PERGAMON

International Journal of Heat and Mass Transfer 43 (2000) 3331–3339

International Journal of HEAT and MASS TRANSFER

www.elsevier.com/locate/ijhmt

Uncertainties in parameter estimation: the optimal experiment design

A.F. Emery[a,*], Aleksey V. Nenarokomov[b], Tushar D. Fadale[c]

[a] *University of Washington, Seattle, WA 98103, USA*
[b] *Department of Space System Engineering, Moscow Aviation Institute, Moscow, Russia*
[c] *Applied Research Associates, Raleigh, NC 27609, USA*

Received 5 February 1999; received in revised form 20 November 1999

Abstract

An extended maximum likelihood principle is described by which inverse solutions for problems with uncertainties in *known* model parameters can be treated. The method introduces the concept of an equivalent experimental noise which differs significantly from the measurement noise when the system response is sensitive to the uncertainties in the *known* parameters. When the equivalent noise varies smoothly and significantly over the range of uncertainty, the inverse solution tends to be independent of the uncertainties. By minimizing the equivalent noise through appropriate choice of a measurement protocol, an optimal experiment can be defined. Examples are given of designing an experiment for estimating conductivity and contact resistance when surface convective coefficients are uncertain.

1. Introduction

The parameters of the model of a system are usually determined by using the inverse method which consists of comparing the measured response of the system to the response predicted with varying values of the sought-after parameters. The predictions are made by using the model with properties and parameters, other than the sought-after parameters, assumed to be *known*. Although existing inverse methods consider the experimental noise, they do not take into account errors or uncertainties which might exist in these presumably *known* parameters of the system. The authors have presented previously [1] a new measure of performance which takes into account such uncertainties in the *known* model parameters by extending the concept of the maximum likelihood principle.

In this paper, we continue the development of this *extended* theory to demonstrate its improved performance and its use in optimal experiment design. Examples are given to investigate the consequences of applying it to sensors which are located at optimal and non-optimal positions or to take readings at times of less than maximum information.

2. Theory

Following [1], let us consider a thermal system which can be modeled by a set of differential equations and model parameters. Of all the parameters involved in the model, let there be P unknown parameters represented by the vector $\mathbf{u}$ which are to be determined by

* Corresponding author. Tel.: +1-206-543-5338; fax: +1-206-685-8047.
E-mail address: emery@u.washington.edu (A.F. Emery).

0017-9310/00/$ - see front matter
PII: S0017-9310(99)00378-6

Nomenclature

b	known parameter vector
Cov[ ]	covariance operator
$E[\ ]$	expected value operator
f	probability density function
F	Fisher information matrix
G	covariance matrix of known parameters
h	heat transfer coefficient
I	identity matrix
$\mathscr{I}$	information
k	thermal conductivity
K	number of observation times
L	likelihood functional
M	proposed Fisher information matrix
N	number of sensors
P	number of parameters
R	contact resistance
S	covariance matrix of measurement noise
t	time
u	unknown parameter vector
$\mathbf{T}_f$	surface temperature of die
V	covariance matrix of temperature
z	measured temperature vector

Greek symbols

$\boldsymbol{\Omega}$	measure of $\mathbf{F}^{-1}$
$\boldsymbol{\Phi}$	predicted temperature vector
$\boldsymbol{\Psi}$	discrepancy between prediction and measurement vector
σ_h	standard deviation of h
σ_n	standard deviation of measurements
$\boldsymbol{\Theta}$	sensitivity matrix of the state of the system to **b**

Subscripts

i, j	sensors
k	time
l, m	parameters
s	substrate

Superscript

T	transpose

the inverse method. The rest of the parameters, say Q parameters, are known a priori and are represented by the vector **b**. In order to estimate the **u** parameters, experimental observations of the state of system are obtained at various locations (x_i, y_i) and times (t_k). Let N be the number of sensors used in the experiment and K the number of readings in time at these sensor locations. Then the experimental observations can be represented by $\mathbf{z}_k$ where,

$$\mathbf{z}_k^{\mathrm{T}} = \{z_i(t_k), i = 1, 2, \ldots, N\} \quad k = 1, 2, \ldots, K \tag{1}$$

Let $\boldsymbol{\Phi}_k$ be a vector of the predicted state of the system at all the sensor locations based on the mathematical model using *known* the parameters (for distributed parameter systems, this involves the solution of the field equations) such that,

$$\boldsymbol{\Phi}_k^{\mathrm{T}} = \{\phi_i(t_k), i = 1, 2, \ldots, N\} \quad k = 1, 2, \ldots, K \tag{2}$$

The idea is to choose the sought-after parameters, **u**, in such a way that the model predictions agree with the experimental measurements according to a specified functional [1–3].

The vector **z** in Eq. (1) results from real measurements, and therefore, always contains errors. The total error is composed of systematic (bias) and random components. The complete analysis of the sources of individual error components is a separate and complicated problem. It is important to note here that state-of-the-art methods and devices used for temperature measurement provide a rather low level of systematic error due to suitable design and manufacturing methods. In some cases, for example, when the sensor location or starting time is not known precisely, it is possible, by comparing model computations to the experimental results, to estimate the systematic error and to include the necessary correction into the measured results [4,5]. However, as pointed out by Moffat [5], such corrections invariably depend upon the judgement of the experimentalist. In the development that follows, we will restrict the analysis to cases where the temperature measurement errors contain only the random component δz.

$$\delta\mathbf{z}_k^{\mathrm{T}} = \{\delta z_i(t_k), i = 1, 2, \ldots, N\} \quad k = 1, 2, \ldots, K \tag{3}$$

where $\delta z_i = z_i - E[z_i]$.

The usual deterministic approach is to minimize the least squares difference between the prediction and the measurements [3].

$$\mathbf{L}^* = \sum_{k=1}^{K} \{\boldsymbol{\Phi}_k - \mathbf{z}_k\}^{\mathrm{T}} \{\boldsymbol{\Phi}_k - \mathbf{z}_k\} \tag{4}$$

In this approach, no information about δz, other than having a zero mean, is utilized. Because inverse problems are poorly posed, the search for the minimum value of $\mathbf{L}^*$ is complicated. A number of papers have described optimal mathematical techniques (often

referred to as regularization) for determining the minimum and criteria for terminating the search [6]. In multi-parameter problems, we terminate the search according to the residual criterion [2].

$$\mathbf{L}^*(\mathbf{u}) \leq \sum_{k=1}^{K} \delta\mathbf{z}_k^{\mathrm{T}} \delta\mathbf{z}_k \tag{5}$$

Eq. (4) can formally be treated as a regression experiment whose purpose is the estimation of $\mathbf{u}$. Since the solution of the inverse problem, as expressed by Eq. (4), is not unique, i.e. generally speaking, a minimum of $\mathbf{L}^*$ can be obtained under various realizations of the experimental conditions and measuring schemes, the problem naturally arises of improving the computational properties of the inverse problem by an optimally designed experiment. Optimal experiment design problems for systems described by partial differential equations were investigated in [7]. There only linear problems were analyzed, and the optimal design was chosen as that experiment which minimizes the variance of the estimated parameter.

3. Maximum likelihood

In solving inverse problems, it seems reasonable to use a criterion which reflects the statistical information available on the errors δz. The theory with which we shall be concerned in this paper has its origins in a paper by Chernoff [8]. During the last 30 years, it was pursued vigorously by Fedorov [7], Box and Lucas [9], Goodwin and Payne [10] and Silvey [11]. The experimental measurements of the state of the system $\mathbf{z}_k$ are considered to be random in nature and the measurement errors are assumed to be independent of each other and normally distributed with a zero mean and a variance of σ_n^2.

If the likelihood is given by a Gaussian probability distribution [10]

$$f(\boldsymbol{\Psi}|\mathbf{u}) = \left[(2\pi)^{NK} \prod_{k=1}^{K} \mathrm{Det}(\mathbf{S}_k) \right]^{-1/2} \exp\left[\sum_{k=1}^{K} -\frac{1}{2} \boldsymbol{\Psi}_k(\mathbf{u})^{\mathrm{T}} \mathbf{S}_k^{-1} \boldsymbol{\Psi}_k(\mathbf{u}) \right] \tag{6}$$

then maximizing the likelihood is equivalent to minimizing

$$\mathbf{L} = \sum_{k=1}^{K} \ln(\mathrm{Det}(\mathbf{S}_k)) + \sum_{k=1}^{K} \boldsymbol{\Psi}_k(\mathbf{u})^{\mathrm{T}} \mathbf{S}_k^{-1} \boldsymbol{\Psi}_k(\mathbf{u}) \tag{7}$$

in which $\boldsymbol{\Psi}_k = \boldsymbol{\Phi}_k - \mathbf{z}_k$ and the covariance matrix $\mathbf{S}_k$ reflects the variance of the experimental measurements. If $\mathbf{S}_k$ is considered to be constant with respect to time and diagonal, i.e, $\mathbf{S}_k = \sigma_n^2 \mathbf{I}$, $\mathbf{L} = \mathbf{L}^*$ and the maximum likelihood and least squares approaches are identical.

According to the Cramer–Rao theorem [10,12], the estimation error has a lower bound of $\mathbf{F}^{-1}$ where $\mathbf{F}$ is the $P \times P$ Fisher Information matrix. If the signal noise is not a function of the sought-after parameter, the elements of $\mathbf{F}$ can be expressed by [13]

$$(\mathbf{F})_{lm} = \sum_{k=1}^{K} \left[\left\{ \frac{\partial \boldsymbol{\Phi}_k}{\partial u_m} \right\}^{\mathrm{T}} \mathbf{S}_k^{-1} \left\{ \frac{\partial \boldsymbol{\Phi}_k}{\partial u_l} \right\} \right] \quad l, m = 1, \ldots, P \tag{8}$$

An optimal experimental protocol can be devised by choosing the experimental conditions, i.e. boundary conditions, sensor location, etc., which minimize a measure $\boldsymbol{\Omega}$ of $\mathbf{F}^{-1}$. When only one parameter is being sought, $\mathbf{F}$ reduces to a scalar. For multiple parameters, $\mathbf{F}$ is a matrix and there are a variety of different measures used to characterize the estimation error. Some of the different measures are discussed by Emery [14,15] for experiments to estimate simultaneously k and c or k and h, respectively.

If only one parameter is being sought and if the signal noise is independent of time and has zero covariance, then the lower bound for the variance of the parameter is

$$\sigma^2(u) \geq \mathbf{F}^{-1} = 1 / \sum_{k=1}^{K} \left[\frac{\{\partial \boldsymbol{\Phi}_k / \partial u\}^2}{\sigma_n^2} \right] = 1 / \sum_{k=1}^{K} \mathscr{I}_k(\mathbf{L}) \tag{9}$$

where $\mathscr{I}_k(\mathbf{L})$ represents the information obtained from each reading when all other model parameters are assumed to be known.

4. Extended maximum likelihood theory

In addition to the experimental errors, the predictions will also exhibit variations which are due to the uncertainties in the *known* parameters used in these models. Thus, the predictions $\boldsymbol{\Phi}_k$ should also be considered to be stochastic in nature. In this extended theory, the unknown parameters are estimated by minimizing the functional $\mathbf{J}$ which is given as,

$$\mathbf{J} = \sum_{k=1}^{K} \ln(\mathrm{Det}(\mathbf{V}_k)) + \sum_{k=1}^{K} \boldsymbol{\Psi}_k(\mathbf{u})^{\mathrm{T}} \mathbf{V}_k^{-1} \boldsymbol{\Psi}_k(\mathbf{u}) \tag{10}$$

The covariance matrix $\mathbf{V}_k$ is defined as,

$$\mathbf{V}_k = E\left[\{\boldsymbol{\Psi}_k - E[\boldsymbol{\Psi}]\} \{\boldsymbol{\Psi}_k - E[\boldsymbol{\Psi}_k]\}^{\mathrm{T}} \right] = \boldsymbol{\Theta}_k \mathbf{G} \boldsymbol{\Theta}_k^{\mathrm{T}} + \mathbf{S}_k \tag{11}$$

where $\mathbf{G}$ is the covariance of the uncertain parameters

and Θ is the sensitivity of the system response with respect to these uncertain parameters. The aspects of the calculation of $\mathbf{V}_k$ are detailed in [16].

It can be shown that the Information matrix, $\mathbf{M}$, for the extended theory given in Eq. (10) is

$$(\mathbf{M})_{lm} = \sum_{k=1}^{K}\left[\left\{\frac{\partial \boldsymbol{\Phi}_k}{\partial u_m}\right\}^{\mathrm{T}}\mathbf{V}_k^{-1}\left\{\frac{\partial \boldsymbol{\Phi}_k}{\partial u_l}\right\} + \frac{1}{2}\mathrm{Tr}\left[\mathbf{V}_k^{-1}\frac{\partial \mathbf{V}_k}{\partial u_l}\mathbf{V}_k^{-1}\frac{\partial \mathbf{V}_k}{\partial u_m}\right]\right] \quad (12)$$

$l, m = 1, \ldots, P$

It is important to note that since $\mathbf{M}$ involves the complicated dependence of $\mathbf{V}_k$ upon the sought-after parameters, the location of the maximum of $\mathbf{M}$ is a highly non-linear optimization problem.

In many of the cases we have studied, $\mathbf{M}$ is only weakly affected by the trace term. In this situation, a comparison of Eqs. (8) and (12) suggests that $\mathbf{V}_k$ can be considered as the *equivalent noise* of the experiment. For only one sought-after parameter, one uncertain parameter, and one sensor, we find

$$\sigma^2(u) \geq \mathbf{M}^{-1} = 1/\sum_{k=1}^{K}\left[\frac{\{\partial \phi_k/\partial u\}^2}{\sigma_n^2 + (\partial \phi_k/\partial b)^2\sigma_b^2}\right] = 1/\sum_{k=1}^{K}\mathscr{I}_k(\mathbf{J}) \quad (13)$$

where $\mathscr{I}_k(\mathbf{J})$ represents the information obtained at each reading when some of the model parameters are uncertain. Even if the signal noise is independent of time, the equivalent noise is a strong function of time through the term $\partial \phi_k/\partial b$.

The information, $\mathscr{I}(\mathbf{J})$ is usually less than $\mathscr{I}(\mathbf{L})$ because $\mathbf{V}_k \geq \mathbf{S}_k$. Attempts to increase the total amount of information by improving the measurement accuracy (i.e. reducing σ_n) are usually ineffective because the term $(\partial \phi_k/\partial b)\sigma_b$ is the dominant component of $\mathbf{V}_k$.

The algorithm for determining the optimal experiment is as follows:

1. Solve the direct problem for $\boldsymbol{\Phi}$.
2. Solve the corresponding sensitivity problem for $\boldsymbol{\Theta}$.
3. Estimate the covariance matrix $\mathbf{G}$ which defines the uncertainty of the a priori known parameters.
4. Choose an initial set of experimental conditions (sensor location, sampling times, starting times, etc.).
5. Determine the maximum value of $\mathbf{M}$, subject to constraints on the experimental conditions.

If the elements of the information matrix $\mathbf{M}$ are independent of the sought-after parameter vector, $\mathbf{u}$, the extremum problem Eq. (10) can be solved without any information on $\mathbf{u}$. This situation, however, is seldom found in thermal studies. Therefore, only the construction of locally optimal experiments is possible [7]. It is necessary to use a priori information on $\mathbf{u}$ as a starting point in an iterative process. As discussed in [15], it is difficult to frequently obtain a global minimum without careful visualization of the functional.

5. Example I

For the first example, let us estimate the conductivity of a homogenous material. This will be done by sampling the temperature of a one-dimensional slab. The slab is of thickness $L = 0.04$ m, has conductivity $k = 1.0$ W/(m °C) and volumetric heat capacity $\rho c = 1.0 \times 10^6$ J/(m^3 °C). It is initially at 0°C temperature,

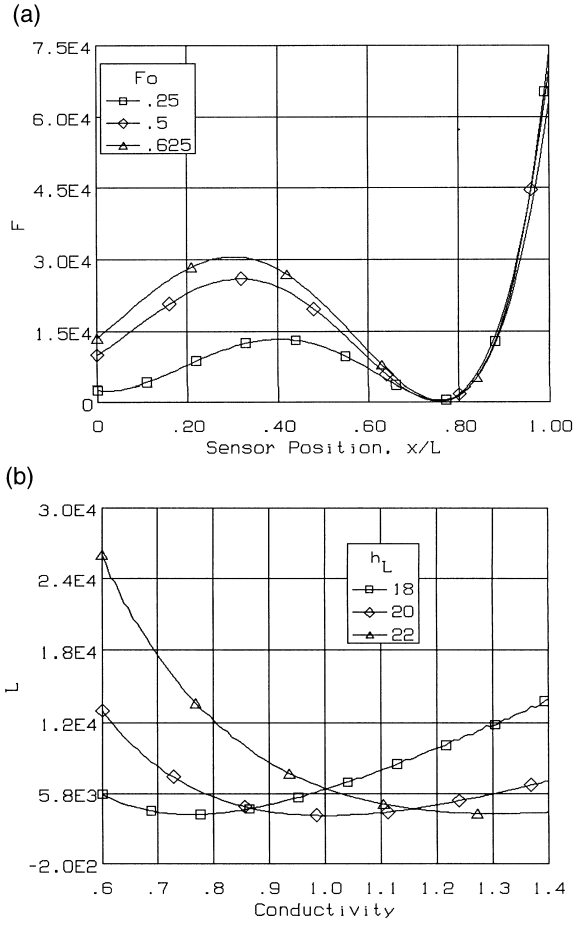

Fig. 1. (a) Values of the Fisher information matrix as a function of total experiment time (Fo) and sensor location (x/L). (b) Variation of $\mathbf{L}$ with conductivity for a sensor at $x = L$, total experiment time (Fo) of 0.625 and 0.1% noise.

has a convective heat transfer coefficient of $h_0 = 5$ W/(m^2 °C) at $x = 0$ and $h_L = 20$ W/(m^2 °C) at $x = L$, and is immersed in a fluid with a temperature 1000°C at time zero. The temperature is measured at a fixed location at 10 equally spaced times over the duration of the experiment. The experimental temperatures are taken to be the analytical temperatures based upon a conductivity of $k = 1$ W/(m °C), corrupted by a Gaussian noise with zero mean and a standard deviation, σ_n of 1°, which corresponds to 0.1% of the maximum temperature.

Let us also consider that the information available on the heat transfer coefficient at $x = L$ is not accurate and that its value is distributed about the mean value $h_L = 20$ with a standard deviation of σ_h.

The measured temperatures will be compared to the analytical temperatures which are computed with different values of conductivity, k. The measured and computed temperatures will be substituted into Eq. (7) and the values of **L** computed. That k for which **L** is a minimum is the estimated conductivity. In general, the location of the minimum when several measurement variables (e.g. sensor locations for multiple sensors) are adjustable requires fairly sophisticated search techniques or regularization. In this example where the optimal design involves only the specification of the location of one sensor, a direct line search is sufficient.

First, it is necessary to determine the optimal sensor location. Fig. 1a shows the values of **F**, the information content of the experiment, as a function of maximum experiment time and sensor location. Referring to Eq. (9), the reciprocal of **F** is a measure of the variance of the conductivity, i.e. the uncertainty in the conductivity estimated from the inverse process. Since the variation in **F** is a direct function of the sensitivity of the measured temperature to the conductivity, it is, as intuitively expected, highest at $x = L$ where the heat flow is the greatest. Surprisingly, it also has a local maximum near $x = 0.3$ where the two heat flows are opposed and tend to raise the temperature. For this experiment, the optimal sensor location is at $x = L$. The value of **F** at $x = L$ indicates that there will be an uncertainty in the estimated value of k of about $\sigma_k/k \approx 0.4\%$ or about four times the standard deviation of the average signal noise.

If one of the *known* parameters used in computing the analytical temperatures, say h_L, is uncertain, then the logical approach would be to repeat the process for different values of h_L. Fig. 1b illustrates the behavior of **L** for different values of h_L for a sensor at $x = L$.

The estimated conductivities are those which minimize **L**. From Fig. 1b, it is clear that our estimate of k is strongly sensitive to the value of h_L assumed, and it is clear that small errors in specifying the surface heat transfer coefficient will lead to large errors in the estimated conductivity.

In this case of uncertainty in the *known* parameters, it would appear reasonable to apply the *extended* theory and to use the value of **J**, Eq. (10). Fig. 2 shows the estimated conductivities found by using **J**, and it is seen that the *extended* theory does not perform any better than the original theory, **L**.

The fallacy in this approach is that the sensor location is not optimal when there is uncertainty in h_L. If the extended theory, **J**, is to be used, one must also determine the optimal position based upon **M**, not **F**. Fig. 3a shows the values of **M** as a function of sensor location for different degrees of uncertainty in h_L. Although the temperature is most sensitive to conductivity at $x = L$ (and thus the high value of **F** there), the increase in uncertainty in the estimated conductivity due to the dependence of the temperature on h_L is also the highest there. The effect is to reduce the information value of readings made there and to increase the standard deviation of the estimated k. At 2.5% uncertainty in h_L, the optimal position is in the interior. As the uncertainty in h_L increases, the optimal position moves towards $x = 0$, since the temperatures there is least affected by h_L.

Using a sensor at $x = 0$ and estimating the conductivity using **J**, gives the results shown in Fig. 3b by the curve marked. (The curve marked **L** will be discussed in the next paragraph.) We see that using the extended theory with measurements at the optimal location gives excellent results even when the value of h_L used is far from correct. Fig. 3c shows how **J** behaves with h_L. It is clear that the minimum points are coincident in terms of k and that accuracy in the value of h_L is not needed. At this point, we do not understand why the curves lose their smoothness away from the mini-

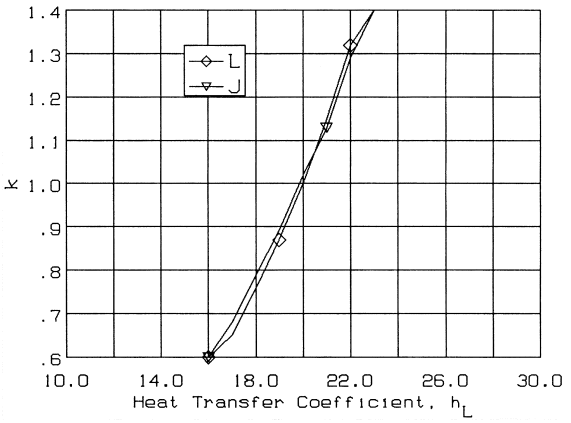

Fig. 2. Estimated conductivity using **L** and **J** for a sensor at $x = L$, for a total experiment time (Fo) of 0.625, 0.1% noise and 10% uncertainty in h_L.

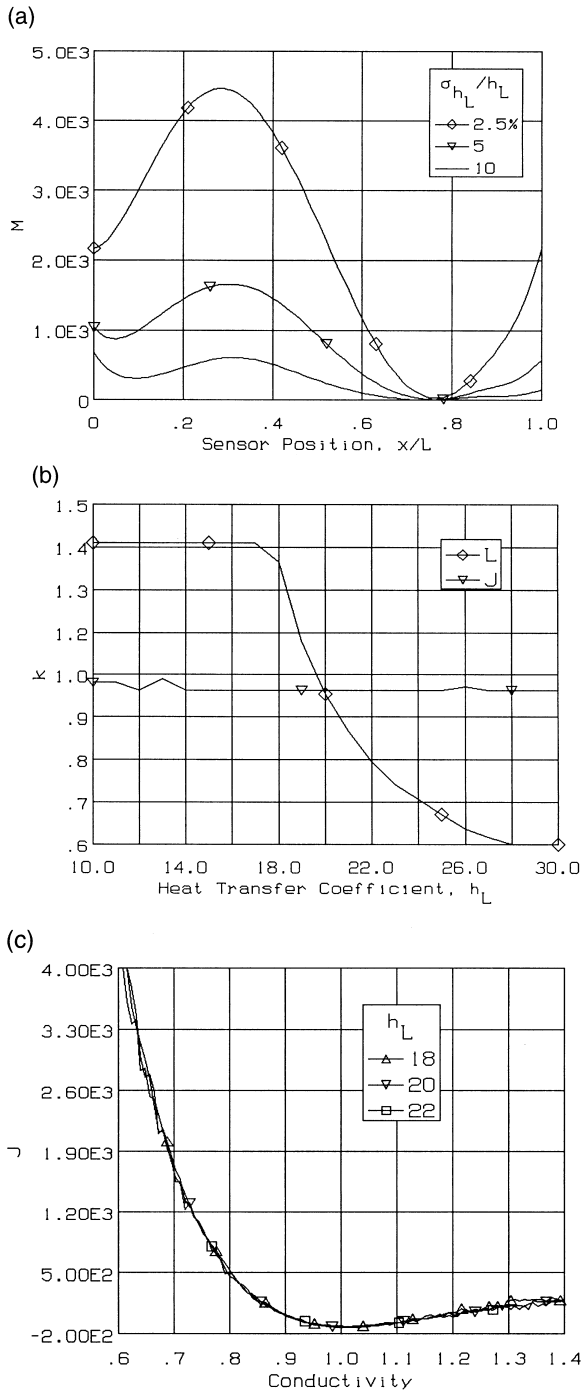

Fig. 3. (a) Values of the extended Fisher information matrix as a function of the uncertainty in h_L for a total experiment time (Fo) of 0.625 and 0.1% noise. (b) Estimated conductivity using **L** and **J** for a sensor at the optimal location, $x = 0$, for a total experiment time (Fo) of 0.625, 0.1% noise, 10% uncertainty in h_L. (c) Variation of **J** with conductivity for a sensor at the optiimal location, $x = 0$, for a total experiment time (Fo) of 0.625, 0.1% noise, and 10% uncertainty in h_L.

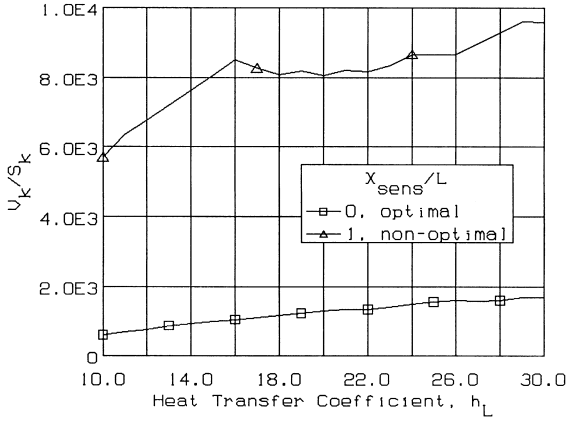

Fig. 4. Comparison of $\mathbf{V}_k$ and $\mathbf{S}_k$ for the sensor at the optimal and non-optimal locations, for a total experiment time (Fo) of 0.625, 0.1% noise, and 10% uncertainty in h_L.

mum point, but in all the cases we have studied, they are sufficiently smooth near the minimum to permit us to estimate the conductivity.

A logical question to ask when estimating k is why not simply put the sensor at a point where we think that the imprecision in h_L has minimal effect and to use the regular approach, i.e, **L**. This would save us from the computation of $\partial\phi_k/\partial b$. Such a point would be at $x = 0$, far from the surface where h_L most strongly affects the temperature. Fig. 3b shows the result of this approach. The conductivities determined from **L** still vary strongly with h_L and the correct result is obtained only when the precise value of h_L is used. The reasons for this are twofold. First, the information obtained at $x = 0$ is small as illustrated in Figs. 1 and 3a. Second, the temperature, and thus **L**, are still reasonably strong functions of h_L. Fig. 4 shows how the ratio $\mathbf{V}_k/\mathbf{S}_k$ (a direct measure of the sensitivity of temperature to h_L) varies with h_L. Since **J** incorporates **V**, the effect of the 3:1 variation of **V** over the range of h_L at the optimal location is to reduce the sensitivity of **J** to h_L.

6. Example II

Consider an encapsulated microelectronic circuit (a die) bonded to an alumina substrate with a contact resistance $\mathbf{R}_0$ between the die and the substrate, Fig. 5. The thickness of the substrate is characterized by its thermal resistance, $\mathbf{R}_s$. When powered, the die generates heat which is conducted into the substrate or conducted through the encapsulation and convected into the ambient fluid, whose temperature remains constant, with a heat transfer coefficient h_0. The measured temperatures will be taken as those computed with h_0 and $\mathbf{R}_0$, respectively, and corrupted by adding uncorrelated,

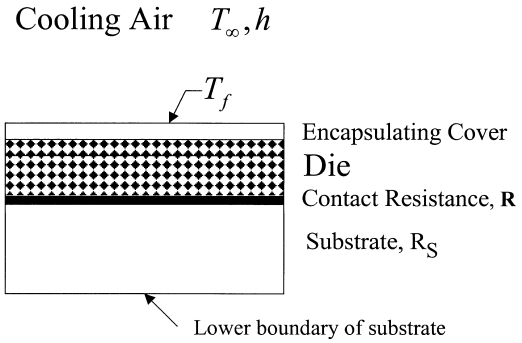

Fig. 5. Schematic of Example II showing location of contact resistance, **R**, and the substrate.

random noise, with zero mean and a standard deviation of 1% of the maximum temperature of the chip.

We will estimate **R** by comparing the simulated measured surface temperature $\mathbf{T}_f$ with those computed using different values of **R** and choosing that value which minimizes either **L** or **J**. We will examine two cases: (1) the lower surface of the substrate is insulated and (2) the lower surface is maintained at the ambient fluid temperature. The properties of the different materials are taken from [17] for an alumina ceramic substrate ($k_s = 25$, $\alpha_s = 1.7 \times 10^6$) and a highly conducting cover.

In estimating the contact resistance **R**, we assume that the convective coefficient h is not precisely known but has an uncertainty σ_h. Thus, we solve for **R** by assuming different values of h, computing $\mathbf{T}_f$ and minimizing either **L** or **J**. The results are sensitive to the ratio of $\mathbf{R}_0/\mathbf{R}_s$ as shown in Fig. 6 for two different thermal boundary conditions at the lower surface of the substrate. For an insulated substrate, we see that for $\mathbf{R}_0/\mathbf{R}_s$ equal to 0.1 and 10 the predicted value of **R** is unacceptably sensitive to our estimate of h, and if there is any uncertainty about the value of h, we will make a serious error in our estimate of **R**.

The reason for this unacceptable sensitivity can be understood by examining the relation for the surface temperature $\mathbf{T}_f$ which is given by

$$\mathbf{T}_f = Q_f/h \tag{14}$$

where Q_f is the heat transferred through the cover. For the properties used in this example, little heat is stored in the encapsulating cover so that Q_f is essentially the difference between the heat generated in the die, Q_d, and that which flows through the attachment to the substrate. When **R** is small as compared to $\mathbf{R}_s$, the heat flow into the substrate is essentially independent of **R**, and thus, Q_f is also insensitive to h. However, from Eq. (14), it is still very sensitive to h. Thus, **L** is insensitive to **R** but highly sensitive to h, and therefore, h must be known precisely to extract the correct value of **R**. When **R** is large in comparison to $\mathbf{R}_s$, essentially no heat flows into the substrate, all flowing into the ambient fluid and again there is no sensitivity to **R** and we must specify h precisely to recover the correct value of **R**. However, when **R** and $\mathbf{R}_s$ are approximately the same, then the sensitivity of Q_f to **R** is relatively high compared to the sensitivity to h and the prediction is better, but still not sufficiently independent of h to be acceptable.

When the surface of the substrate is maintained at a fixed temperature, the heat flow into the substrate is much more sensitive to **R** and the estimation of **R** becomes essentially independent of the value of h used to compute $\boldsymbol{\Phi}$.

6.1. Information analysis

It is clearly advantageous to analyze the experiment before conducting it in order to maximize the accuracy

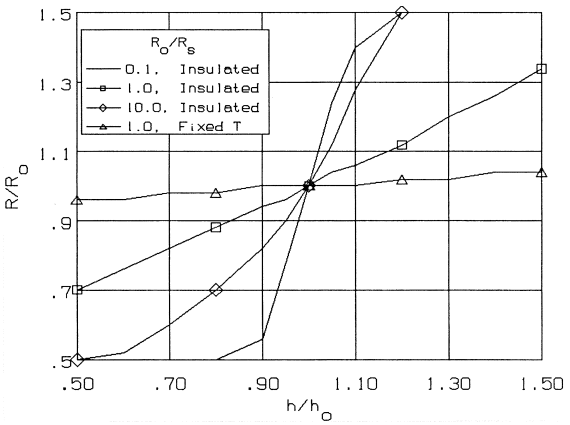

Fig. 6. Estimated values of **R** based on 20 samples of $\mathbf{T}_f$ equally spaced in time over the period $0.35 \le Fo \le 3.5$.

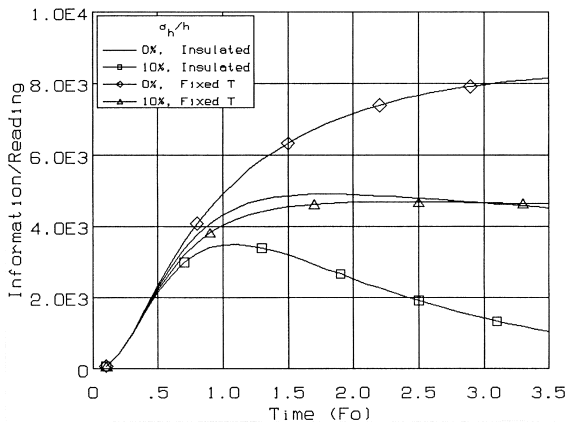

Fig. 7. Variation of the information per reading, $\mathscr{I}_k$, with respect to time.

of the predicted value of **R**. This is best done by looking at the information which we can obtain from the experiment.

Fig. 7 shows how the information varies with time for the two different boundary conditions for $\mathbf{R}_0/\mathbf{R}_s = 1$. First, we see that the information for the constant temperature boundary conditions is approximately twice as much as the maximum information obtained for the insulated substrate and remains essentially constant at long times, even if h is uncertain. On the other hand, when there is an uncertainty in h, the information decreases dramatically at longer times and readings taken shortly after the information has peaked contribute essentially no new information. Thus, the estimation error of **R** remains high.

If the value of **R** is extracted using **J** instead of **L**, then we obtain the results shown in Fig. 8. Even when readings are taken at long times, where the information is markedly reduced, the use of **J** compensates for the deleterious effect of an uncertain h and the predicted value of **R** is in error by less than 10% over a range of h/h_0 from 0.5 to 1.5.

It might appear reasonable to confine the sampling to that time interval during which the information content is a maximum. This is not possible since the method depends upon sensing the time variation of $\mathbf{T}_f$, and if the samples are too close together, the temporal variation will be overcomed by the noise. If we do restrict the sampling to a range of time for which $\mathbf{T}_f$ varies sufficiently, but one during which the information content is reasonable ($0.35 \le Fo \le 1.5$) the estimation error is reduced as indicated by the second curve labeled **L** in Fig. 8. However, the accuracy is no better than that obtained using **J** which permits more latitude in choosing the times at which the measurements are taken.

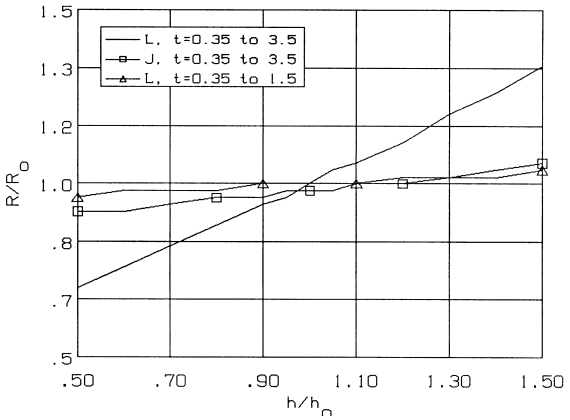

Fig. 8. Estimated values of **R** using **J** and **L** for two different sampling intervals.

7. Conclusions

The *extended* theory has been shown to be able to account successfully for uncertainties in surface heat transfer coefficients when estimating conductivity and contact resistance. Although results are presented only for variations in h_L, we have found similar results for variations in fluid temperature, sensor locations, and independent variations in h_L and h_0 and for problems in which multiple parameters are estimated. We thus conclude that this extended theory is applicable in all cases in which some of the prescribed parameters, whether they be properties or boundary conditions, are uncertain. However, the success is achieved only when the sensor is placed at the position which is defined as optimal by the corresponding extended Fisher information matrix, **M**.

The theory is useful only when $\mathbf{V}_k$ differs significantly from $\mathbf{S}_k$ and is a strong function of the sought-after parameters and the uncertain parameters; thus, it cannot correct for uncertainty when the measurement noise is high.

Since the information content of a reading is such a strong function of the uncertain parameters and the parameter sought, the first task in designing an experiment should be the determination of the information content of the measurements in terms of the parameter sought. Once this is known, then the experimentalist can decide on the specific experiment, taking into account the trade-offs between information and practicality. If there are no uncertain parameters, the process is straightforward. However, in the presence of uncertainty, it is important to make use of the extended information matrix and maximum likelihood function. Doing so gives a more accurate estimate of the effect of the uncertainty on the estimation error, but more importantly provides a method for reducing the error.

Acknowledgements

A portion of this work was done under the sponsorship of NSF grant CTS-9322215.

References

[1] T.D. Fadale, A.V. Nenarokomov, A.F. Emery, Uncertainties in parameter estimation: the inverse problem, The International Journal of Heat and Mass Transfer 38 (3) (1995) 511–518.
[2] O.M. Alifanov, Inverse Heat Transfer Problems, Springer-Verlag, Berlin/Heidelberg, 1994.
[3] H.W. Sorenson, Parameter Estimation: Principles and Problems, Marcel Dekker, New York, 1980.
[4] R.B. Abernethy, R.P. Benedict, R.B. Dowdell, ASME

measurement uncertainty, ASME J. Fluids Engineering 26 (1985) 161–164.
[5] R.J. Moffat, Describing the uncertainties in experimental results, Experimental Thermal and Fluid Science 1 (1988) 3–17.
[6] J.V. Beck, B. Blackwell, C.R. St. Clair, Inverse Conduction Ill-posed Problems, Wiley–Interscience, New York, 1985.
[7] V.V. Fedorov, Theory of Optimal Experiment, Academic Press, New York, 1972.
[8] H. Chernoff, Locally optimal design for estimating parameters, The Annals of Mathematical Statistics 24 (4) (1953) 586–602.
[9] G.E.P. Box, H.L. Lucas, Design of experiments in nonlinear situations, Biometrika 46 (1) (1959) 77–90.
[10] G.E. Goodwin, R.L. Payne, Dynamic System Identification. Experiment Design and Data Analysis, Academic Press, New York, 1977.
[11] S.D. Silvey, Optimal Design: An Introduction to the Theory for Parameter Estimation, Chapman & Hall, London, 1983.
[12] Strang, G., Introduction to Applied Mathematics, Wellesley–Cambridge, 1986.
[13] O.M. Alifanov, E.A. Artyukhin, S.V. Rumyantsev, Extreme Methods for Solving Ill-posed Problems with Applications to Inverse Problems, Begell House, Wallingford, UK, 1995.
[14] A.F. Emery, T.D. Fadale, Design of experiments using uncertainty information, J. Heat Transfer, Trans. of ASME 118 (1996) 532–538.
[15] T.D. Fadale, A.V. Nenarokomov, A.F. Emery, Two approaches to optimal sensor locations, J. Heat Transfer, Trans. of ASME 117 (5) (1995) 373–379.
[16] T.D. Fadale, A.F. Emery, Transient effects of uncertainties on the sensitivities of temperatures and heat fluxes using stochastic finite elements, J. Heat Transfer, Trans. of ASME 116 (1994) 808–814.
[17] M. Pecht, R.K. Agarwal, Electronic packaging materials and their properties, in: R.J Hanermann, A.D. Kraus, M. Pecht (Eds.), Physical Architecture of VLSI Systems, Wiley, New York, 1994.