_{-j}^{(i+1)})}\right\}$ and

$$q_j(\theta_j^{(*)} | \theta_{-j}^{(i+1)}, \theta_j^{(i)}) = \frac{1}{\sqrt{2\pi}\sigma_{qj}} \exp\left\{-\frac{1}{2\sigma_{qj}^2}(\theta_j^{(*)} - \theta_j^{(i)})^2\right\} \quad (24)$$

with σ_{qj} is the std of the j^{th} proposal pdf. The reason for updating a single component of θ at each MCMC step is to improve the acceptance probability. Using properly chosen σ_{qj} in each of the proposal pdfs, the entire posterior state space can be visited with a reasonable acceptance ratio and with fast convergence speed. This sampler is essentially a cycle of m symmetric MCMC samplers [23].

Since each run of MH requires a direct computation of the transient temperature field, it is essential to introduce reduced-order modeling.

5 Reduced-order modeling

For the convenience of reduced-order modeling implementation, the direct problem is separated into a homogeneous part and an inhomogeneous part, i.e. $T = T^I + T^h$ and $I = I^I + I^h$. The homogeneous fields T^h and I^h are governed by:

$$\rho C_p \frac{\partial T^h}{\partial t} = k \nabla^2 T^h - \nabla \cdot \vec{q}_r + g(t)G(x - x^*, y - y^*, z - z^*) \quad (25)$$

$$\vec{s} \cdot \nabla I^h + (\kappa + \sigma)I^h - \frac{\sigma}{4\pi} \int_{4\pi} I^h(\vec{r}, \vec{s}') d\Omega' = \kappa I_b - \kappa I_b^I \quad (26)$$

$$I^h = \frac{1 - \epsilon}{\pi} \int_{\vec{n} \cdot \vec{s}' < 0} |\vec{n} \cdot \vec{s}'| I^h(\vec{r}, \vec{s}') d\Omega', \quad \vec{n} \cdot \vec{s} > 0 \quad (27)$$

$$T^h = 0, \quad \text{on } S \quad (28)$$

The governing equations of inhomogeneous fields T^I and I^I can be obtained by subtracting equations (25) - (28) from complete equations in section 2. The reduced-order models are constructed for homogeneous T^h and I^h since the inhomogeneous fields only need to be solved once in the inverse procedure.

The POD method is considered in the current work for the reduced-order modeling. In this approach, the direct simulation result at each time step is expressed as a linear combination of a set of orthonormal basis functions. The coefficients associated with the basis field functions are then computed from the solution of ordinary differential equations (ODEs) derived by Galerkin projection. The basis functions can be extracted from computational or experimental database through solving the following eigenvalue problem [25]:

$$\frac{1}{N_e} \sum_{i=1}^{N_e} \int_V U^{(i)} U^{(i)} \Psi dv = \mu \Psi \quad (29)$$

where $U^{(i)}$ is the i^{th} field function (temperature or intensity field) from the database, N_e is the size of the database, μ is the eigenvalue of operator $K\Psi = \frac{1}{N_e} \sum_{i=1}^{N_e} \int_V U^{(i)} U^{(i)} \Psi dv$ and Ψ is the corresponding eigenfunction. In this study, the basis is obtained using the method of snapshots as follows:

- Take an ensemble set $\{U^{(1)}, U^{(2)}, \dots, U^{(N_e)}\}$, where $U^{(i)}$ is the full-model solution of the PDEs at the i^{th} time step. For temperature, $U^{(i)}$ is in fact $T^h(t = i\Delta t)$. For intensity, $U^{(i)}$ is $I^h(t = i\Delta t)$.
- Solve the eigenvalue problem $CV = V\mu$, where C is an $N_e \times N_e$ matrix with $C_{ij} = \frac{1}{N_e} \int_V U^{(i)} U^{(j)} dv$, μ is a $N_e \times N_e$ diagonal matrix with the i^{th} diagonal entry μ_i is the i^{th} eigenvalue of C , and the corresponding eigenvector V_i is the i^{th} column of $N_e \times N_e$ matrix V .
- Compute the basis functions as $\Psi_i = \sum_{j=1}^{N_e} V_i(j) U^{(j)} / (N_e \mu_i)$.

The set $\{\Psi_1, \Psi_2, \dots, \Psi_{N_e}\}$ is orthonormal [25]. Note that the intensity I^h is a function of both space and orientation, therefore, the volume integration in Eq. (29) and the followed eigenvalue analysis should be extended to $\int_V \int_{4\pi} dv d\Omega$ for model reduction of I^h .

The idea of POD-based model reduction is to express the solutions of Eqs. (25)-(28) as linear combinations of the basis functions. Let $\{\Psi_1^T, \Psi_2^T, \dots, \Psi_{K_T}^T\}$ denote the basis functions of T^h and $\{\Psi_1^I, \Psi_2^I, \dots, \Psi_{K_I}^I\}$ denote the basis functions of I^h , where K_T and K_I are the numbers of basis functions used for temperature and intensity fields, respectively. The solutions of the reduced-order model can be written as:

$$T^h(t, \vec{r}) = \sum_{i=1}^{K_T} a_i(t) \Psi_i^T(\vec{r}) \quad (30)$$

$$I^h(t, \vec{r}, \vec{s}) = \sum_{i=1}^{K_I} b_i(t) \Psi_i^I(\vec{r}, \vec{s}) \quad (31)$$

Substituting the above expressions into Eqs. (25) and (26), the following ordinary differential equations are obtained:

$$M_j \frac{da_j}{dt} + \sum_{i=1}^{K_T} H_{ji} a_i = -S_j + Q_j g(t), \quad j = 1 : K_T \quad (32)$$

$$\sum_{i=1}^{K_I} A_{ji} b_i - \sum_{i=1}^{K_I} B_{ji} b_i = D_j, \quad j = 1 : K_I \quad (33)$$

The related quantities in the above ODEs are defined as,

$$M_j = \rho C_p \int_V (\Psi_j^T)^2 dv \quad (34)$$

$$H_{ji} = k \int_V \nabla \Psi_j^T \cdot \nabla \Psi_i^T dv \quad (35)$$

$$S_j = \int_V \nabla \cdot \vec{q}_r \Psi_j^T dv \quad (36)$$

$$Q_j = \int_V \Psi_j^T G(x - x^*, y - y^*, z - z^*) dv \quad (37)$$

$$A_{ji} = \int_V \int_{4\pi} \{ \vec{s} \cdot \nabla \Psi_i^I \Psi_j^I + (\kappa + \sigma) \Psi_i^I \Psi_j^I \} d\Omega dv \quad (38)$$

$$B_{ji} = \int_V \int_{4\pi} \{ (\int_{4\pi} \Psi_i^I d\Omega') \Psi_j^I \} d\Omega dv \quad (39)$$

$$D_j = \int_V \int_{4\pi} (\kappa I_b - \kappa I_b^I) \Psi_j^I d\Omega dv \quad (40)$$

Solving Eqs. (32) and (33), the reduced-order solution can be obtained as follows:

$$T = T^I + \sum_{i=1}^{K_T} a_i \Psi_i^T \quad (41)$$

$$I = I^I + \sum_{i=1}^{K_I} b_i \Psi_i^I \quad (42)$$

It is seen that the total degrees of freedom are reduced to $K_T + K_I$, which is extremely small compared to the degrees of freedom of the full-order model simulation. Using this reduced-order solver of the direct analysis, we are now ready to investigate the inverse problem of interest.

6 Numerical examples

A numerical example is presented in this section to demonstrate the developed methodologies. The schematic of the problem is shown in Fig. 2. The boundary conditions associated with Eqs. (1) and (2) are the following:

$$T = 800 \text{ K}, \quad \text{on } x = 0, 1, \quad y = 0, 1, \quad z = 0, 1 \quad (43)$$

$$I(\vec{r}, \vec{s}) = \epsilon I_b + \frac{1 - \epsilon}{\pi} \int_{\vec{n} \cdot \vec{s}' < 0} |\vec{n} \cdot \vec{s}'| I(\vec{r}, \vec{s}') d\Omega', \quad \vec{n} \cdot \vec{s} > 0$$

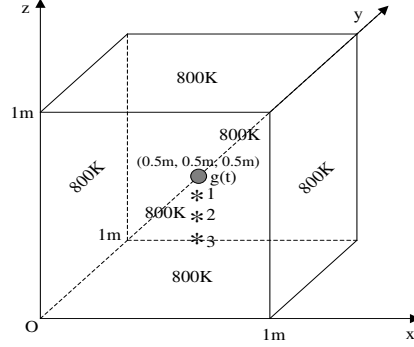


Figure 2: Schematic of the numerical example.

$$\text{on } x = 0, 1, \quad y = 0, 1, \quad z = 0, 1 \quad (44)$$

Three thermocouples are mounted at the locations 1 – (0.5, 0.5, 0.45), 2 – (0.5, 0.5, 0.4) and 3 – (0.5, 0.5, 0.35), as seen in Fig. 2. The heat source is located at (0.5, 0.5, 0.5). The spatial distribution of the heat source is approximated as follows:

$$G(x - x^*, y - y^*, z - z^*) = \exp\left\{-\frac{1}{0.05^2}(x - 0.5)^2(y - 0.5)^2(z - 0.5)^2\right\} \quad (45)$$

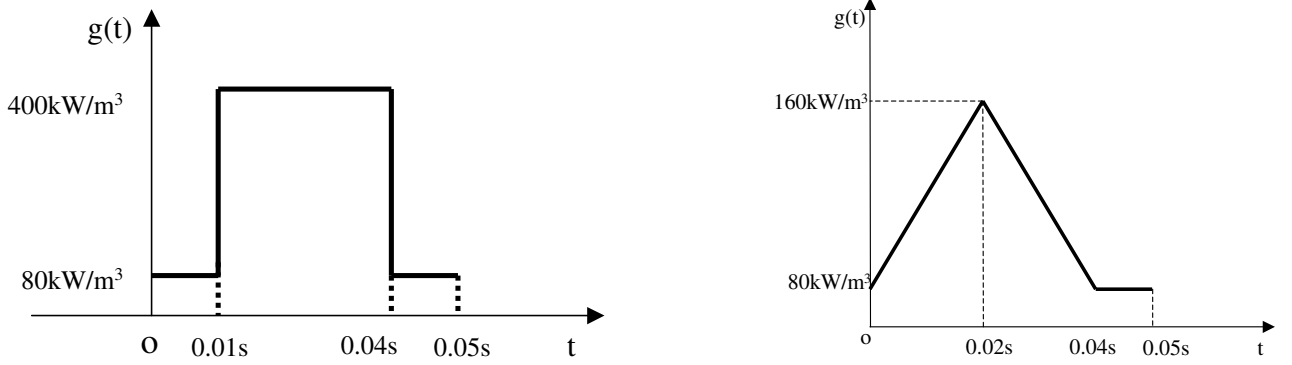


Figure 3: (a) Profile of the step heat source (b) Profile of the triangular heat source.

The material properties are taken as follows: $\rho = 0.4 \text{ kg/m}^3$, $C_p = 1100 \text{ J/kg} \cdot \text{K}$, $k = 44 \text{ W/m} \cdot \text{K}$, $\kappa = 0.5$, $\sigma = 0.5$ and $\epsilon = 0.5$. The steady-state solution obtained when $g(t) = 80 \text{ kW/m}^3$ and

$$G(x - x^*, y - y^*, z - z^*) = \exp\left\{-\frac{1}{0.25^2}(x - 0.5)^2(y - 0.5)^2(z - 0.5)^2\right\} \quad (46)$$

is taken as the initial condition.

With the above specified conditions and a step heat source strength as shown in Fig. 3a, the full-order direct model is first solved on a $26 \times 26 \times 26$ grid from $t = 0$ to $t = 0.05\text{s}$ at 100 time steps. All 100 temperature and intensity fields are recorded as snapshots to obtain the reduced order basis. Eigenfunctions corresponding to the first 6 largest eigenvalues are used in the reduced-order model. Fig. 4 shows the 1st, 3rd and 6th eigenfunctions of I^h on $y = 0.5$ along

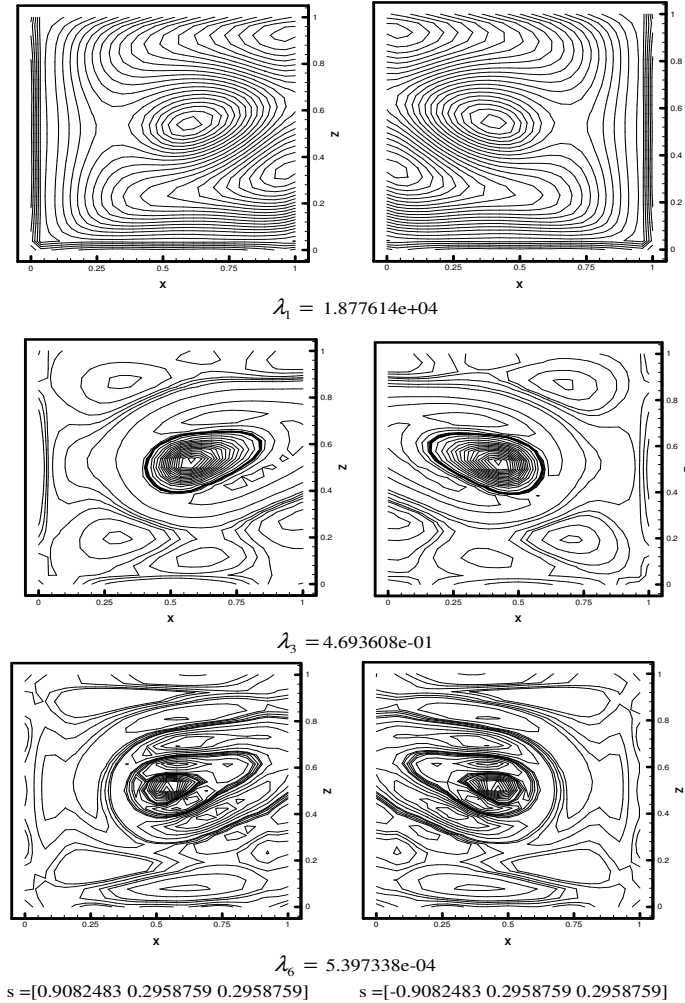


Figure 4: Eigenfunctions of I^h on $y = 0.5$ along directions $[0.9082483 \ 0.2958759 \ 0.2958759]$ and $[-0.9082483 \ 0.2958759 \ 0.2958759]$

the specified directions. The 1st, 3rd and 6th eigenfunctions of T^h on $y = 0.5$ are plotted in Fig. 5. To verify the accuracy of the POD method, temperatures at the thermocouple locations computed by both full-order and reduced-order model simulations are plotted in Fig. 6. It is obvious that the two obtained solutions are almost indistinguishable.

Simulation data are generated by adding i.i.d. zero-mean random Gaussian noise to the full-order direct model solution at the thermocouple locations. For all studied cases, the temperature is assumed to be measured from $t = 0$ to $t = 0.05$ s with a sampling interval $\Delta t = 0.001$ s, hence, there are totally 150 measurements for each case. 26 basis functions are used in the discretization of $\hat{g}(t)$ with equal step size of $dt = 0.002$ s.

To obtain a good starting point of the modified MH sampler, an initialization step is first conducted by running the sampling algorithm while solely increasing the likelihood (i.e. instead of using the MCMC acceptance criterion, we calculate the likelihood of each sample – if the likelihood is increased, we kept it, otherwise, we disregarded it). A couple hundred runs of this procedure is enough to provide a good initial guess of θ .

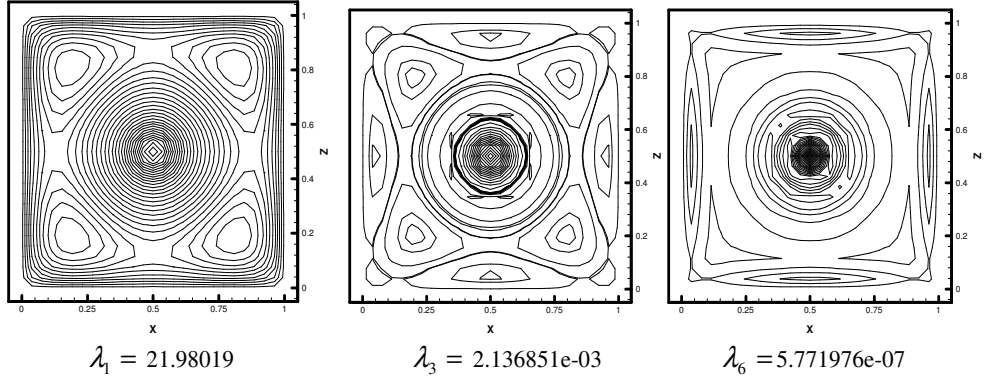


Figure 5: Eigenfunctions of T^h on $y = 0.5$.

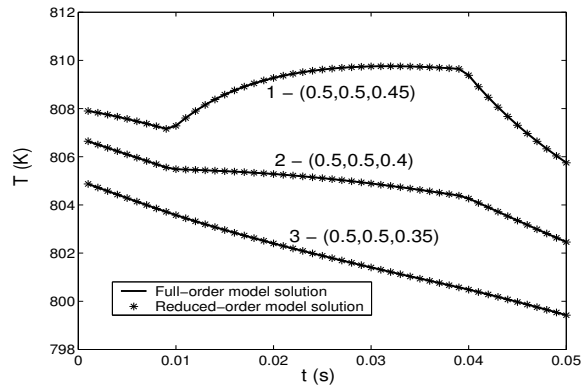


Figure 6: Temperature evolution at thermocouple locations for step heat source.

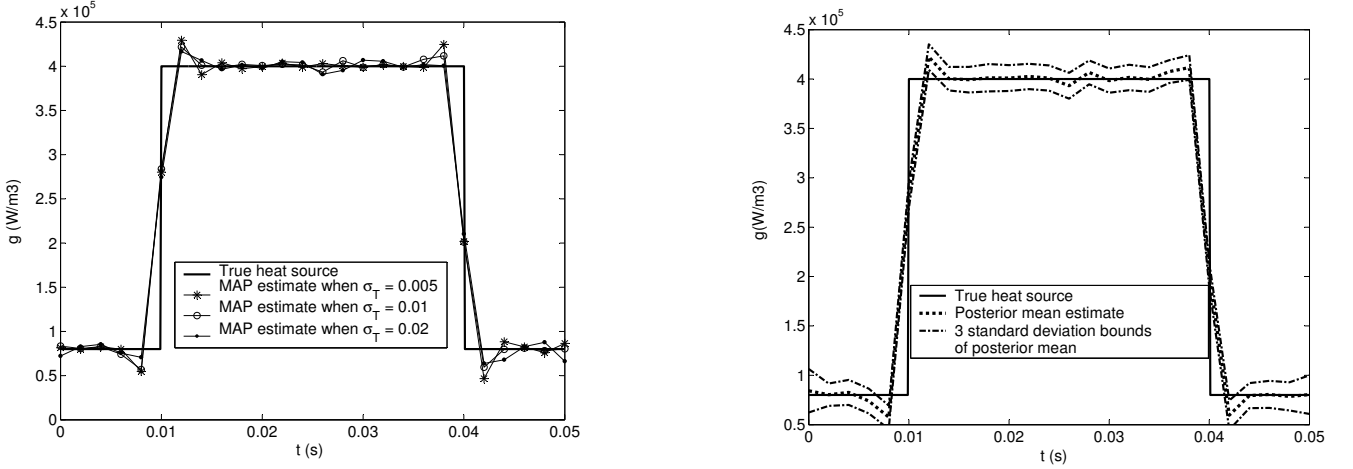


Figure 7: (a) MAP estimates for the step heat source (b) Posterior mean estimate of the step heat source and probability bounds of the posterior distribution when $\sigma_T = 0.01$.

Figure 7a shows the MAP estimates of the step heat source using MCMC sampler when σ_T has different values. It is seen that the MAP estimates are stable to the magnitudes of errors. In Fig. 7b, the posterior mean estimate when $\sigma_T = 0.01$ is plotted. The estimates are achieved using 10000 converged MCMC samples. The upper and lower bounds plotted in the same figure are the values at 3 standard deviations from the sample mean, which is an indicator of the highest density region of the posterior state space. The σ_{qj} used in the proposal distribution is 1% of the magnitude of $\theta_j^{(i)}$. This is to guarantee that the proposal distribution can fully explore the posterior state space while concentrating on the highest density region. The regularization constant, λ is chosen to be $8.0e-9$, $5.0e-9$ and $2.0e-9$, respectively for the above three cases by using the heuristic method described in [20] (selecting the range of regularization parameter within which the computed point estimate remains practically unchanged). The overall acceptance ratio for the chain used in Fig. 7b is around 77.5%.

A triangular profile of heat source as shown in Fig. 3b is also reconstructed following the same procedures. Figure 8a plots the MAP estimates of triangular heat source when σ_T has different values. The values of λ used here are $5.0e-7$, $1.0e-7$, $5.0e-8$ corresponding to values of σ_T 0.005, 0.01, 0.02, respectively. It is again seen that the estimates are relatively stable to the change of magnitude of noise. Figure 8b plots the posterior mean estimate when $\sigma_T = 0.01$. The same proposal distribution as in the previous cases is used for this run. The overall acceptance of the Markov chain is around 77.4%. It is seen that with simulated noise, the posterior mean estimate approximates the true heat flux quite well.

7 Discussion and conclusion

An inverse radiation problem is solved using a Bayesian statistical inference method. The posterior distribution of an unknown heat source is computed from temperature measurements by modeling the sensor error as Gauss random noise. A modified Metropolis-Hastings sampler is used to explore the posterior state space along which the POD method to reduce the computational cost. A Markov random field model is used to regularize the ill-posed inverse problem. The

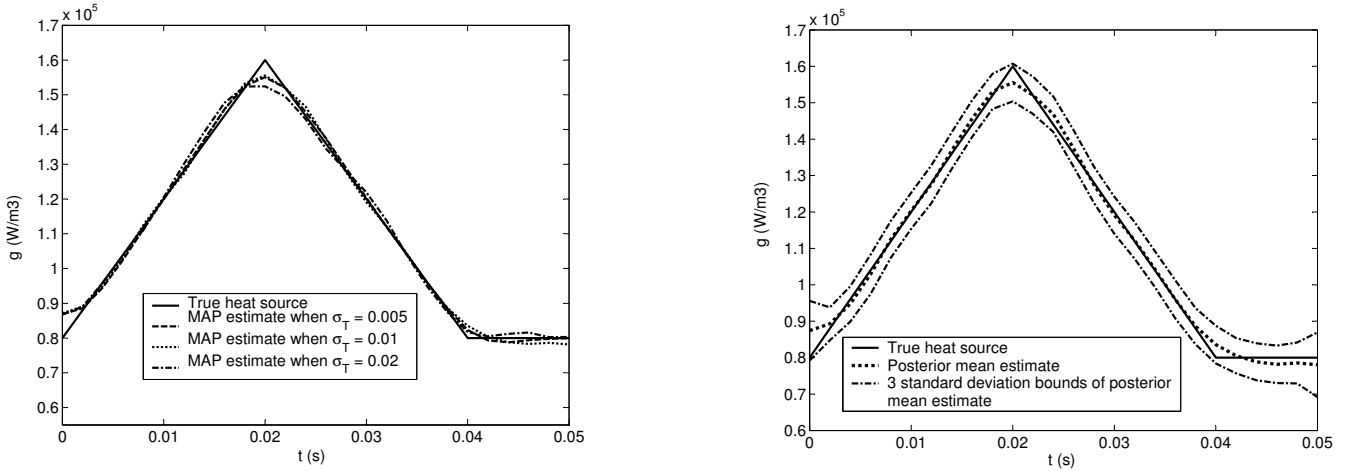


Figure 8: (a) MAP estimates for the triangular heat source case (b) Posterior mean estimate of the triangular heat source and associated probability bounds when $\sigma = 0.01$.

simulation results indicate that the method can provide accurate point estimate of the unknown heat source as well as the complete statistical information. The presented methodology provides an example of the use of Bayesian inference and spatial statistics for the solution of stochastic optimization problems and uncertainty quantification in complex continuum systems.

Acknowledgement

This work was partially supported by NASA, Office of Biological and Physical Sciences Research (98-HEDS-05), the Air Force Office of Scientific Research (grants F49620-00-1-0373 and FA9550-04-1-0070) and by the National Science Foundation (grant DMI-0113295). This work was completed while the senior author was supported by the Statistical and Applied Mathematics Science Institute (SAMSI). This research was conducted using the resources of the Cornell Theory Center.

References

- [1] T.A. Zang, M.J. Hemsch, M.W. Hilburger, S.P. Kenny, J.M. Luckring, P. Maghami, S.L. Padula and W.J. Stroud, Needs and opportunities for uncertainty-based multidisciplinary design methods for aerospace vehicles, NASA/TM-2002-211462 technical report series, Langley Research Center, Hampton, Virginia, 2002.
- [2] S.F. Wojtkiewicz, M.S. Eldred, R.V. Field, Jr. A. Urbina and J.R. Red-Horse, Uncertainty quantification in large computational engineering models, AIAA paper, AIAA-2001-1455, 2001.
- [3] R.G. Hills and T.G. Trucano, Statistical validation of engineering and scientific models: Background, Technical report SAND99-1256, Sandia national laboratories, 1999.

- [4] L. Huyse, Solving problems of optimization under uncertainty as statistical decision problems, AIAA paper 2001-1519 in 42nd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference and exhibit, Seattle, WA, 2001.
- [5] C. P. Robert, The Bayesian Choice, From Decision-Theoretic Foundations to Computational Implementation, Second Edition, Springer-Verlag, New York, Inc., 2001.
- [6] J. Besag, P. Green, D. Higdon, K. Mengersen, Bayesian computation and stochastic systems, *Statistical Science* 10(1) (1995) 3-41.
- [7] J. Moler (editor), *Spatial Statistics and Computational Methods*, Springer-Verlag, New York, Inc., 2003.
- [8] J. Wang, N. Zabaras, Using Bayesian statistics in the estimation of heat sources in radiation, *International Journal of Heat and Mass transfer*, submitted for publication.
- [9] H.M. Park, M.C. Sung, Sequential solution of a three-dimensional inverse radiation problem, *Compt. Meth. Appl. Mech. Eng.* 192 (2003) 3689-3704.
- [10] H.M. Park, T.Y. Yoon, Solution of the inverse radiation problem using a conjugate gradient method, *International Journal of Heat and Mass Transfer* 43 (2000) 1767-1776.
- [11] R. Siegel, J.R. Howell, *Thermal radiation heat transfer*, third edition, Hemisphere publishing corporation, 1992.
- [12] N.J. McCormick, Inverse radiative transfer problems : a review, *Nuclear Science and Engineering* 112 (1992) 185-198.
- [13] M.F. Modest, *Radiative Heat Transfer*, McGraw-Hill, Inc., 1993.
- [14] O. M. Alifanov, *Inverse Heat Transfer Problems*, Springer-Verlag, Berlin Heidelberg, 1994.
- [15] J.V. Beck, B. Blackwell, C. R. St-Clair Jr., *Inverse Heat Conduction: Ill-posed Problems*, Wiley-Interscience, New York, 1985.
- [16] S.M.H. Sarvari, J.R. Howell, S.H. Mansouri, Inverse boundary design conduction-radiation problem in irregular two-dimensional domains, *Numerical Heat Transfer Part B - Fundamentals* 44(3) (2003) 209-224.
- [17] A. N. Tikhonov, *Solution of ill-posed problems*, Halsted Press, Washington, 1977.
- [18] 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.
- [19] C. Ferrero, K. Gallagher, Stochastic thermal history modelling. 1. Constraining heat flow histories and their uncertainty, *Marine and Petroleum Geology* 19 (2002) 633-648.
- [20] J.B. Wang, N. Zabaras, A Bayesian inference approach to the stochastic inverse heat conduction problem, *International Journal of Heat and Mass Transfer*, accepted for publication.
- [21] A.F. Emery, Stochastic regularization for thermal problems with uncertain parameters, *Inverse Problems in Engineering* 9 (2001) 109-125.

- [22] 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, September 2002, Statistical and Applied Mathematical Sciences Institute at Duke University.
- [23] C. Andrieu, N. DE Freitas, A. Doucet, M. I. Gordan, An introduction to MCMC for machine learning, *Machine Learning* 50 (2003) 5-43.
- [24] P. Holmes, J.L. Lumley, G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, 1998.
- [25] S.S. Ravindran, A reduced-order approach for optimal control of fluids using proper orthogonal decomposition, *International Journal for Numerical Methods in Fluids* 34 (2000) 425-448.
- [26] H.V. Ly, H.T. Tran, Modeling and control of physical processes using proper orthogonal decomposition, *Mathematical and Computer Modeling* 33 (2001) 223-236.
- [27] A.N. Brooks, T.J.R. Hughes, Streamline-upwind/Petrov-Galerkin formulation for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equation, *Comput. Methods Appl. Mech. Eng.* 32 (1982) 199-259.
- [28] 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 (3) (2002).
- [29] P. Brémaud, *Markov chains, Gibbs Fields, Monte Carlo Simulation, and Queues*, Springer-Verlag, New York, 1999.
- [30] I. Beichl, F. Sullivan, The Metropolis algorithm, *Computing in Science and Engineering* 2(1) (2000) 65-69.