Using Bayesian statistics in the estimation of heat source in radiation

Jingbo Wang and Nicholas Zabaras[1]
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, USA

Abstract

An unknown transient heat source in a three-dimensional participating medium is reconstructed from temperature measurements using a Bayesian inference method. The heat source is modeled as a stochastic process. The joint posterior probability density function (PPDF) of heat source values at consecutive time points is computed using the Bayes' formula. The errors in thermocouple readings are modeled as independent identically distributed (i.i.d.) Gauss random variables. 'Maximum A Posteriori' (MAP) and posterior mean estimates of the heat source are then computed using a Markov chain Monte Carlo (MCMC) simulation method. The designed MCMC sampler is composed of a cycle of symmetric MCMC kernels. To improve the sampling speed, a model-reduction technique is used in the direct computation of temperatures at thermocouple locations given a guessed heat source, i.e. in the likelihood computation. Two typical heat source profiles are reconstructed using simulated data to demonstrate the presented methodologies. 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

Study of thermal radiation 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 [1].

In participating media, radiation is accompanied by heat conduction and convection. To simulate such processes, a coupled system of partial differential equations (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 the temperature, the radiation intensity varies in space and in direction, and the radiation intensity equation is an integro-differential equation [2]. 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 [2]. The inverse radiation problem in a participating medium that is of interest here is defined as the identification of the heat source given temperature measurements within the domain. Distinctly different from the well-posed direct problem, this inverse problem is in general ill-posed, i.e., its solution may not be unique and/or may be unstable to small errors in the given data [3, 4]. Special techniques are thus required to compute solutions to such inverse problems.

The usual solution approaches restate the inverse problem as a least-squares minimization problem [5, 6]. The objective function is formulated by minimizing the error between the computed temperatures with guessed inverse solution (in this work, a heat source) and the temperature measurements at given thermocouple locations. The error can be defined using various norms in either finite- or infinite-dimensional spaces [7, 8]. Gradient optimization techniques are introduced, and appropriate continuum or discrete sensitivity and/or adjoint problems are required [9, 10]. Other methods, such as Monte Carlo method, have also been developed for solving inverse radiation problems [11]. For review of inverse techniques for heat transfer problems, one can consult Alifanov [12] and Beck et al. [13]. The ill-posedness of these inverse problems can be addressed using appropriate regularization techniques including Tikhonov regularization [14, 15], the function specification method by Beck et al. [13], Zabaras and Liu [16] or the iterative regularization technique by Alifanov [12].

A new stochastic outlook to inverse thermal problems has recently been introduced using spectral stochastic methods [17] and Bayesian inference [18]. Stochastic inverse methods can account for uncertainties and are able to provide point estimates to the inverse solution with probability bounds [18]. In this work, we emphasize the use of Bayesian statistical inference [18, 19]. In Bayesian inference, a prior distribution model is combined with the likelihood to formulate the posterior probability density function (PPDF) [20, 21]. A Bayesian inference approach provides a complete probabilistic description of the unknown quantities given all related observations. The method regularizes the ill-posed inverse problem through prior distribution modeling (Emery [22]) and in addition provides means to estimate the statistics of uncertainties.

With the recent propagation of Markov chain Monte Carlo (MCMC) simulation methods [23], the application of Bayesian inference to engineering inverse problems becomes tractable. MCMC provides large sample data set drawn from the PPDF. These samples can be used to approximate the expectation of any function of the heat source. Running a Markov chain usually involves repetitive solution of the direct problem, which is not feasible for most nonlinear transient problems. In such situations, reduced-order models are needed [24, 25]. One widely used approach of model-reduction is the computation of the proper orthogonal decomposition (POD) basis using the method of snapshots [26, 27].

In this work, a Bayesian inference method is used to identify the strength of a transient heat source in participating media in three-dimensions (3D) through temperature measurements. A MCMC sampler is designed to explore the posterior state space. The kernel of the MCMC sampler is composed of a cycle of symmetric MCMC kernels.

Nomenclature

A	acceptance probability of MCMC	w	direction weight in S_4 method
C_p	thermal capacity	W	covariance matrix of MRF
E	expectation		test function in Galerkin formulation
F	direct simulation solver	$\tilde{W}$	test function in SUPG formulation
g	heat source	Y	temperature measurement vector
$\hat{g}$	estimate of heat source		
$G(\cdot)$	spatial approximation of point	*Greek symbols*	
	heat source	$\delta(\cdot)$	Dirac delta function
h	linear finite element basis function	ϵ	emissivity
I	radiation intensity	θ	parameter form of unknown heat source
I^h	homogeneous part of I	$\hat{\theta}$	estimate of θ
I^I	inhomogeneous part of I	κ	absorption coefficient
I_b	black body radiation intensity	λ	scaling constant of Gauss MRF
k	thermal conductivity	μ	eigenvalue in POD expansion
L	number of MCMC samples	ρ	mass density
m	dimension of θ	σ	scattering coefficient
M	number of thermocouples	σ_b	Stefan-Boltzmann constant
n	total number of measurements	σ_q	standard deviation of proposal distribution
$\vec{n}$	unit normal to the boundary vector	σ_T	standard deviation of ω
N	number of measurement steps	ω	measurement noise
N_e	number of snapshots	Ω	solid angle
$p(\cdot)$	probability density function	Φ	kernel function of MRF
$\vec{q}_r$	radiative thermal flux	Ψ	eigenfunction of POD expansion
$\vec{r}$	position vector		
$\vec{s}$	direction vector		
S	surface of 3D domain	*Superscripts*	
t	time	(i)	i^{th} iteration or i^{th} time step
$\hat{t}$	time of measurement	T	transpose
δt	thermocouple sampling interval	$*$	candidate
dt	time interval in the discretization of g		
Δt	time step size in direct simulation		
T	temperature	*Subscripts*	
T_h	homogeneous part of T	i	i^{th} component
T_I	inhomogeneous part of T	$i \sim j$	site neighborhood
u	random number	max	maximum
U	standard uniform distribution	post	posterior mean
$U^{(i)}$	i^{th} snapshot	MAP	maximum a posteriori
V	3D domain		

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. Section 3 briefly describes the full- and reduced-order finite element models used for the direct analysis. The formulation of the likelihood is presented in Section 4 together with the prior distribution model and the PPDF under a Bayesian inference framework. The design of the MCMC sampler is discussed in Section 5 including the exploration of the posterior state space. 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 and some related issues.

2 Heat source reconstruction in 3D participating media

In many high-temperature applications such as industrial combustion chambers and nuclear reactors, the strength of the heat source cannot be determined explicitly. The development of inverse techniques, however, makes it possible to reconstruct the heat source through temperature measurements at a few locations within the domain. In this work, the situation where thermal conduction and radiation occur simultaneously in participating media with diffusively reflecting boundaries is considered. The schematic of the problem of interest is given in Fig. 1. Inside the 3D domain V, heat conduction occurs simultaneously with absorption, scattering and emission of the electromagnetic waves. On the boundary surface S, the temperature is known and the electromagnetic waves

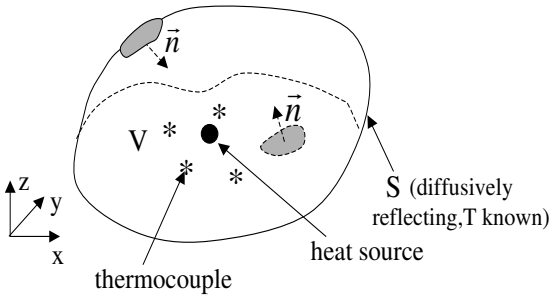


Figure 1: Schematic of the inverse radiation problem. The objective is to compute the point heat source $g(t)$ given initial conditions, boundary conditions on the surface and temperature measurements at a number of points within the domain.

are diffusively reflected. The transient heat source will be estimated through temperature measurements at sensor (thermocouple) sites within the domain. 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 governed 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 boundary 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 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. $Y = [Y_1^{(1)}, Y_2^{(1)}, ..., Y_M^{(1)}, Y_1^{(2)}, Y_2^{(2)}, ..., Y_M^{(2)}, ..., ..., Y_1^{(N)}, Y_2^{(N)}, ..., Y_M^{(N)}]^T$, where

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

where $i = 1,\ldots,M$, $j = 1,\ldots,N$ 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 with this optimal source estimate can match Y in some sense. For instance, most deterministic approaches will solve for $\hat{g}(t)$ by minimizing the least-squares error between Y and the computed temperatures.

3 Direct simulation and reduced-order modeling

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. It is seen that Eq. (1) is a nonlinear partial differential equation (PDE) and Eq. (2) has an integral term. They are coupled by the expressions in Eqs. (3) and (4). The iterative process at each time step to solve the coupled Eqs. (1) and (2) is summarized next:

1. Set $T_{\text{guess}}^{(i)} = T^{(i-1)}$;
2. Substitute $T_{\text{guess}}^{(i)}$ 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 $T_{\text{guess}}^{(i)}$ with the solution;
6. If the solutions converged, set $T_{\text{guess}}^{(i)}$ 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 field) and $T_{\text{guess}}^{(i)}$ is the guessed temperature solution. In each iteration of the above procedure, the integro-differential Eq. (2) is solved using the S_4 method [2]. In this approach, the intensity I at each spatial point is discretized into 24 directions. The integration over solid angles (directions) is approximated as weighted sum in these 24 directions. The direction vectors and associated weights are specified in [2]. 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:

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

where w_j is the weight associated with the j^{th} direction. For any given temperature field, 24 equations as Eq. (8)

with fixed direction vectors, $\vec{s}_i$'s, need to be solved iteratively to obtain I. It is noticed that Eq. (8) contains an advection term $\vec{s}_i \cdot \nabla I_i$, hence the streamline-upwind/Petrov-Galerkin (SUPG) formulation [28] is used to derive stabilized FEM equations. In summary, the weak formulations of temperature Eq. (1) and intensity Eq. (8) can be written as follows:

$$\int_V \rho C_p T^{(i)} W dv + \Delta t \int_V k \nabla T^{(i)} \cdot \nabla W dv =$$

$$\Delta t \int_V (-\nabla \cdot \vec{q}_r + g(t) G(x - x^*, y - y^*, z - z^*)) W dv +$$

$$\int_V \rho C_p T^{(i-1)} W dv, \tag{10}$$

and

$$\int_V \vec{s}_i \cdot \nabla I_i \tilde{W} dv + \int_V (\kappa + \sigma) I_i \tilde{W} dv =$$

$$\int_V \kappa I_b \tilde{W} dv + \int_V \frac{\sigma}{4\pi} \sum_{j=1}^{24} I_j w_j \tilde{W} dv, \tag{11}$$

where W and $\tilde{W}$ are the test (basis) functions for classical Galerkin and SUPG formulations, respectively.

Using the above direct simulation framework, the total number of degrees-of-freedom for the system becomes $N_n^3 \times 25$, 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 the above full-order direct model solver will be computationally intensive. To solve the stochastic inverse problem, a large number of direct simulations is required. Therefore, reduced-order modeling needs to be introduced for the direct simulation.

For the convenience of implementation, the direct problem is separated into a homogeneous part and an inhomogeneous part, i.e. $T = T^I + T^h$ and $I = I^I + I^h$. These fields are defined as follows:

For the inhomogeneous fields T^I and I^h:

$$k \nabla^2 T^I = 0 \tag{12}$$

$$\vec{s} \cdot \nabla I^I + (\kappa + \sigma) I^I - \frac{\sigma}{4\pi} \int_{4\pi} I^I(\vec{r}, \vec{s}^{'}) d\Omega^{'} = \kappa I_b^I \tag{13}$$

$$I_b^I = \frac{\sigma_b (T^I)^4}{\pi} \tag{14}$$

$$I^I = \epsilon I_b^I + \frac{1 - \epsilon}{\pi} \int_{\vec{n} \cdot \vec{s}^{'} < 0} |\vec{n} \cdot \vec{s}^{'}| I^I(\vec{r}, \vec{s}^{'}) d\Omega^{'}, \quad \vec{n} \cdot \vec{s} > 0 \tag{15}$$

$$T^I = T_w, \qquad \text{on } S \tag{16}$$

For the homogeneous fields T^h and I^h:

$$\rho C_p \frac{\partial T^h}{\partial t} = k \nabla^2 T^h - \nabla \cdot \vec{q}_r + g(t) G(x - x^*, y - y^*, z - z^*) \tag{17}$$

$$\vec{s} \cdot \nabla I^h + (\kappa + \sigma) I^h - \frac{\sigma}{4\pi} \int_{4\pi} I^h(\vec{r}, \vec{s}^{'}) d\Omega^{'} = \kappa I_b - \kappa I_b^I \tag{18}$$

$$I^h = \frac{1 - \epsilon}{\pi} \int_{\vec{n} \cdot \vec{s}^{'} < 0} |\vec{n} \cdot \vec{s}^{'}| I^h(\vec{r}, \vec{s}^{'}) d\Omega^{'}, \quad \vec{n} \cdot \vec{s} > 0 \tag{19}$$

$$T^h = 0, \qquad \text{on } S \tag{20}$$

The reduced-order models are constructed for homogeneous T^h and I^h only since the steady state Eqs. (12)-(16) 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 each basis function are computed from the solution of ordinary differential equations (ODEs) derived by Galerkin projection. The basis functions can be extracted from computational or experimental snapshots available in a database through solving the following eigenvalue problem [26]:

$$\frac{1}{N_e} \sum_{i=1}^{N_e} \int_V U^{(i)} U^{(i)}(\vec{r}^{'}) \Psi(\vec{r}^{'}) dv^{'} = \mu \Psi \tag{21}$$

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

- Take an ensemble set $\{U^{(1)}, U^{(2)}, ..., U^{(N_e)}\}$, 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 a $N_e \times N_e$ 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 $\{\Psi_1, \Psi_2, \ldots, \Psi_{N_e}\}$ is orthonormal [26]. Note that the intensity I^h is a function of both space and orientation, therefore, the volume integration in Eq. (21) and the followed eigenvalue analysis should be replaced with $\int_V \int_{4\pi} dv d\Omega$ for model reduction of I^h. Finally note that the beauty of the POD-based model-reduction is that in most situations, it is sufficient to take only a small number of basis functions (those corresponding to the larger eigenvalues). Convergence and optimality properties of POD expansions can be found in [25].

Let $\{\Psi_1^T, \Psi_2^T, ..., \Psi_{K_T}^T\}$ denote the basis functions of T^h and $\{\Psi_1^I, \Psi_2^I, ..., \Psi_{K_I}^I\}$ denote the basis functions of I^h, where K_T and K_I are the number of basis functions used for expanding temperature and intensity fields, respectively. The solutions of the reduced-order model are written as follows:

$$T^h(t, \vec{r}) = \sum_{i=1}^{K_T} a_i(t) \Psi_i^T(\vec{r}) \tag{22}$$

$$I^h(t, \vec{r}, \vec{s}) = \sum_{i=1}^{K_I} b_i(t) \Psi_i^I(\vec{r}, \vec{s}) \tag{23}$$

Substituting the above expressions into Eqs. (17) and (18), the following ODEs are obtained:

$$M_j \frac{da_j}{dt} + \sum_{i=1}^{K_T} H_{ji} a_i = -S_j + Q_j g(t), \quad j = 1 : K_T \quad (24)$$

$$\sum_{i=1}^{K_I} A_{ji} b_i - \sum_{i=1}^{K_I} B_{ji} b_i = D_j, \quad j = 1 : K_I \quad (25)$$

where the following definitions have been introduced:

$$M_j = \rho C_p \int_V (\Psi_j^T)^2 dv \quad (26)$$

$$H_{ji} = k \int_V \nabla \Psi_j^T \cdot \nabla \Psi_i^T dv \quad (27)$$

$$S_j = \int_V (\nabla \cdot \vec{q_r}) \Psi_j^T dv \quad (28)$$

$$Q_j = \int_V \Psi_j^T G(x - x^*, y - y^*, z - z^*) dv \quad (29)$$

$$A_{ji} = \int_V \int_{4\pi} \{(\vec{s} \cdot \nabla \Psi_i^I) \Psi_j^I + (\kappa + \sigma) \Psi_i^I \Psi_j^I\} d\Omega dv \quad (30)$$

$$B_{ji} = \int_V \int_{4\pi} \{(\int_{4\pi} \Psi_i^I d\Omega^{'}) \Psi_j^I\} d\Omega dv \quad (31)$$

$$D_j = \int_V \int_{4\pi} (\kappa I_b - \kappa I_b^I) \Psi_j^I d\Omega dv \quad (32)$$

Solving Eqs. (24) and (25), the reduced-order solution can be obtained as follows:

$$T = T^I + \sum_{i=1}^{K_T} a_i \Psi_i^T \quad (33)$$

$$I = I^I + \sum_{i=1}^{K_I} b_i \Psi_i^I \quad (34)$$

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

4 Bayesian inverse formulation

From a Bayesian point of view, the inverse solution is not solely a point estimate $\hat{g}$ but the probability density function of $\hat{g}$ given the observation Y. To introduce the Bayesian formulation, the unknown heat source function is first discretized using linear finite element basis functions in time as follows:

$$\hat{g}(t) = \sum_{i=1}^{m} h_i(t) \theta_i \quad (35)$$

where h_i's are as shown in Fig. 2, 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

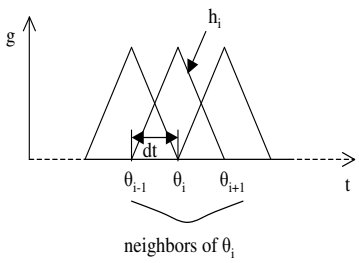


Figure 2: Basis functions and neighbor sites in the discretization of $\hat{g}$.

the estimation of the joint distribution of a 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 the Bayes's formula as:

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{p(Y)} \quad (36)$$

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

$$\hat{\theta}_{\text{MAP}} = augmax_\theta \;\; p(\theta|Y) \quad (37)$$

and the posterior mean estimate:

$$\hat{\theta}_{\text{postmean}} = E \;\; \theta|Y \quad (38)$$

In general, the probability $p(Y)$ is not explicit 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 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,

$$p(\theta|Y) \propto p(Y|\theta)p(\theta) \quad (39)$$

The likelihood function can be obtained from the following relationship,

$$Y = F(\theta) + \omega \quad (40)$$

where F is the a numerical solver that computes the temperatures at thermocouple locations given the heat source using the reduced-order model introduced in the previous section. F_i represents the temperature at the same location and time as Y_i does. In this work, we regard measurement errors (ω) as independent identically distributed (i.i.d.) Gauss random variables with zero mean and standard deviation (std) σ_T. It is assumed that the numerical errors are much less in magnitude than measurement errors. Subsequently, the likelihood can be written as,

$$p(Y|\theta) = \frac{1}{(2\pi)^{n/2}\sigma_T^n} \exp\{-\frac{(Y - F(\theta))^T (Y - F(\theta))}{2\sigma_T^2}\} \quad (41)$$

The prior distribution reflects the knowledge, if there is any, of the heat source, 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 [18]. In the current study, a specific form of Markov random fields (MRF) [29] is adopted for the prior modeling of θ. In general, the MRF can be mathematically expressed as follows:

$$p(\theta) \propto \exp\{-\sum_{i\sim j} W_{ij}\Phi(\gamma(\theta_i - \theta_j))\} \tag{42}$$

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 as shown in Fig. 2, and $W'_{ij}s$ are specified non-zero weights [21]. Let $\Phi(u) = \frac{1}{2}u^2$, the MRF can then be rewritten as:

$$p(\theta) \propto \lambda^{m/2} \exp\{-\frac{1}{2}\lambda\theta^T W\theta\} \tag{43}$$

In the one-parameter model of Eq. (43), 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. λ is a scaling constant. This MRF model is equivalent to Tikhonov regularization provided the measurement errors are Gaussian and the objective is to maximize the posterior probability (MAP) [18].

With the specified likelihood function in Eq. (41) and prior distribution in Eq. (43), the PPDF for the inverse problem can then be formulated as,

$$p(\theta|Y) \propto \exp\{-\frac{1}{2\sigma_T^2}[F(\theta) - Y]^T[F(\theta) - Y]\}$$

$$\cdot \exp\{-\frac{1}{2}\lambda\theta^T W\theta\} \tag{44}$$

In the above formulation, all the normalizing constants are neglected because the numerical algorithm introduced in later section allows to explore the posterior state space without knowing these constants. Eq. (44) is the Bayesian formulation investigated for the inverse radiation problem of interest. Both point estimates of MAP (Eq. (37)) and posterior mean (Eq. (38)) and probability bounds of the posterior distributions are computed based on this formulation.

5 MCMC sampler

For point estimates like MAP, deterministic optimization algorithms such as the conjugate gradient method can be used to find the approximate solutions. However, for obtaining the posterior mean estimate, or for estimating higher order statistics of the random unknown, statistical sampling algorithms such as Markov chain Monte Carlo (MCMC) simulation must be introduced to explore the posterior state space.

The idea of general Monte Carlo simulation is to approximate the expectation or higher order statistics of any function $f(\theta)$ by the sample mean and sample statistics from a large set of i.i.d. samples $\{\theta^{(i)}, i = 1 : L\}$ drawn from a target distribution $p(\theta)$ (PPDF in the current example), where L is the size of the sample set. Then by the strong law of large numbers, the following convergence holds:

$$E_L f(\theta) = \frac{1}{L}\sum_{i=1}^{L} f(\theta^{(i)}) \longmapsto_{L\to\infty} Ef(\theta) = \int f(\theta)p(\theta)d\theta \tag{45}$$

Obviously, the posterior mean estimate of Eq. (44) can be obtained through the above approximation. The MAP estimate can be approximated as:

$$\hat{\theta}_{MAP} = argmax_{\theta^{(i)}} \;\; p(\theta^{(i)}) \tag{46}$$

For Eq. (44), the key step in Monte Carlo simulation is to draw the sample set from this high dimensional and implicit distribution function. MCMC provides such sampling strategy using the Markov chain mechanism [23, 30]. Only the basic form of MCMC, the Metropolis-Hastings (MH) algorithm [31], 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 $u < A(\theta^{(*)}, \theta^{(i)}) = \min\{1, \frac{p(\theta^{(*)})q(\theta^{(i)}|\theta^{(*)})}{p(\theta^{(i)})q(\theta^{(*)}|\theta^{(i)})}\}$

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

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

In the above algorithm, Nmcmc is the total number of runs, u is a random number generated from standard uniform distribution $U(0, 1)$, $p(\theta)$ is the target distribution (PPDF here) and $q(*|i)$ is a proposal distribution that has standard form and generates candidate sample conditional on the previous sample. By its design, the algorithm guarantees that the samples will converge to the target distribution for any proposal distribution. However, careful design of $q(*|i)$ can accelerate convergence. Once convergence of the chain is achieved, the samples obtained can be regarded to belong to the target distribution. In principle, if the full conditional distribution of each component θ_i is available and in a standard form, it is advantageous to use the Gibbs sampler, which uses the full conditional distribution as the proposal distribution. However, this is not feasible for Eq. (44) since $F(\theta)$ is implicit.

In this study, a modified MH sampler is designed which takes advantage of the idea of Gibbs sampler, namely, to update the vector θ one component at each time. The following notation is introduced:

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

in which, the superscript (i) 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, $A(\theta_j^{(*)}, \theta_j^{(i)}) = \min\{1, \frac{p(\theta_j^{(*)}|\theta_{-j}^{(i+1)})q(\theta_j^{(i)}|\theta_j^{(*)}, \theta_{-j}^{(i+1)})}{p(\theta_j^{(i)}|\theta_{-j}^{(i+1)})q(\theta_j^{(*)}|\theta_j^{(i)}, \theta_{-j}^{(i+1)})}\}$ and

$$q_j(\theta_j^{(*)}|\theta_{-j}^{(i+1)}, \theta_j^{(i)}) = \frac{1}{\sqrt{2\pi}\sigma_{qj}} \exp\{-\frac{1}{2\sigma_{qj}^2}(\theta_j^{(*)} - \theta_j^{(i)})^2\} \tag{47}$$

with σ_{qj} is the std of the j^{th} proposal distribution. The reason for updating a single component of θ at each MCMC step is to improve the acceptance probability. In fact, by updating the entire vector at the same time, it is rather difficult to get the candidate accepted. This sampler is essentially a cycle of m symmetric MCMC samplers [23].

Since each run of above MH step requires a direct computation of the transient temperature field, it is now clear that model-reduction is essential.

6 Numerical examples

A numerical example is presented in this section to demonstrate the developed methodologies. The example considered is similar to that discussed in Park and Sung [24] but with different spatial approximation of the point heat source and with a reduced number of thermocouples. The schematic of the problem is shown in Fig. 3. The boundary conditions associated with Eqs. (1) and (2) are the following:

$$T = 800\ K, \qquad \text{on}\ \ x = 0, 1, \quad y = 0, 1, \quad z = 0, 1 \tag{48}$$

$$I(\vec{r}, \vec{s}) = \epsilon I_b + \frac{1-\epsilon}{\pi} \int_{\vec{n}\cdot\vec{s}'<0} |\vec{n} \cdot \vec{s}'| I(\vec{r}, \vec{s}') d\Omega', \qquad \vec{n} \cdot \vec{s} > 0$$
$$\text{on}\ \ x = 0, 1, \quad y = 0, 1, \quad z = 0, 1 \tag{49}$$

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

$$G(x - x^*, y - y^*, z - z^*) = exp\{-\frac{1}{0.05^2}$$
$$(x - 0.5)^2(y - 0.5)^2(z - 0.5)^2\} \tag{50}$$

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

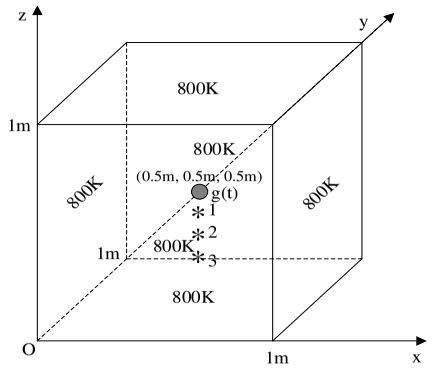


Figure 3: Schematic of the numerical example.

$$G(x - x^*, y - y^*, z - z^*) = exp\{-\frac{1}{0.25^2}$$
$$(x - 0.5)^2(y - 0.5)^2(z - 0.5)^2\} \tag{51}$$

is taken as the initial condition.

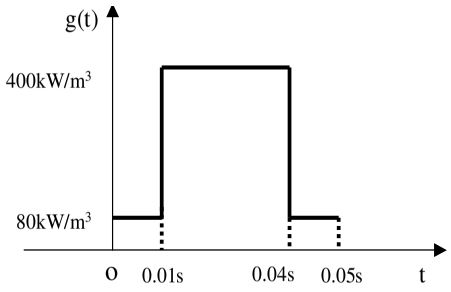
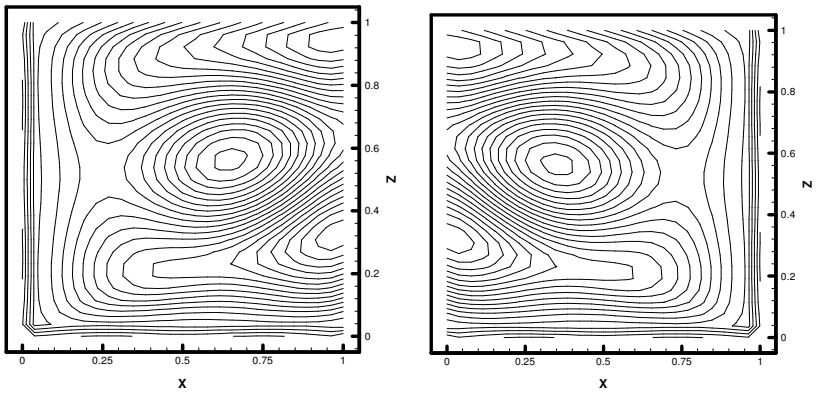


Figure 4: Profile of the step heat source.

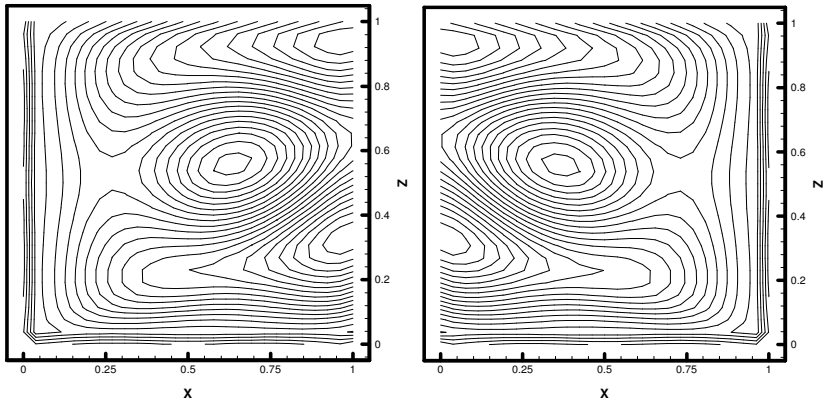
With the above specified conditions and a step heat source profile of $g(t)$ as shown in Fig. 4, the full-order direct model is first solved on a $26 \times 26 \times 26$ grid from $t = 0$ to $t = 0.05$ s at 100 time steps. Fig. 5 shows the computed homogeneous radiation intensities on cross section $y = 0.5$ at different times along the specified directions. The homogeneous temperature fields on the same cross section at different times are plotted in Fig. 6.

All 100 temperature and intensity fields are recorded as snapshots to obtain the eigenfunctions. Eigenfunctions corresponding to the first 6 largest eigenvalues are used in the reduced-order model. Fig. 7 shows the 1st, 3rd and 6th eigenfunctions of I^h on $y = 0.5$ along the specified direction. The 1st, 3rd and 6th eigenfunctions of T^h on $y = 0.5$ are plotted in Fig. 8. To verify the accuracy of the POD method, the temperature fields on $y = 0.5$ obtained by solving the reduced-order model with a heat source as in Fig. 4 are given in Fig. 9. Fig. 10 shows the evolution of the temperature at the thermocouple locations computed by both full-order and reduced-order model simulations. It is obvious that the two solutions are almost indistinguishable.

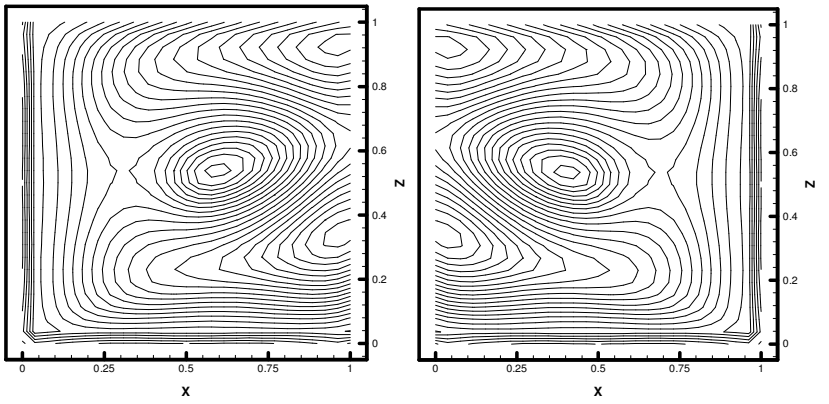
To demonstrate the Bayesian method for inverse reconstruction of the heat source profile of Fig. 4, simulation



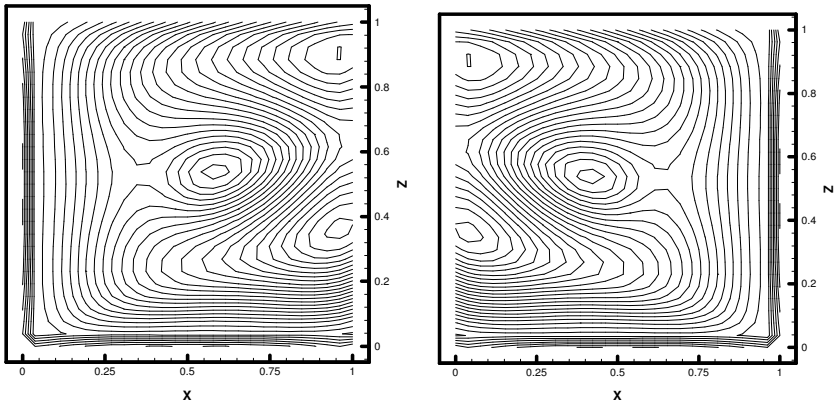
(a) t = 0.005s



(b) t = 0.01s



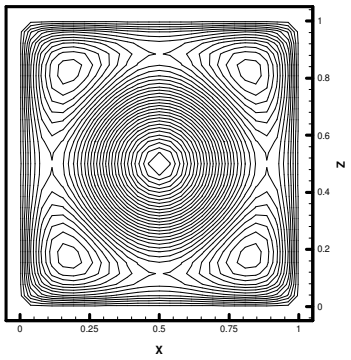
(c) t = 0.025s



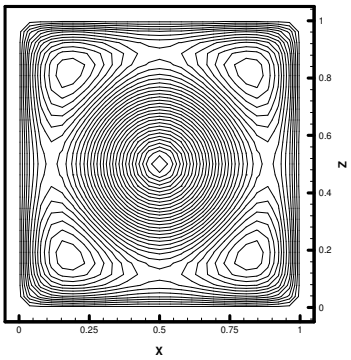
(d) t = 0.05s

s =[0.9082483 0.2958759 0.2958759] s =[-0.9082483 0.2958759 0.2958759]

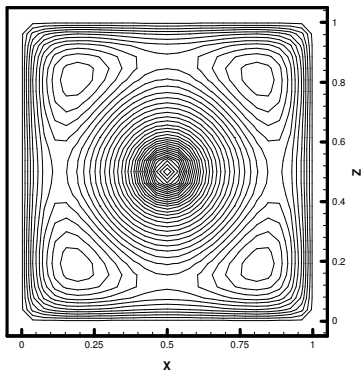
Figure 5: Homogeneous intensity fields on $y = 0.5$ along directions [0.9082483 0.2958759 0.2958759] and [−0.9082483 0.2958759 0.2958759] for step heat source.



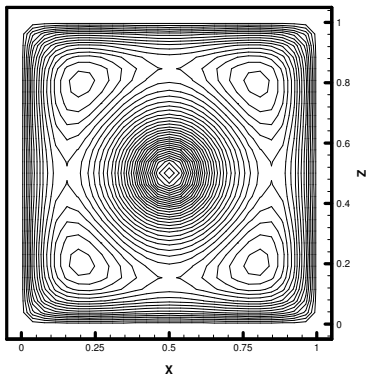
(a) t = 0.005s



(b) t = 0.01s

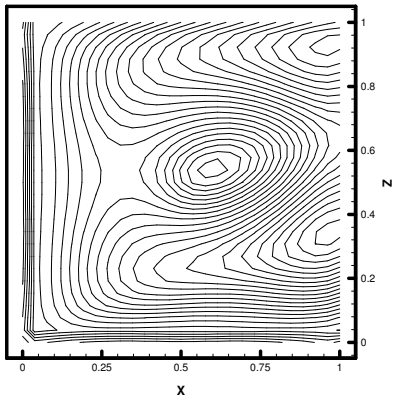


(c) t = 0.025s

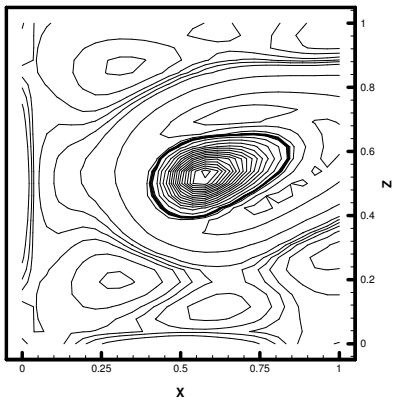


(d) t = 0.05s

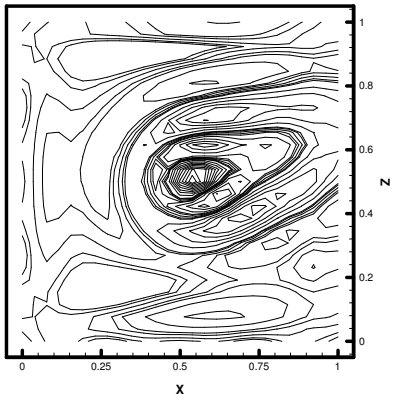
Figure 6: Homogeneous temperature fields on $y = 0.5$ for step heat source.



$\lambda_1 =$ 1.877614e+04



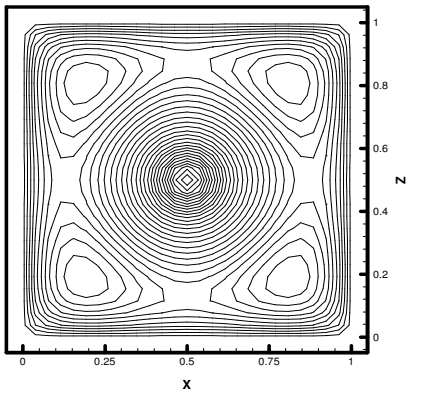
$\lambda_3 =$ 4.693608e-01



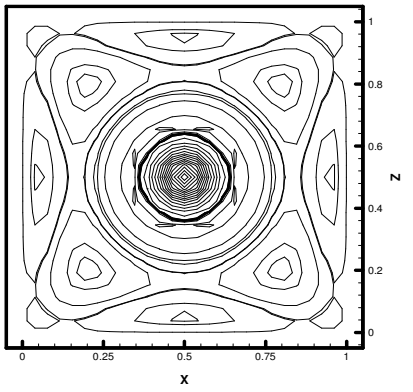
$\lambda_6 =$ 5.397338e-04

s =[0.9082483 0.2958759 0.2958759]

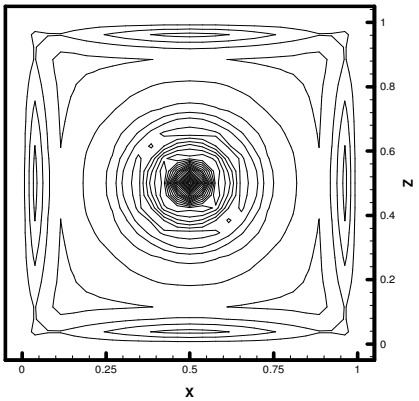
Figure 7: Eigenfunctions of I^h on $y = 0.5$ along direction [0.9082483 0.2958759 0.2958759].



$\lambda_1 = 21.98019$

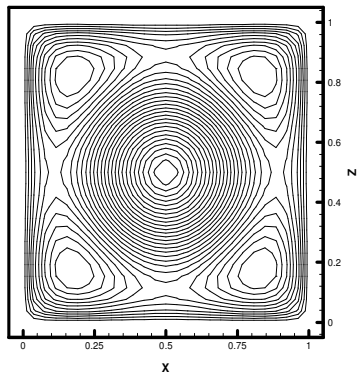


$\lambda_3 = 2.136851\text{e-}03$

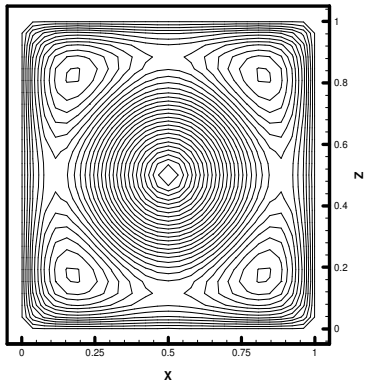


$\lambda_6 = 5.771976\text{e-}07$

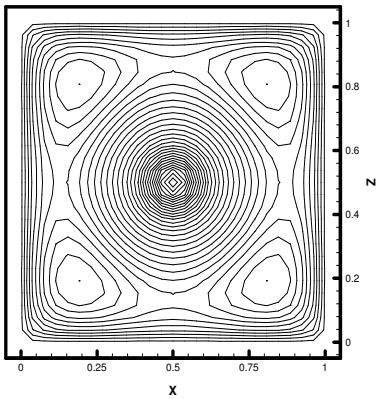
Figure 8: Eigenfunctions of T^h on $y = 0.5$.



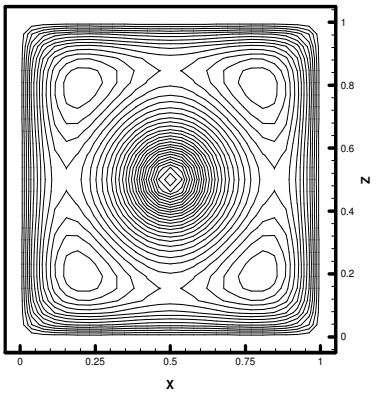
t = 0.005s



t = 0.01s



t = 0.025s



t = 0.05s

Figure 9: Homogeneous temperature field computed using the POD method on $y = 0.5$ for step heat source.

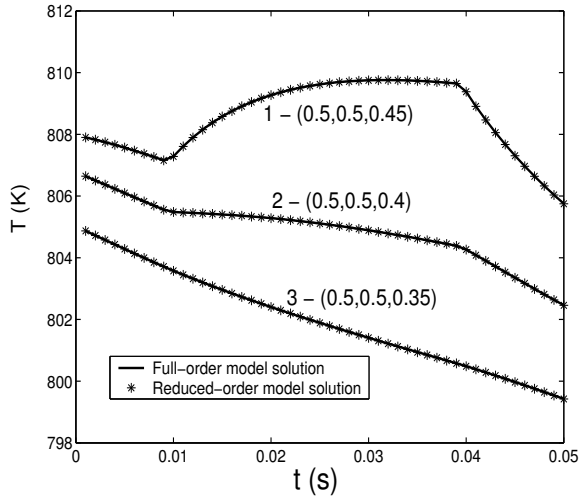


Figure 10: Temperature evolution at thermocouple locations for step heat source.

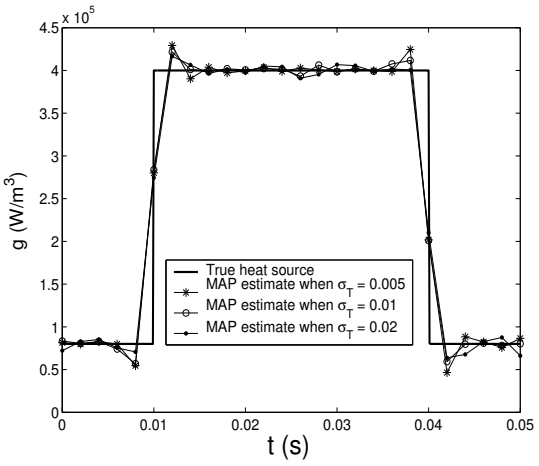


Figure 11: MAP estimates for the step heat source.

data are generated by adding Gauss random noise with zero mean and standard deviation σ_T to the full-order direct model solution at the thermocouple locations. For all following cases, the temperature is assumed to be measured from $t = 0$ to $t = 0.05$s with a sampling interval $\delta t = 0.001s$, 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 $dt = 0.002s$.

To obtain a good starting point for the MH sampling, an initialization step is first conducted by running the sampling algorithm while solely increasing the likelihood. A few hundred runs of this procedure is enough to provide a good initial guess of θ.

Fig. 11 plots the MAP estimates of the step heat source using MCMC samples when σ_T has different values. It is seen that the MAP estimates are stable to various magnitudes of errors. In Fig. 12, 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 indication 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 method described in [18] (selecting the range of regularization parameter within which the computed point estimate remains practically unchanged). The regularization parameter λ can be treated as a hyperparameter in a hierarchical augmented Bayesian formulation thus avoiding any need for its priori selection. This approach, however, was not followed here to limit the discussion to the fundamental aspects of Bayesian inference. The overall acceptance ratio for the chain used in Fig. 12 is around 77.5%.

A triangular profile of heat source as shown in Fig. 13 is

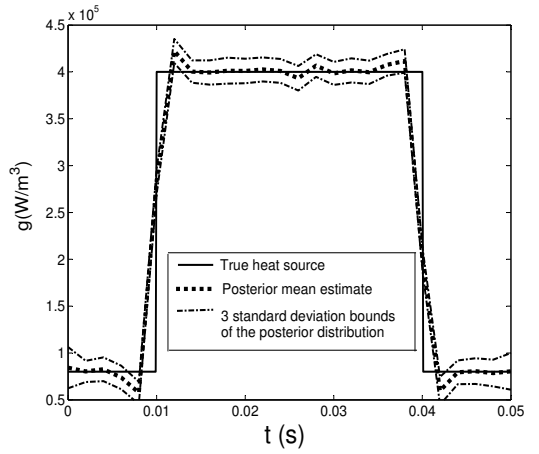


Figure 12: Posterior mean estimate of the step heat source and probability bounds of the posterior distribution when $\sigma_T = 0.01$.

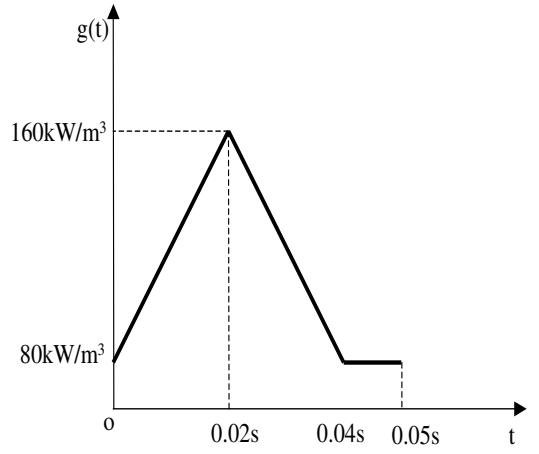


Figure 13: Profile of the triangular heat source.

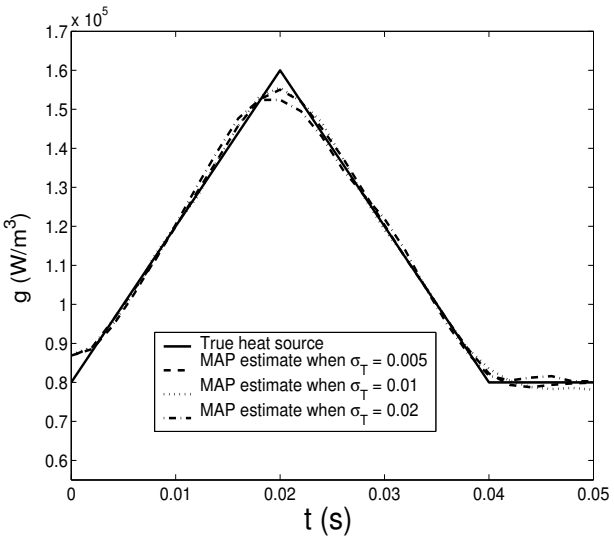


Figure 14: MAP estimates for the triangular heat source case.

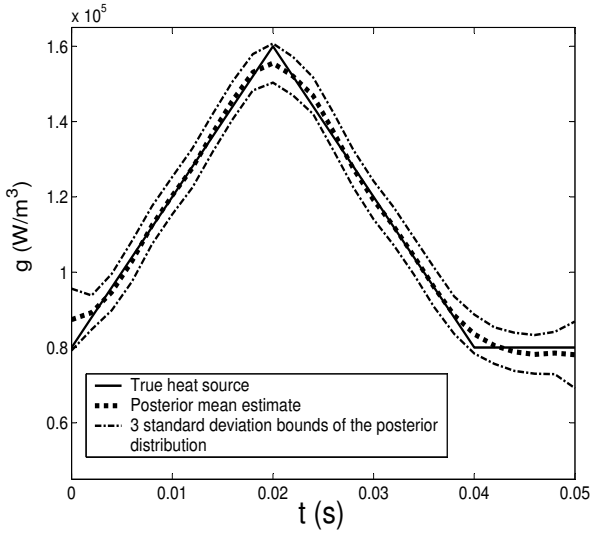


Figure 15: Posterior mean estimate of the triangular heat source and probability bounds of the posterior distribution when $\sigma = 0.01$.

also reconstructed following the same procedures. Fig. 14 plots the MAP estimates of triangular heat source when σ_T has different values. It is again seen that the estimates are relatively stable to the change of magnitude of noise. Fig. 15 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 strength is computed from temperature measurements by modeling the measurement errors as i.i.d. Gauss random variables. The Metropolis-Hastings algorithm was used to explore the posterior state space and the POD method to reduce the computational cost. A Markov random fields model was used to regularize the ill-posed inverse problem. The simulation results indicate that the method can provide accurate point estimates of the unknown heat source as well as complete statistical information. Although the study is devoted toward point heat source estimation, the methodologies can be extended to reconstruction of distributed heat sources as well by using multiscale Markov random fields models in the prior distribution modeling, where the inherent length scales in temporal and spatial directions are explored. Finally, in the situation where thermal properties are dependent on the temperature and large temperature variation is observed, the Bayesian computation is still applicable. However, a general methodology for proper selection of snapshots in the POD modeling that can capture the dynamics of the temperature and intensity fields needs to be developed.

Acknowledgements

Our preliminary work on Bayesian techniques was partially supported from the Advanced Mechanical Technologies Program at General Electric Global Research Center (GE-GRC). Additional support was provided by NASA, Office of Biological and Physical Sciences Research (NAG8-1671), the Air Force Office of Scientific Research (grant FA9550-04-1-0070) and by the National Science Foundation (grant DMI-0113295). This research was conducted using the computing resources of the Cornell Theory Center.

References

[1] R. Siegel, J.R. Howell, Thermal radiation heat transfer, third edition, Hemisphere publishing corporation, 1992.

[2] M.F. Modest, Radiative Heat Transfer, McGraw-Hill, Inc., 1993.

[3] C.H. Ho, M.N. Özisik, An inverse radiation problem, International Journal of Heat and Mass Transfer, 32(1989) 335-341.

[4] N.J. McCormick, Inverse radiative transfer problems : a review, Nuclear Science and Engineering, 112 (1992) 185-198.

[5] L.H. Liu, H.P. Tan, Inverse radiation problem in three-dimensional complicated geometric systems with opaque boundaries, Journal of Quantitative Spectroscopy and Radiative Transfer, 68 (2001) 559-573.

[6] C.E. Siewert, An inverse source problem in radiative transfer, Journal of Quantitative Spectroscopy and Radiative Transfer, 50 (1993) 603-609.

[7] T. Viik, N.J. McCormick, Numerical test of an inverse polarized radiative transfer algorithm, Journal

of Quantitative Spectroscopy and Radiative Transfer, 78 (2003) 235-241.

[8] M. Prud'homme, S. Jasmin, Determination of a heat source in porous medium with convective mass diffusion by an inverse method, International Journal of Heat and Mass Transfer 46 (2003) 2065-2075.

[9] H.M. Park, T.Y. Yoon, Solution of the inverse radiaiton problem using a conjugate gradient method, International Journal of Heat and Mass Transfer 43 (2000) 1767-1776.

[10] 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.

[11] S. Subramaniam, M.P. Mengüc, Solution of the inverse radiation problem for inhomogeneous and anisotropically scattering media using a Monte Carlo technique, International Journal of Heat and Mass Transfer 34 (1991) 253-266.

[12] O. M. Alifanov, Inverse Heat Transfer Problems, Springer-Verlag, Berlin Heidelberg, 1994.

[13] J.V. Beck, B. Blackwell, C. R. St-Clair Jr., Inverse Heat Conduction: Ill-posed Problems, Wiley-Interscience, New York, 1985.

[14] A. N. Tikhonov, Solution of ill-posed problems, Halsted Press, Washington, 1977.

[15] H.E., Ofodike, A. Ezekoye, J.R. Howell, Comparison of three regularized solution techniques in a three-dimensional inverse radiation problem, Journal of Quantitative Spectroscopy and Radiative Transfer, 73 (2002) 307-316.

[16] 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.

[17] V.A. Badri Narayanan, N. Zabaras, Stochastic inverse heat conduction using a spectral approach, International Journal for Numerical Methods in Engineering, 60(7) (2004).

[18] J.B. Wang, N. Zabaras, A Bayesian inference approach to the stochastic inverse heat conduction problem, International Journal of Heat and Mass Transfer, 47 (2004).

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

[21] J. Besag, P. Green, D. Higdon, K. Mengersen, Bayesian computation and stochastic systems, Statistical Science 10(1) (1995) 3-41.

[22] A.F. Emery, Stochastic regularization for thermal problems with uncertain parameters, Inverse Problems in Engineering 9 (2001) 109-125.

[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] H.M. Park, M.C. Sung, Sequential solution of a three-dimensional inverse radiation problem, Compt. Meth. Appl. Mech. Eng. 192 (2003) 3689-3704.

[25] P. Holmes, J.L. Lumley, G. Berkooz, Turbulence, Coherent Structures, Dynamical Systems and Symmetry, Cambridge University Press, 1998.

[26] 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.

[27] H.V. Ly, H.T. Tran, Modeling and control of physical processes using proper orthogomal decomposition, Mathematical and Computer Modeling 33 (2001) 223-236.

[28] 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.

[29] 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) 230-241.

[30] P. Brémaud, Markov chains, Gibbs Fields, Monte Carlo Simulation, and Queues, Springer-Verlag, New York, 1999.

[31] I. Beichl, F. Sullivan, The Metropolis algorithm, Computing in Science and Engineering 2(l) (2000) 65-69.