Chapter 5
Radiative Transfer and Heat Conduction

We show in this chapter how to estimate the value of several material properties, such as single scattering albedo and thermal conductivity, present in the heating process of a material with thermal radiation.

The inverse problems considered here deal with the estimation of a finite number of parameters, related to an infinite dimensional model. These problems are solved as finite dimensional optimization problems, using the Levenberg-Marquadt method, which is a variant of Newton's method for non-linear systems of equations. In an intermediary step, this method includes a regularization, similar to Tikhonov's regularization (Section 3.4).

This chapter deals with inverse identification problems. According to Table 2.3, page 48, they are classified as Type III *inverse problems*.

5.1 Mathematical Description

Consider the example depicted in Fig. 5.1, of a material body subjected to heating due to thermal radiation. Assume we can regulate the intensity and the position of the heat sources and that we perform temperature measurements in some material body points, using temperature sensors[1].

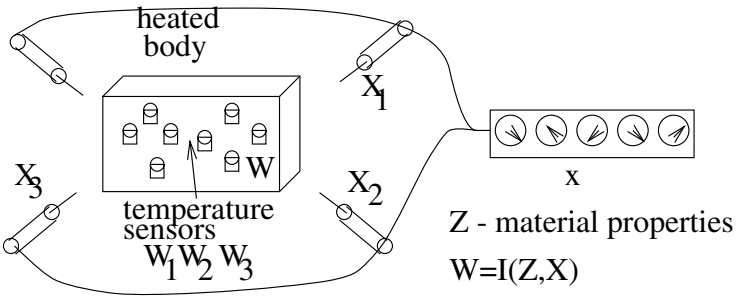

Fig. 5.1 Schematic representation of a participating medium (absorbent, emittant, and scatterer) being subject to radiation originated in external sources. Here **Z** represents material properties, **X** intensity of external sources of radiation and **W** = **I**(**Z**,**X**) measurements of temperature at several positions within the material body.

[1] In this example the measurements involve only temperature. Notwithstanding, problems involving thermal radiation measurements are discussed within this chapter, as well.

We denote by $\mathbf{W}$ a vector representing the output of this physical system, given by the resulting set of temperature measurements, comprising measurements in several specific points. By $\mathbf{X}$ we denote values representing levels and types of the regulation of the system (external sinks or sources), i.e. systems input, and by $\mathbf{Z}$ the physical/material properties of the body which influence the system's output.

The question is: given temperature measurements, $\mathbf{W}$, and knowing the value of the regulation, $\mathbf{X}$, obtain an estimate, $\mathbf{Z}$, of the physical properties values.

In a general way, consider the question of obtaining an estimate, $\mathbf{Z}$, for some constants, present in a mathematical model for a physical system's output. This output is represented by a physical magnitude, $\mathbf{W}$, obtained from experimental measurements of the real output of the physical system.

The hypotheses used throughout this chapter are

$$\mathbf{Z} = (Z_1, \ldots, Z_N)^T \in \mathbb{R}^N \,, \;\; \mathbf{W} = (W_1, \ldots, W_M)^T \in \mathbb{R}^M \,,$$

and $\mathbf{X}$ can be an element of $\mathbb{R}^K$, or even a function. We represent the dependence of $\mathbf{W}$ in $\mathbf{Z}$ and $\mathbf{X}$ in a functional form by $\mathbf{W} = \mathbf{I}(\mathbf{Z}, \mathbf{X})$.

The dependence of $\mathbf{I} = (I_1, \ldots, I_M)^T$ on $\mathbf{Z}$ can be explicit —as was the case considered in Chapter 1, when the relationship between the system inputs, $\mathbf{x}$, and the system outputs, $\mathbf{y}$, was given by a linear function, $\mathbb{R}^3 \ni \mathbf{x} \mapsto A\mathbf{x} \in \mathbb{R}^3$,— or implicit, if $\mathbf{W}$ and $\mathbf{Z}$ satisfy a given, possibly non-linear, system of equations, which would be written as

$$\mathbf{G}(\mathbf{W},\mathbf{Z},\mathbf{X}) = 0 \,,$$

for some known function $\mathbf{G} : \mathbb{R}^M \times \mathbb{R}^N \times \mathbb{R}^K \to \mathbb{R}^M$.

Another possibility is when $\mathbf{Z}$ is a parameter of a differential or integro-differential equation, and $\mathbf{W}$ represents some *observation* of its solution (such as the value of the solution in some points of the domain, or an average of the solution for some part of the domain) [68, 81]. In this case, $\mathbf{X}$ can be the value of an initial or boundary condition, or a source. Unless we can find the solution explicitly, we would say that $\mathbf{I}$ is given implicitly (as the solution of the appropriate equation).

In the first two cases the inverse problem is of Type I, while in the last one the inverse problem is of Type III, in accordance with the classification given in Table 2.3, page 48. This is the type with which we will deal in this chapter.

More generally, the relation between $\mathbf{W}$, $\mathbf{Z}$ and $\mathbf{X}$ implies a relation of cause-effect (stimulus-reaction), linear or non-linear, that can be explicit or implicit. Here $\mathbf{I}$ represents the *solution operator*—the abstract object that explicitly renders $\mathbf{W}$ as a function of $\mathbf{Z}$ and $\mathbf{X}$. In practice, $\mathbf{I}$ may be impossible to obtain explicitly. Sometimes its existence, or even its uniqueness and smooth dependence on data, can be abstractly proven. That information may be insufficient for application purposes, and must be complemented by some numerical solution. Fortunately, the qualitative theoretical results can bring to light behaviour and properties of the algorithm for numerical solution.

Given material properties $\mathbf{Z}$ and parameters $\mathbf{X}$ the i-th *predicted* output of the system is $I_i(\mathbf{Z}, \mathbf{X})$. This is compared with the effectively measured quantity (experimental data), W_i, defining the *residual*,

$$R_i = R_i(\mathbf{Z},\mathbf{X}) = I_i(\mathbf{Z},\mathbf{X}) - W_i \,. \tag{5.1}$$

Here, $\mathbf{Z} = (Z_1, \ldots, Z_N)^T \in \mathbb{R}^N$ is the vector of unknowns of the problem.

The inverse problem is solved as an optimization problem, on a space of finite dimension, in which we pursue the minimization of the functional representing half the sum of the squared residuals,

$$\begin{aligned} Q = Q(\mathbf{Z}) &= \frac{1}{2}|\mathbf{R}|^2 = \frac{1}{2}\,\mathbf{R}^T\mathbf{R} = \frac{1}{2}|\mathbf{I}(\mathbf{Z}) - \mathbf{W}|^2 \\ &= \frac{1}{2}\sum_{i=1}^{M}[I_i(\mathbf{Z}) - W_i])^2 \,, \end{aligned} \tag{5.2}$$

where $\mathbf{R} = (R_1, \ldots, R_M)^T \in \mathbb{R}^M$ represents the residual between computed magnitudes $\mathbf{I}$ and measurements (experimental data) $\mathbf{W}$, and M is the total number of experimental data available.

This formulation is similar to the minimization problems presented in Sections 2.6 and 3.4, and is an instance of the *least squares method*.

5.2 Modified Newton's Method

The functional given by Eq. (5.2), is minimized by finding its critical point, $\nabla Q = 0$, that is,

$$\frac{\partial Q}{\partial Z_k} = 0 \quad \text{for } k = 1, 2, \ldots, N \,, \tag{5.3}$$

which constitutes a system of N non-linear equations and N unknowns, $\mathbf{Z} = (Z_1, \ldots, Z_N)$. From Eqs. (5.1) to (5.3), the *critical point equation* is rewritten as

$$\sum_{i=1}^{M} R_i \frac{\partial I_i}{\partial Z_k} = 0 \quad \text{for } k = 1, 2, \ldots, N \,. \tag{5.4}$$

We solve Eq. (5.4) by means of a modified *Newton's method*, that can be deduced as follows. Using a Taylor's expansion of $\mathbf{R}$ around $\mathbf{Z}^n$, where n will be the index of the iterations in the iterative procedure, and keeping only the zero and first order terms, we have

$$R_i^{n+1} = R_i^n + \sum_{j=1}^{N} \frac{\partial R_i^n}{\partial Z_j} \Delta Z_j^n \,, \quad \text{for } i = 1, 2, \ldots, M \,. \tag{5.5}$$

Here, $\mathbf{R}^n$ represents the evaluation of $\mathbf{R}$ in $\mathbf{Z}^n$,

$$\mathbf{R}^n \;=\; \mathbf{R}(\mathbf{Z}^n) \;=\; \mathbf{I}(\mathbf{Z}^n) - \mathbf{W} \,, \tag{5.6}$$

and,

$$\Delta\mathbf{Z}^n \quad = \quad \mathbf{Z}^{n+1} - \mathbf{Z}^n \,,$$

or, in coordinates, $R_i^n = R_i(\mathbf{Z}^n)$ and $\Delta Z_j^n = Z_j^{n+1} - Z_j^n$.

Using Eq. (5.5) in the system of equations (5.4), and noticing that $\partial R_i/\partial Z_j = \partial I_i/\partial Z_j$, we obtain,

$$\sum_{i=1}^{M}\left(R_i^n + \sum_{j=1}^{N}\left.\frac{\partial I_i}{\partial Z_j}\right|_{Z=Z^n}\Delta Z_j^n\right)\left.\frac{\partial I_i}{\partial Z_k}\right|_{Z=Z^n} = 0\,, \tag{5.7}$$

for $k = 1, 2, \ldots, N$.

We look at $\mathbf{I}$ as a function of $\mathbf{Z}$ only, making $\mathbf{X}$ constant. By definition, the *Jacobian matrix* of $\mathbf{I}$ with respect to $\mathbf{Z}$, $\mathcal{J} = \mathcal{J}\mathbf{I}|_Z$, has entries

$$\mathcal{J}_{ij} = \partial I_i/\partial Z_j\,,$$

for $i = 1, \ldots, M$, and $j = 1, \ldots, N$, that is,

$$\mathcal{J} = \begin{pmatrix} \frac{\partial I_1}{\partial Z_1} & \cdots & \frac{\partial I_1}{\partial Z_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial I_M}{\partial Z_1} & \cdots & \frac{\partial I_M}{\partial Z_N} \end{pmatrix}.$$

The system of equations (5.7) can, therefore, be rewritten in the form known as *normal equation*

$$(\mathcal{J}^n)^T\mathcal{J}^n\Delta\mathbf{Z}^n = -(\mathcal{J}^n)^T\mathbf{R}^n\,, \tag{5.8}$$

where $\mathcal{J}^n$ represents $\mathcal{J}\mathbf{I}|_{Z=Z^n}$.

An iterative procedure can be constructed to determine the vector of unknowns $\mathbf{Z}$ that minimizes the functional Q, knowing the experimental data, $\mathbf{W}$, and computed values, $\mathbf{I}$, which depend on the unknowns to be determined, $\mathbf{Z}$.

Starting from an initial estimate, $\mathbf{Z}^0$ and measurements $\mathbf{W}$, residuals are computed from Eq. (5.6), and corrections are computed sequentially from Eq. (5.8), where n is the iteration counter. The algorithm can be written as[2],

$$\mathbf{R}^n = \mathbf{I}(\mathbf{Z}^n, \mathbf{X}) - \mathbf{W} \tag{5.9a}$$

$$(\mathcal{J}^n)^T\mathcal{J}^n\Delta\mathbf{Z}^n = -(\mathcal{J}^n)^T\mathbf{R}^n \tag{5.9b}$$

$$\mathbf{Z}^{n+1} = \mathbf{Z}^n + \Delta\mathbf{Z}^n\,. \tag{5.9c}$$

[2] The method described here is a modification of Newton's method (presented in Section 4.3, page 92). We remark that the goal is not Newton's method, but the solution of Eq. (5.4), that is obtained here by means of the modified Newton's method, given by Eq. (5.9). For the problem treated here Newton's method demands the computation of second order derivatives of $\mathbf{I}$, while the method based on Eq. (5.8) avoids that. See Exercise 5.1.

The iterative procedure is interrupted when a convergence criterion defined *a priori* is satisfied. For example,

$$\left|\Delta Z_j^n / Z_j^n\right| < \varepsilon \,, \text{ for all } j \text{ with } j = 1, 2, \ldots, N \,.$$

Here, ε is a sufficiently small value, say 10^{-5}. Another possibility is the use of the vector of corrections norm, $|\Delta \mathbf{Z}^n| < \epsilon$ as considered in Newton's method.

Observe that in every iteration the values $I_i^n = I_i(\mathbf{Z}^n)$, $i = 1, 2, \ldots, M$, are computed using the estimates for the unknowns, $\mathbf{Z}^n$. This involves, in the examples discussed later in this chapter, the solution of differential or integro-differential equations. This makes them implicit inverse problems.

Finally, we remark that the solution of Eq. (5.9b) can be written as

$$\Delta \mathbf{Z}^n = -\left[(\mathcal{J}^n)^T \mathcal{J}^n\right]^{-1} (\mathcal{J}^n)^T \mathbf{R}^n \tag{5.10}$$

Equation 5.10 represents explicitly the solution of Eq. (5.8), making use of the inverse[3] of $(\mathcal{J}^n)^T \mathcal{J}^n$.

5.3 Levenberg-Marquardt's Method

The methods presented in Section 4.3 and the algorithm presented in the previous section, are Newton-like, and can encounter convergence difficulties if the initial estimate for the vector of unknowns, $\mathbf{Z}^0$, is not adequately selected. Choosing an adequate initial estimate can prove extremely difficult.

In 1963, Marquardt [54] designed the algorithm that will be described presently, with the objective of reaching convergence with a wider range of initial estimates. One of the referees of his work noticed that, in 1944, Levenberg had made a similar proposal: adding a term in the diagonal of matrix $\mathcal{J}^T \mathcal{J}$. The method came to be known as the *Levenberg-Marquardt method*.

Based on Eq. (5.8), the Levenberg-Marquardt method considers the determination of the corrections $\Delta \mathbf{Z}^n$ by means of the following equation,

$$\left[(\mathcal{J}^n)^T \mathcal{J}^n + \lambda^n \mathcal{I}\right] \Delta \mathbf{Z}^n = -(\mathcal{J}^n)^T \mathbf{R}^n \,. \tag{5.11}$$

Here $\lambda = \lambda^n$ is the *damping factor* and $\mathcal{I}$ represents the identity matrix. Observe that this formulation is similar to the Tikhonov's regularization, Eq. (3.22).

Similar to the developments in the previous section, an iterative algorithm is built to determine the vector of unknowns $\mathbf{Z}$ that should minimize the functional Q. The procedure is based on Eq. (5.11). From an initial estimate, $\mathbf{Z}^0$, corrections are sequentially computed,

$$\Delta \mathbf{Z}^n = -\left[(\mathcal{J}^n)^T \mathcal{J}^n + \lambda^n \mathcal{I}\right]^{-1} (\mathcal{J}^n)^T \, \mathbf{R}^n \,, \text{ for } n = 0, 1, \ldots, \tag{5.12}$$

[3] In computations, one rarely inverts a matrix due to its high computational cost. It is preferable to solve the system of equations. Therefore, the corrections $\Delta \mathbf{Z}^n$ are computed by solving the linear algebraic system of equations (5.8). For theoretical considerations it is sometimes useful to have closed form solutions, that is, to have the solution written in terms of a solution operator.

where n is an iteration counter, and the new estimates for the vector of unknowns are computed by Eq. (5.9c). The iterations are interrupted when a convergence criterion established *a priori* is satisfied.

It should be noticed that the solution of the problem described by Eq. (5.11) differs from the one given by Eq. (5.8). On the other hand, our aim is to solve Eq. (5.4). To guarantee the convergence, at the beginning of the iterative process, a relatively high value is assigned to λ, λ^0, thus emphasizing the importance of the diagonal of matrix $(\mathcal{J}^T\mathcal{J} + \lambda\mathcal{I})$ relative to the information contained in the elements outside the diagonal. Through the iterative procedure, the value of the damping factor λ is to be reduced, in such a way that its value approaches zero as the procedure approaches its conclusion. In the light of the notion of regularization, described in Chapter 3, Eq. (5.11) is a kind of regularization of Eq. (5.8).

An algorithm to control the value of the damping factor will be described shortly, [54]. Let $c > 1$, $d > 1$ and $\xi > 1$. Let also $Q^n = Q(\mathbf{Z}^n)$. When $Q^{n+1} \leq Q^n/c$, the reduction $\lambda^{n+1} = \lambda^n/d$ is performed. Otherwise, $\lambda^n = \xi\lambda^{n-1}$ is taken, and a new estimate for the vector of unknowns $\mathbf{Z}^{n+1}$ is computed for the same value of the iteration counter n, using again the previous estimate $\mathbf{Z}^n$. Silva Neto and Özişik, [74, 75, 79], used $c = d = \xi = 2$ in the applications in heat transfer by means of thermal radiation.

5.4 Confidence Intervals for Parameter Estimates

Folowing Gallant [34, 38, 59], the *confidence intervals* of the estimates of the parameters $\mathbf{Z}$ are computed using the *sensitivity coefficients*,

$$\partial I_i/\partial Z_j \,, i = 1, \ldots, M \;\text{ and }\; j = 1, \ldots, N \,,$$

and the *standard deviation*, σ, of the error present in the experimental data.

Let $\nabla\mathbf{I}$ be the $M \times N$ matrix whose entries are given by the sensitivity coefficients,

$$(\nabla\mathbf{I})_{ij} \;=\; \left.\frac{\partial I_i}{\partial Z_j}\right|_Z \,.$$

In this case, the square of the *standard deviation* of the estimators of the parameters are given by

$$\sigma_Z^2 \;=\; \left(\sigma_{Z_1}^2, \ldots, \sigma_{Z_N}^2\right)^T = \sigma^2 \left\{\text{diag}\left[(\nabla\mathbf{I})^T\,\nabla\mathbf{I}\right]^{-1}\right\} \tag{5.13a}$$

where $\text{diag}(A)$ represents the vector whose elements are the elements of the diagonal of a matrix A.

Assuming a normal distribution for the experimental errors, with zero mean, the 99 % confidence intervals for the estimates Z_j are [33]

$$\left]Z_j - 2.576\sigma_{Z_j},\, Z_j + 2.576\sigma_{Z_j}\right[\qquad \text{for } j \,=\, 1, \ldots, N \,. \tag{5.13b}$$

In general, smaller confidence intervals are associated with larger sensitivity coefficients and smaller experimental errors, thus producing better estimates.

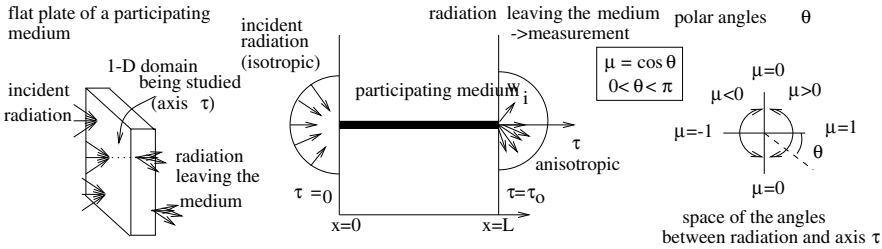

Fig. 5.2 Radiative transfer in a participating medium. Optical variable τ is related (sometimes linearly) to the spatial variable x.

5.5 Phase Function, Albedo and Optical Thickness

Different types of radiation, like neutral particles, gamma rays and photons have been used to identify objects in industry (through non-destructive tests) and also in medicine (for diagnosis and therapy).

Heat transfer by thermal radiation in a *participating medium*, that is, one that emits, absorbs and scatters radiation, schematically represented in Figs. 5.1 and 5.2, is modeled according to the linear version of the Boltzmann equation [60, 80, 69, 57].

It is worthwhile to mention that the physical phenomena relevant to neutron transport in nuclear reactors, or to tomography with scattering (NIROT—*Near Infrared Optical Tomography*) can also be represented mathematically by the linear Boltzmann equation.

Consider the situation depicted in Fig. 5.2 representing a flat plate made of a scattering anisotropic, gray material with transparent boundary surfaces, subject to external isotropic radiation on its left surface, in a permanent regimen (steady-state — it does not depend on time).

A material is *anisotropically* scattering when the scattering depends on angle and it is *gray* if the properties do not depend on the radiation's wavelength. A material has a *transparent* surface when this surface does not reflect radiation.

In this case, and also considering azimuthal symmetry and a cold medium (no emission), the linear *Boltzmann equation* is written as [60] (see Fig. 5.2)

$$\mu \frac{\partial I}{\partial \tau}(\tau,\mu) + I(\tau,\mu) = \frac{\omega}{2}\int_{-1}^{1} p(\mu,\mu')I(\tau,\mu')d\mu' \,, \tag{5.14a}$$

in $0<\tau<\tau_0$, $-1\le\mu\le 1$, and

$$I(0,\mu) = 1, \quad \mu > 0, \text{ and } I(\tau_0,\mu) = 0, \text{ for } \mu < 0 \,. \tag{5.14b}$$

In this equation, $I(\tau,\mu)$ is the radiation intensity in position τ, following direction represented by μ. Here, τ is the spatial optical variable, and μ is the cosine of the polar angle θ formed between the direction of the radiation beam and the τ axis. Also, ω is the single scattering *albedo* (the ratio between the scattering and the extinction coefficients, σ_s and β, respectively, with $\beta = \sigma_s + k_a$, and k_a is the absorption

coefficient), τ_0 is the medium's *optical thickness* (related to the geometrical thickness of the medium) and $p(\mu, \mu')$ is the *anisotropic scattering phase* function.

We remark that $\frac{1}{2}p(\mu, \mu')$ represents the probability density of an incident beam with direction μ' to be scattered following the direction represented by μ. More explicitly, the probability that the scattered direction μ is between μ_1 and μ_2 given that the incident direction is μ' is given in terms of $p(\mu,\mu')$ by

$$P(\mu_1 \leq \mu \leq \mu_2 \mid \mu') = \frac{1}{2} \int_{\mu_1}^{\mu_2} p(\mu,\mu')\, d\mu \,.$$

The medium is *isotropic* when the scattering is uniform in all directions, i.e., when $p(\mu, \mu') = c$, for all $\mu,\mu' \in [-1,1]$, where c is a constant value. Since

$$P(-1 \leq \mu \leq 1 \mid \mu') = 1 \,, \tag{5.15}$$

it is then necessary that $c = 1$.

The second term in the left hand side of Eq. (5.14a) represents the absorption and scattering of radiation by the medium away from the direction represented by μ (out scattering) and the right hand side represents the way the radiation is scattered by the medium into such direction (in scattering). The emission of radiation by the medium can be neglected if compared to the incident radiation in $\tau = 0$. We recall that we are considering here a *steady state* problem (it does not depend on time).

When the operator, the medium's geometry (in this case, the optical thickness τ_0 for the plane-parallel medium), the material properties (here, ω and $p = p(\mu, \mu')$) and the boundary conditions (given in this example by Eq. (5.14b)) we say that the model is characterized. That is, it is modeled by steady state linear Boltzmann equation, with specific type Dirichlet boundary conditions and identified (all constants and auxiliary functions are given). In this case we deal with a direct problem, and the radiation intensity $I(\tau, \mu)$ can be computed in all of the spatial domain $0 \leq \tau \leq \tau_o$ and the angular domain $-1 \leq \mu \leq 1$.

Different analytic and numerical techniques have been developed to solve the linear radiative *transport* equation, Eq. (5.14). Wick [93] and Chandrasekhar [20, 21] created the *discrete-ordinates method*, by replacing the right-side term of Eq. (5.14a) by a Gaussian quadrature term. This leads to a system of ordinary differential equations with as many equations as points used in the quadrature. Silva Neto and Roberty [80] presented a comparison between spherical harmonics expansion methods, P_N, Galerkin, global base and discrete-ordinates (i.e., finite differences + Gaussian quadrature), S_N, for the case of isotropic scattering. Chalhout et al. [19] considered three variations of the discrete ordinates method and performed a comparison with the Monte Carlo method. Moura Neto and Silva Neto [57, 63] presented solutions using methods with integrating factor and operator splitting.

We use the customary representation of the scattering phase function by expansion in Legendre polynomials [60, 79, 47],

$$p(\mu, \mu') = \sum_{l=0}^{L} (2l + 1)\, f_l P_l(\mu) P_l(\mu') \,, \tag{5.16}$$

where f_l, $l = 0, 1, \ldots, L$ are the coefficients of expansion with $f_0 = 1$.

In the example presented here, we consider the inverse problem of estimating simultaneously the medium's optical thickness, τ_0, the single scattering albedo, ω, and the scattering phase function, $p(\mu,\mu')$, by means of its coefficients, f_l, $l=1,2,\ldots,L$, in the expansion in Legendre polynomials, Eq. (5.16) [79, 47].

The vector of unknowns is made up of the following elements

$$\mathbf{Z} = (\tau_0, \omega, f_1, \ldots, f_L)^T \ ,$$

to be determined using the experimental measurements of the radiation that leaves the medium, W_i, $i = 1,2,\ldots,M$, by minimizing the functional

$$\begin{aligned} Q &= Q(\tau_0, \omega, f_1, \ldots, f_L) \\ &= \frac{1}{2}|\mathbf{R}|^2 = \frac{1}{2}\sum_{i=1}^{M} [I_i[\mathbf{Z}] - W_i]^2 \ . \end{aligned}$$

Here $I_i[\mathbf{Z}]$, $i = 1,2,\ldots,M$ are the values of the radiation intensities computed by the solution of the direct problem described by Eq. (5.14) and evaluated in the same directions in which the radiation leaving the medium is measured, μ_i, $i = 1,2,\ldots,M$ which are pre-defined directions, assuming the parameters $\mathbf{Z} = (\tau_0, \omega, f_1, f_2, \ldots, f_L)$ are known,

$$I_i[\mathbf{Z}] = I[\tau_0, \omega, f_1, f_2, \ldots, f_L](\mu_i) \ , \text{ for } i = 1, 2, \ldots, M \ .$$

The Levenberg-Marquardt method, described in Section 5.3, is used to solve this optimization problem in finite dimension.

Since we do not have experimental data on this problem, we use *synthetic* or artificial data. For that we mean data generated from solving the direct problem with known parameters and adding random values to simulate experimental errors. In the example we are considering, we assume that the parameters $\mathbf{Z} = (\tau_0, \omega, f_1, \ldots, f_L)^T$ are known in advance and we use them to solve the direct problem, Eq. (5.14). The synthetic data is then determined by

$$W_i = I_i[\tau_0, \omega, f_1, \ldots, f_L](\mu_i) + c\epsilon_i \ ,$$

where ϵ_i is a realization of a uniform random variable in the interval $[-1,1]$,

$$c = \gamma \max_i I(\tau_0, \mu_i), \text{ for } i = 1, \ldots, M \ ,$$

and γ is a maximum percentage error parameter.

It should be remarked that the solution of inverse problems with synthetic data makes it possible to verify that the computational procedure is correct, before applying it to real problems.

The value of the phase function is related to the angle subtended between the directions of the incident radiation and the scattered radiation [22]. The results presented in Figs. 5.3 and 5.4 where computed assuming the incident radiation to be exclusively in the direction of vector $(1, 0)$. A representation in polar coordinates is

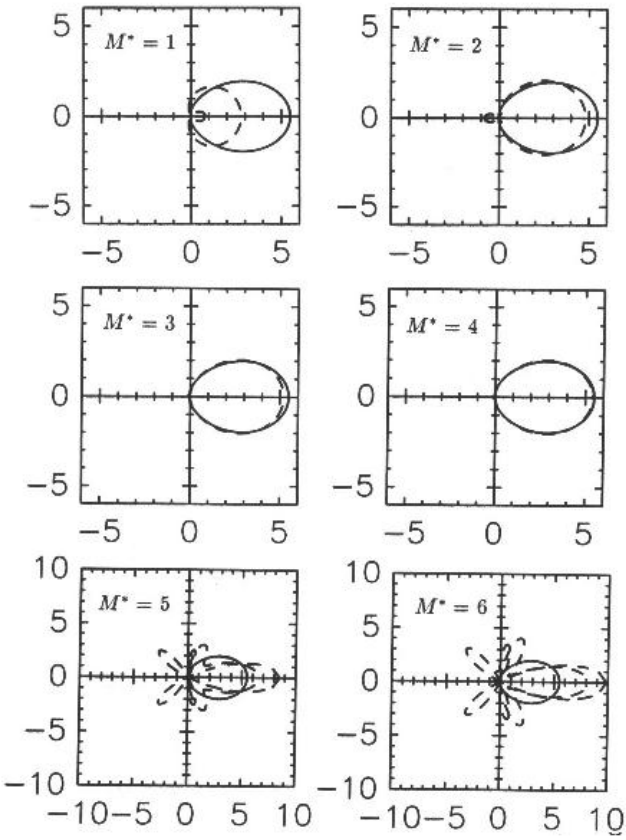

Fig. 5.3 Estimation of an anisotropic scattering phase function, with $L = 7$, characterizing forward scattering, $\omega = 0.5$ and $\tau_0 = 2.0$. For forward scattering, the phase function representation is restricted to the 1^{st} and 4^{th} quadrant. The experimental error reached 6.2 % of the highest intensity value that was measured. Here, M^* is the number of coefficients of the phase function that were considered in the estimations. Solid line represents the exact phase function, and dashed line represents the estimates for the phase function obtained with different values for M^*. The best result occurs with $M^* = 4$.

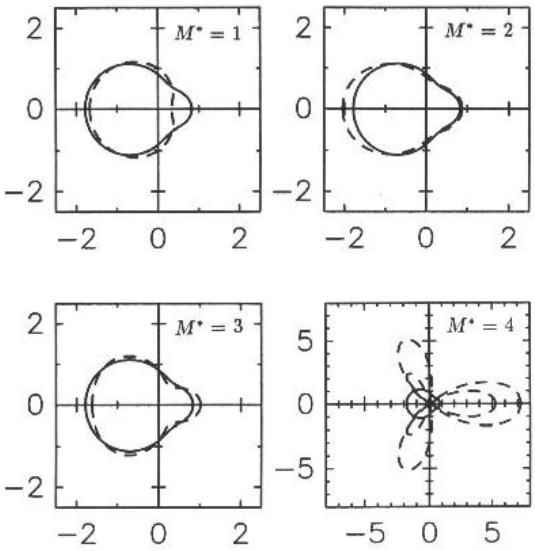

Fig. 5.4 Estimation of a phase function with $L = 5$, with preferential backward scattering, $\omega = 0.1$ and $\tau_0 = 10.0$. In a preferred backward scattering, one should have $\int_{-1}^{0} p(\mu,1)\,d\mu > \int_{0}^{1} p(\mu,1)\,d\mu$, which is the case if $p(\mu,1) < p(-\mu,1)$, for $\mu > 0$, as shown in the graph of the phase function. The experimental error reached 4.1 % of the highest intensity measured value. The number of coefficients of the phase function, M^*, were chosen as 1 to 4. Solid line represents the exact phase function, and dashed line represents the phase function obtained estimate with different values for M^*. Fairly good results were obtained with $M^* = 2$ and $M^* = 3$.

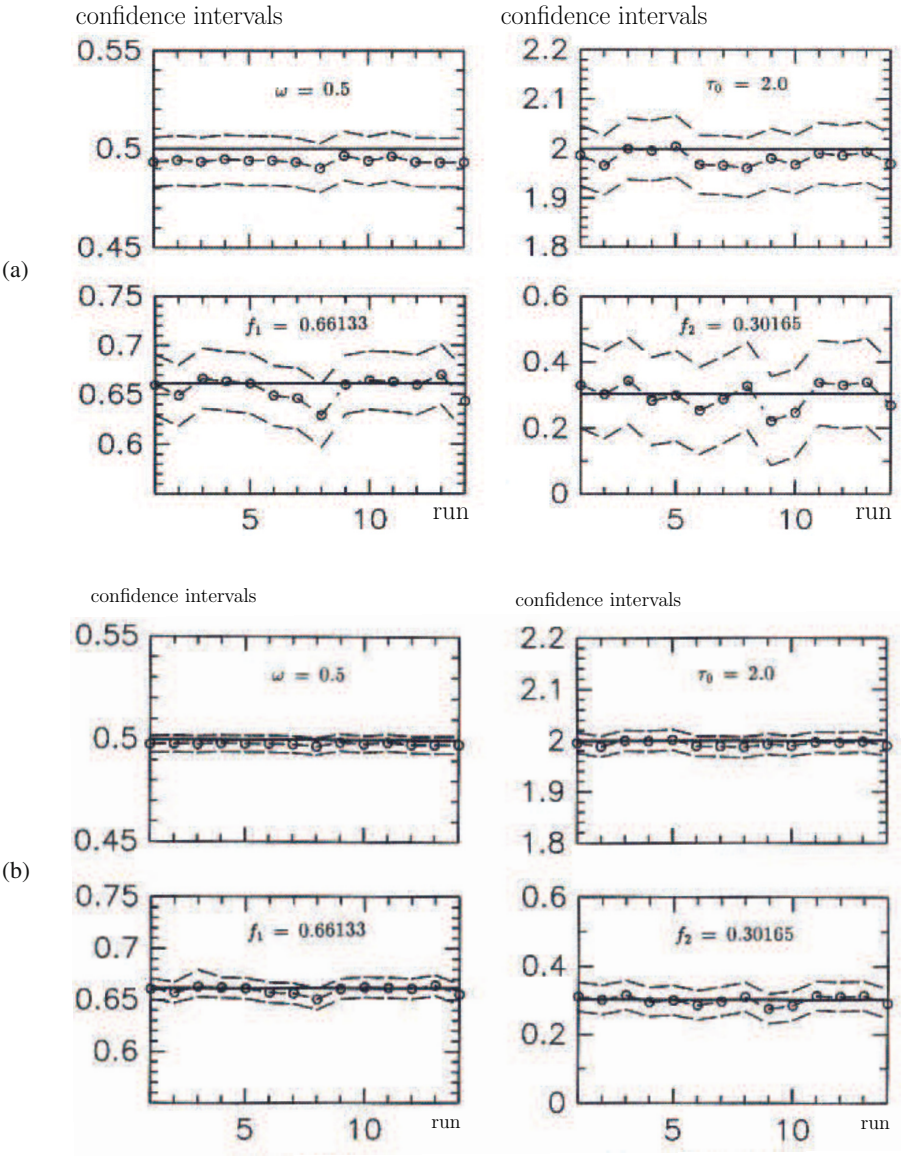

Fig. 5.5 Confidence intervals for the single scattering albedo (ω), optical thickness (τ_0), and first two coefficients of the anisotropic scattering phase function expansion (f_1 and f_2) estimates when the phase function has $L = 7$ terms.
– exact values, - - - confidence intervals, -●-●- estimates. Experimental error in the highest mean value of the measured intensities: (a) 6.2 %, and (b) 2.1 %.

given in which the distance from the point on the graph to the origin provides the probability density for the radiation to be scattered according to that direction.

Figure 5.3 represents the phase function estimation of an anisotropic scattering, with preferential forward scattering with $L = 7$, (see Eq. (5.16)). Due to the experimental error occurring in the measurements, $\mathbf{W}$, of $\gamma = 6.2\,\%$ of the highest value of the measured intensity (where, 6.2 % is, therefore, the lowest percentage of experimental error in the measurements being considered), it is not possible to recover all the coefficients. The highest-order coefficients, with relatively low numerical values, are more rapidly affected, causing the estimation to deteriorate. However, the relevant information for the design of thermal equipments is the shape of the phase function, and not necessarily the value of each coefficient separately.

As a matter of fact, the number of coefficients, L, of the anisotropic scattering phase function represented in Eq. (5.16) is also an unknown of the problem. Silva Neto and Özişik [79] developed a criterion to choose the number of coefficients in the expansion M^* in such a way to obtain the best possible estimated phase function, considering the available experimental data.

In the test case represented in Fig. 5.3, $M^* = 3$ or 4 values would be chosen. It must be pointed out that, in the absence of experimental errors (an ideal situation that does not happen in practice) all seven coefficients, i.e., $M^* = L = 7$, were estimated within the precision established *a priori* in the stopping criterion.

Figures 5.5a and b show the estimates for ω, τ_0, and for the first two coefficients, f_1 and f_2, of the phase function represented in Fig. 5.3. The different executions of the computational code correspond to estimates due to different sets of experimental data. Figure 5.5a presents the results when the lowest experimental error reaches 6.2 % and, in Fig. 5.5b, the lowest experimental error reaches 2.1 %.

Figure 5.4 presents the results on the estimation of a phase function with $L = 5$, corresponding to a medium with preferential backward scattering. The lowest experimental error here considered is 4.1 %. In this example we would choose $M^* = 2$ or 3.

Using the expression for the confidence intervals, Eq. (5.13), we are able to compute them for the various parameters being estimated, and we represent them in graphical form in Fig. 5.5. As expected, the estimates, in the examples given here, have narrower confidence intervals, when the experimental data presents lower levels of noise.

5.6 Thermal Conductivity, Optical Thickness and Albedo

Silva Neto and Özişik [74] solved an inverse problem involving heat transfer due to conduction and radiation in a *participating medium*, considering only isotropic scattering (not depending on the polar angle). Lobato et al. [48] dealt with a similar problem, but using stochastic methods [69] for the minimization of the squared residues functional.

Consider the situation illustrated in Fig. 5.6. A plane-parallel, gray, isotropically scattering medium with transparent boundary surfaces, is subject to incident external

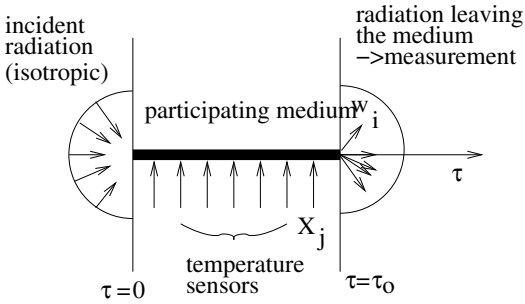

Fig. 5.6 Heat transfer by thermal conduction and radiation in a participating medium

isotropic radiation that reaches surface $\tau = 0$. The surfaces at $\tau = 0$ and $\tau = \tau_0$ are kept at constant temperatures T_1 and T_2, respectively.

The mathematical formulation of the heat transfer problem due to one dimensional, steady-state heat conduction, in which the coupling with the transfer due to radiation in the participating medium is achieved by the source term, is given, in a dimensionless formulation, by a boundary value problem of an ordinary differential equation (*Poisson's equation*, with a non-linear source term and Dirichlet boundary conditions) [60]

$$\frac{d^2\Theta}{d\tau^2} - \frac{(1-\omega)}{N}\left[\Theta^4(\tau) - G^*[I](\tau)\right] = 0\,, \quad \text{in } 0 < \tau < \tau_0\,, \tag{5.17a}$$

with boundary conditions,

$$\Theta(0) = 1 \text{ and } \Theta(\tau_0) = T_2/T_1\,. \tag{5.17b}$$

Here

$$G^*[I](\tau) = \frac{1}{2}\int_{-1}^{1} I(\tau,\mu)\,d\mu\,, \quad N = \frac{k\beta}{4\,n^2\,\overline{\sigma}\,T_1^3}\,, \quad \Theta = \frac{T}{T_1}\,, \tag{5.17c}$$

where Θ is the dimensionless temperature, N is the conduction-radiation parameter, ω is the simple scattering albedo, k is the thermal conductivity, β is the extinction coefficient (absorption + scattering), n is the medium's refractive index, and $\overline{\sigma}$ is the Stefan-Boltzmann's constant.

Modeling of the radiative transfer in the participating medium is achieved by means of the *linear* Boltzmann's equation [60],

$$\mu\,\frac{\partial I}{\partial \tau}(\tau,\mu) + I(\tau,\mu) = H(\Theta(\tau)) + \frac{\omega}{2}\int_{-1}^{1} I(\tau,\mu')\,d\mu'\,, \tag{5.17d}$$

in $0 < \tau < \tau_0$ and $-1 \le \mu \le 1$, and

$$I(0,\mu) = 1\,, \text{ for } \mu > 0, \text{ and } I(\tau_0,\mu) = 0\,, \text{ for } \mu < 0\,, \tag{5.17e}$$

where the source term, $H(\Theta)$, is related to the medium's temperature distribution,

$$H(\Theta) \quad = \quad (1-\omega)\,\Theta^4\,, \tag{5.17f}$$

and the remaining symbols have already been defined. We observe that since the medium is an isotropic scatterer, the phase function of scattering is

$$\frac{1}{2}p(\mu,\mu') = \frac{1}{2} \text{ for all } \mu,\ \mu'\ .$$

Equation (5.17) provides a complete mathematical formulation for the one dimensional heat transfer problem, in steady state regime, by the combined mode of conduction and radiation. The problem of conduction, Eqs. (5.17a)–(5.17c), and the radiation problem, Eqs. (5.17d)–(5.17f), are coupled by means of the source terms, given respectively by

$$G^* = G^*[I] \text{ and } H = H(\Theta)\ .$$

To solve Eq. (5.17), we use an iterative procedure. Starting with a first estimate of I, we solve Eqs. (5.17a)–(5.17c) to obtain an estimate of Θ. From this estimate of Θ, we solve Eqs. (5.17d)–(5.17f) to obtain a new estimate of I. This is done until convergence is reached.

In the solution of the direct problem, Silva Neto and Özişik [74] used the iterative procedure described with the Galerkin method, global basis for the part of the problem related to heat transfer due to radiation in a participating medium, Eqs. (5.17d)–(5.17f), and the finite difference method for the part of the problem related to heat transfer by conduction, Eqs. (5.17a)–(5.17c).

In the inverse problem just presented, we consider the simultaneous estimation of the optical thickness, τ_0, the single scattering albedo, ω, and the conduction-radiation parameter, N. We use synthetic experimental measurements of the radiation, W_i, $i = 1, 2, \ldots, M$, and of temperature inside the medium, represented by X_j, $j = 1, 2, \ldots, K$. The vector of unknowns $\mathbf{Z} = (\tau_0, \omega, N)^T$ is determined — the model is identified — as the minimum point by minimization of the functional

$$\begin{aligned} Q = Q(\tau_0,\omega,N) = & \frac{1}{2}\sum_{i=1}^{M}\left[I_i\left(\tau_0,\omega,N\right) - W_i\right]^2 \\ & + \frac{1}{2}\sum_{j=1}^{K}\left[\Theta_j\left(\tau_0,\omega,N\right) - X_j\right]^2\ , \end{aligned} \tag{5.18}$$

where $I_i(\tau_0,\omega,N)$, $i = 1, 2, \ldots, M$, are the radiation intensities computed in the same surface and in the same directions in which the radiation is measured, W_i, $i = 1, 2, \ldots, M$. Here, $\Theta_j(\tau_0,\omega,N)$, $j = 1, 2, \ldots, K$, are temperatures computed in the same positions where the temperatures are measured, X_j, $j = 1, 2, \ldots, K$. The radiation and temperature intensities are computed by solving Eq. (5.17), following the procedure already described.

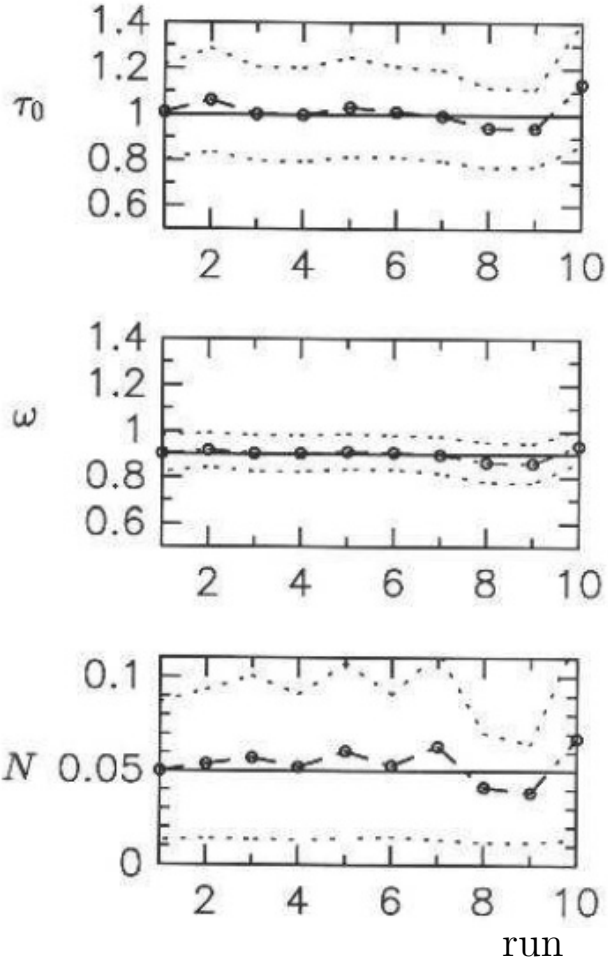

Fig. 5.7 Confidence intervals for the optical thickness (τ_0), single scattering albedo (ω), and conduction-radiation parameter (N) estimates in the combined conduction-radiation heat transfer model.
– exact values, - - - confidence intervals, -●-●- estimates. Experimental error of 4 % of the largest value of the magnitudes that were measured.

Figure 5.7 presents the results of the parameters (τ_0, ω, N) estimation, for a test case in which the exact values,

$$(\tau_0, \omega, N) = (1.0,\ 0.9,\ 0.05)\,,$$

were known *a priori*. The different results considered represent, as before, estimates, obtained with the execution of the computational code with different sets of synthetic experimental data. The artificially generated experimental error is set to 4 % of the value of the largest measured magnitude.

Conduction-radiation parameter, N, is relatively small, which indicates a dominance of the radiative heat transfer mechanism. This fact is proved by the relatively large size of confidence intervals[4] of the estimates of N (N depends on the medium's thermal conductivity, k), when compared to those obtained for the parameters τ_0 and ω.

5.7 Refractive Index and Optical Thickness

Consider a gray flat plate in *radiative equilibrium*, with two gray boundary surfaces, *opaque* —diffuse emittant and reflector (non-specular),— with emissivity ε and reflectivity ρ, which are kept at constant temperatures T_0 and T_L (see Fig. 5.8).

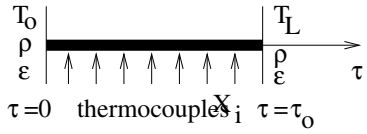

Fig. 5.8 Schematic representation of a one dimensional medium in radiative thermal equilibrium

The temperature distribution inside the medium, $T(\tau)$ satisfies [75],

$$\frac{T^4(\tau) - T_L^4}{T_0^4 - T_L^4} = \frac{\theta(\tau) + \left[\frac{1}{\varepsilon} - 1\right] S}{1 + 2\left[\frac{1}{\varepsilon} - 1\right] S}\,, \tag{5.19}$$

where the function $\theta(\tau)$ satisfies the *integral equation*

$$\theta(\tau) = \frac{1}{2}\left[E_2(\tau) + \int_0^{\tau_0} \theta(\tau')\, E_1(|\tau - \tau'|)\, d\tau'\right]. \tag{5.20}$$

Here, $E_m(\tau)$ represents the m-th *integral exponential function*, given by

$$E_m(\tau) = \int_0^1 \eta^{m-2}\, e^{-\frac{\tau}{\eta}}\, d\eta\,, \tag{5.21a}$$

[4] The confidence intervals are related to the sensitivity coefficients $\partial I/\partial\tau_0$, $\partial I/\partial\omega$, $\partial I/\partial N$, $\partial\Theta/\partial\tau_0$, $\partial\Theta/\partial\omega$, and $\partial\Theta/\partial N$.

and the constant S is given by

$$S = 1 - 2\int_0^{\tau_0} \theta(\tau')\, E_2(\tau')\, d\tau' \,. \tag{5.21b}$$

Given the optical thickness of the medium, τ_0, function $\theta(\tau)$ is computed by Eq. (5.20). If $\theta(\tau)$ is known, parameter S is computed using Eq. (5.21b). The temperature distribution inside the medium is then computed by Eq. (5.19).

For opaque and gray surfaces, $\varepsilon = 1-\rho$. When the refractive index of the medium, n_m, is higher than that of its environment, n_e, the reflectivity ρ is related to the relative refractive index $n = n_\mathrm{m}/n_\mathrm{e}$ as follows,

$$\rho(n) = 1 - \frac{1}{n^2}\left\{\frac{1}{2} - \frac{(3n+1)(n-1)}{6(n+1)^2} - \frac{n^2(n^2-1)^2}{(n^2+1)^3}\ln\left(\frac{n-1}{n+1}\right) + \frac{2n^3(n^2+2n-1)}{(n^2+1)(n^4-1)} - \frac{8n^4(n^4+1)}{(n^2+1)(n^4-1)^2}\ln(n)\right\} . \tag{5.22}$$

In the inverse problem presented here, the simultaneous estimation of the relative refractive index, n, and the medium's optical thickness, τ_0, is considered. The vector of unknowns $\mathbf{Z} = (n, \tau_0)^T$, is to be determined from the experimental measurements inside the medium, X_i, $i = 1, 2, \ldots, M$, by minimizing the functional

$$Q(n, \tau_0) = \frac{1}{2}|\mathbf{R}|^2 = \frac{1}{2}\sum_{i=1}^{M} [T_i\,(n, \tau_0) - X_i]^2 \,,$$

where $T_i(n, \tau_0)$, $i = 1, 2, \ldots, M$ are the temperatures computed in the same positions in which the experimental data are measured, X_i, using the problem described by Eqs. (5.19) through (5.22).

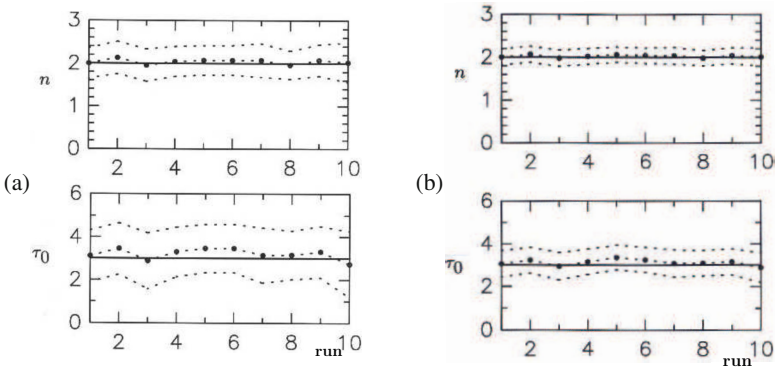

Fig. 5.9 Confidence intervals for the estimates.
– exact values, - - - confidence intervals, -●-●- estimates. Experimental error of the highest temperature measurement: (a) 5 %; (b) 2.5 %.

Results of the estimates of (n, τ_0) are presented in Fig. 5.9, in which the exact values for the test case, $(2.0, 3.0)$, were known *a priori*. The results correspond to estimates obtained by means of different sets of experimental synthetic data. Two levels of experimental errors corresponding to 5 % and 2.5 % are used, and the results are presented, respectively, in Figs. 5.9a and b.

As expected, these estimates are better (smaller confidence intervals for the same confidence level) when the experimental data present lower noise levels.

Exercises

5.1. We recall that Newton's method, presented in Section 4.3, for the system of equations $\mathbf{G}(\mathbf{Z}) = 0$, where $\mathbf{G} : \mathbb{R}^N \to \mathbb{R}^N$ is a vector-valued function, is given by the iterative scheme,

$$\begin{aligned} \mathcal{J}\mathbf{G}^n \Delta\mathbf{Z}^n &= -\mathbf{G}^n \\ \mathbf{Z}^{n+1} &= \mathbf{Z}^n + \Delta\mathbf{Z}^n \,, \end{aligned}$$

where $\mathcal{J}\mathbf{G}^n$ represents the Jacobian matrix of $\mathbf{G}$, evaluated at $\mathbf{Z}^n$

$$\mathcal{J}\mathbf{G}^n = \mathcal{J}\mathbf{G}|_{Z=Z^n} \,,$$

and $\mathbf{G}^n = \mathbf{G}(\mathbf{Z}^n)$. We remark that $\Delta\mathbf{Z}^n$ satisfies a linear system of equations.

In the case of Eq. (5.4), the function $\mathbf{G}$ has the following structure,

$$G_k(\mathbf{Z}) = \sum_{i=1}^{M} R_i \; \partial I_i / \partial Z_k \,,$$

where, for simplicity, we are omitting the dependency of $\mathbf{G}$ in $\mathbf{X}$.

Show that the equation for $\Delta\mathbf{Z}^n$ reads

$$[(\mathcal{J}^n)^T \mathcal{J}^n + A^n]\Delta\mathbf{Z}^n = -(\mathcal{J}^n)^T \mathbf{R}^n \,,$$

where

$$A^n_{jk} = \sum_{i=1}^{n} R_i(\mathbf{Z}^n) \left(\partial^2 I_i / \partial Z_j \partial Z_k\right) \,.$$

(Compare this procedure with Eq. (5.8) and note that here, second order derivatives of $\mathbf{I}$ are needed in order to compute matrices A^n, which make this algorithm more expensive and more prone to numerical errors than the other.)

5.2. Deduce the Levenberg-Marquardt method considering the functional given by Eq. (3.20).

5.3. Assume that the participant medium is isotropic, i.e., $p(\mu,\mu') = c$ is a constant.

(a) Use Eq. (5.15) to determine the value of constant c.

(b) Obtain the simplified version of the Boltzmann equation, Eq. (5.14), for $I = I(\tau, \mu)$.

(c) Show that there is not a solution of the integro-differential equation for the isotropic medium, with the prescribed boundary conditions, depending only on τ, i.e, $I = I(\tau)$.

(d) Obtain a solution $I = I(\tau)$ if only the first requirement in Eq. (5.14b) is asked for.

5.4. An alternative approach for the solution of Eq. (5.14a) uses an integrating factor. Define a new dependent variable as

$$J(\tau,\mu) = e^{\frac{\tau}{\mu}} I(\tau,\mu)$$

and show that Eq. (5.14a) may be written as

$$\mu \frac{\partial J}{\partial \tau}(\tau,\mu) = \frac{\omega}{2} \int_{-1}^{1} p(\mu,\mu') e^{\tau\left(\frac{1}{\mu} - \frac{1}{\mu'}\right)} J(\tau,\mu')\, d\mu'$$

Write the boundary conditions given by Eq. (5.14b) in terms of the new dependent variable, [57].