

Analysis of parameter sensitivity and experimental design for a class of nonlinear partial differential equations

Michael L. Anderson^{1,‡}, Wolfgang Bangerth^{1,2,§} and Graham F. Carey^{1,*,†}

¹*Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin TX 78712, U.S.A.*

²*Institute for Geophysics, John A. and Katherine G. Jackson School of Geosciences, University of Texas at Austin, Austin TX 78712, U.S.A.*

SUMMARY

The purpose of this work is to analyse the parameter sensitivity problem for a class of nonlinear elliptic partial differential equations, and to show how numerical simulations can help to optimize experiments for the estimation of parameters in such equations. As a representative example we consider the Laplace–Young problem describing the free surface between two fluids in contact with the walls of a bounded domain, with the parameters being those associated with surface tension and contact. We investigate the sensitivity of the solution and associated functionals to the parameters, examining in particular under what conditions the solution is sensitive to parameter choice. From this, the important practical question of how to optimally design experiments is discussed; i.e. how to choose the shape of the domain and the type of measurements to be performed, such that a subsequent inversion of the measured data for the model parameters yields maximal accuracy in the parameters. We investigate this through numerical studies of the behaviour of the eigenvalues of the sensitivity matrix and their relation to experimental design. These studies show that the accuracy with which parameters can be identified from given measurements can be improved significantly by numerical experiments. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: parameter sensitivity; Laplace–Young equation; optimal experimental design

*Correspondence to: Graham F. Carey, Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin TX 78712, U.S.A.

†E-mail: carey@ices.utexas.edu

‡E-mail: michaela@ices.utexas.edu

§E-mail: bangerth@ices.utexas.edu

Contract/grant sponsor: Institute for Computational Engineering and Sciences

Contract/grant sponsor: Sandia National Laboratories; contract/grant number: 56522

Received 23 February 2004

Revised 13 January 2005

Accepted 14 January 2005

1. INTRODUCTION

Even if a system is qualitatively well understood, one usually needs to identify essential parameters before its behaviour can be predicted quantitatively in an accurate way. These parameters may be coefficients in a constitutive relation (for example heat conductivity), material parameters (such as density), or boundary coefficients (e.g. the wetting energy for a fluid in contact with a wall). When such properties cannot be measured directly, one can instead measure the response of the system to a known input and identify the parameters by using the theory of inverse problems, effectively tuning the parameters until the model predicts the measurements correctly. This is a computationally expensive task, but it is theoretically well understood (see, e.g. References [1, 2]).

However, the question of how to perform measurements so that parameters can be identified most accurately is not as well understood for models described by partial differential equations [3–5]. In essence, we want to set up experiments so that the subsequent inversion of the data for the values of parameters such as the material properties yields maximum accuracy; i.e. we want to maximize the sensitivity of measurements to the actual values of model or system parameters. Note that this contrasts with the usual desire to have models be rather insensitive to parameter changes when solving the forward problem, since then the results of computations do not change too much if the values of parameters are not exactly known.

In the present study we consider the questions of parameter sensitivity and optimal experimental designs. Of particular interest are models that are described by nonlinear elliptic partial equations, and, as a representative example, we take the Laplace–Young problem describing free surface behaviour of a fluid in contact with walls. First, this problem is representative of a class of elliptic boundary value problems of general interest that includes, for example, the minimal surface problem for soap films [6–8], certain subsonic potential flow approximations [9–14], and certain other problems arising in deformation of membranes, flow through porous media, eddy current electromagnetics, and elsewhere. Moreover, the Laplace–Young problem is sufficiently nonlinear to be interesting and non-trivial. For example in certain cases the solution can be singular near corners [15]. This equation also displays some of the ways that parameters typically enter equations. Consequently, the parameter optimization, sensitivity analysis and experimental design problem are interesting but remain of modest complexity. Moreover, the problem is easy to interpret and also amenable to simulation as well as practical laboratory experiment.

For the Laplace–Young equation, the two parameters of interest are a fluid property related to the surface tension, and the contact angle between fluid and wall. While there are several well established techniques to measure these parameters [16, 17], they typically work in a different manner to the inverse problem technique examined here, and do not seem to have been the subject of optimal experimental design. In the context of this equation, knowledge of the two parameters is of significant practical interest mainly in handling molten materials, such as in crystal growth or soldering applications. Also, the wetting behaviour of fluids is important in microfluidics, cleaning of surfaces for microelectronics applications, and for lubrication. An application for a related equation is the measurement of gas parameters in potential flow.

Optimal experimental design is of great practical interest, since it guides the design of experiments so that we measure the responses of the system in a configuration that allows us to identify the system parameters with the greatest possible accuracy [3, 4]. The results

of optimal experimental design can also help to avoid experiments that would yield data that is insensitive to one or more parameters, or insensitive to certain combinations of these parameters, as shown later. In practice, experimental design involves the following:

1. *Initial design*: given physical intuition and experience, design an initial experimental setup.
2. *Measurements*: perform the experiment and measure output variables.
3. *Inverse problem*: use all previously made measurements to determine most likely values of system parameters and their uncertainties.
4. *Experimental design*: assuming that the identified parameters are correct, determine a modified experimental design that would minimize the uncertainty in the parameters.
5. *Implementation*: implement this modified experiment and go back to the measurement phase in step 2.

In this paper, we will be mostly concerned with step 4, although we discuss step 3 in Section 4 as well, since it provides the basis for experimental design. It is worth noting that both steps 3 and 4 in general depend nonlinearly on the output of the previous step. Thus, the optimal experiment determined in step 4 is based on the assumption that the parameters identified in step 3 are already correct and that we only want to minimize their uncertainty. In most cases, this assumption will not hold, and the next round of measurements and solving the inverse problem will yield improved parameters which in turn will again yield a different optimal design. For the purpose of considering only one iteration of optimal experimental design, we can, however, assume the values of the parameters in the model as known and fixed. We will then use the numerical experiments discussed in this paper to calculate the sensitivity of the inverse problem to various experimental design factors, and to determine optimal values for these factors.

An outline of the paper is as follows: in the following two sections we first describe the Laplace–Young model equation and the role of the parameters, and then briefly discuss the behaviour of solutions for domains of different shapes. In Section 4, we introduce the idea of measurement operators and misfit functionals used in the inverse problem to identify the parameters in the model. The sensitivity matrix associated with the inverse problem is derived in Section 5, and the role of its eigenvalues is discussed. Following this, we present results for the eigenvalues of the sensitivity matrix as a function of the geometry of the domain in which measurements are performed, and discuss the implications on experimental design. The implications of our analysis and computations are summarized in the Conclusions. An appendix shows how we computed sensitivity matrices numerically, using adaptively refined meshes on a domain with a corner.

2. PROBLEM STATEMENT

To introduce the main concepts, let us consider models that can be described as minimizing a convex energy functional $E(u, \nabla u)$, where E also depends on certain parameters. The minimization over all admissible functions naturally leads to a (possibly nonlinear) elliptic partial differential equation. As stated in the introduction, we consider this class of equations due to its wide applicability in modelling, as well as the fact that such problems are generally well-understood from a theoretical viewpoint [18].

While this equation class covers a number of broader applications, we choose as a specific case study the Laplace–Young problem describing the height u of a fluid in a bounded domain Ω with vertical walls. For this system, surface tension, gravity, and wall contact determine the shape of the free fluid surface. These factors can be described by two relevant parameters, κ and σ . Here, $\kappa = \beta/g\rho$ is the ratio of surface energy and gravitational energy, where β is the energy per unit surface of the fluid interface (between the lower fluid and some upper fluid, typically air), g the gravitational acceleration, and ρ the density difference between the two fluids. Thus, κ may be considered a property of the fluid. On the other hand, for an interface with wall contact angle γ , $\beta\sigma = \beta \cos \gamma$ is the energy per unit length of the fluid–boundary contact and is therefore a model parameter rather than a material property. Note that while there are often small variations in σ along the boundary due to differing roughness or surface coatings of the container, we will consider it as constant for simplicity. In addition, the range of possible values of $\sigma = \cos \gamma$ is $[-1, 1]$, where we have introduced σ in favor of γ to make the equations linear in the parameters.

Accumulating respective contributions due to surface tension, gravity, and contact, the energy functional corresponding to the Laplace–Young problem is

$$E(u, \nabla u) = \int_{\Omega} \sqrt{1 + |\nabla u|^2} \, d\Omega + \int_{\Omega} \frac{1}{2} \kappa u^2 \, d\Omega - \int_{\partial\Omega} \sigma u \, ds$$

and the stationarity condition is given by the following variational statement in terms of a semilinear form $a(\kappa, \sigma; u, \varphi) = E'(u, \nabla u)(\varphi, \nabla \varphi)$: find $u \in V$ such that

$$a(\kappa, \sigma; u, \varphi) = \left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}}, \nabla \varphi \right)_{\Omega} + \kappa(u, \varphi)_{\Omega} = \sigma(1, \varphi)_{\partial\Omega} \quad \forall \varphi \in V \quad (1)$$

where $(\phi, \psi)_{\Sigma} = \int_{\Sigma} \phi \psi \, d\Sigma$, for a space V of appropriate test functions. The choice of this space depends on the geometry of the domain, as discussed in the next section. The Euler–Lagrange equation arising from the weak variational statement (1) is the quasi-linear elliptic Laplace–Young equation,

$$-\nabla \cdot \frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} + \kappa u = 0 \quad \text{in } \Omega \quad (2)$$

with natural boundary condition

$$\mathbf{n} \cdot \frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} = \sigma \quad \text{on } \partial\Omega \quad (3)$$

This problem exhibits a number of interesting features that are often associated with nonlinear equations. In particular, it has unbounded behavior at corners with small angles, corresponding to fluids climbing (for wetting fluids) or depressing (for non-wetting fluids) near the contact boundaries of the enclosing container, as well as nonlinear dependence of the solution to changes in the parameters. As we will see, this implies a varying sensitivity to parameters.

To be specific, we consider solving the Laplace–Young equation in a pie-shaped domain of radius R and with corner angle θ (Figure 1). This domain geometry may be considered as the shape of a container in which measurements on the fluid are performed. The figure also

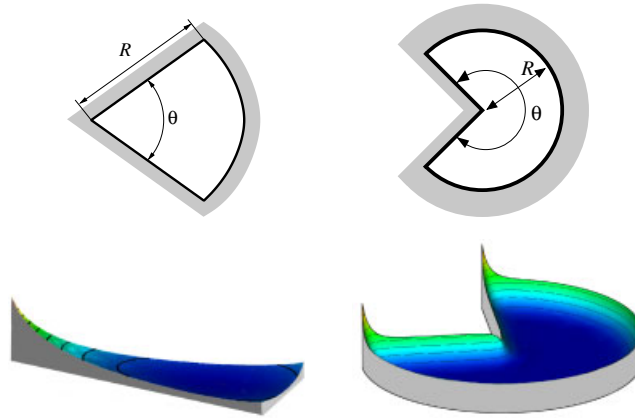


Figure 1. Top: the pie-shaped domains used for the calculations. Bottom: typical solutions computed on very fine grids showing the corner and boundary behaviour. The two plots are not at the same vertical scale, the left one growing to a much larger height.

shows surface plots of typical solutions, demonstrating the behaviour of a wetting fluid close to walls and corners of both oblique and obtuse angles.

We consider this setting to explore various ideas concerning the best choice of the quantities to be measured and the associated parameter sensitivity issues. Note, that similar equations describe different physical applications and many of the ideas to be presented below can be transferred to these settings. While different types of measurements would be chosen in those situations, the general concepts and techniques remain the same.

For predictive computations, one needs to know the exact values of the parameters κ and σ of the fluid–wall system. They can be inferred using inverse-problem methods (see, for example, References [1, 2]). In this approach, one probes various properties of the fluid surface, such as the mean elevation, the fluid height along the wedge boundary, etc., to determine these parameter values. Two relevant questions then are: (i) which shape of the domain is best suited for this parameter identification, and (ii) which measurements should be performed. Accordingly, we first examine the effect of geometry on the sensitivity of measurements, as well as the sensitivities that different types of measurements provide. In particular, we look for an optimal geometry and optimal measurements among some given set.

3. SOLUTION BEHAVIOUR

The behaviour of solutions to the Laplace–Young equation in a corner is known to depend strongly on the opening angle of the corner, θ , and the contact angle, γ , with a long experimental history of exploring this relationship [19–21] (for recent experiments, see, for example, Reference [22]). Furthermore, it has been shown that there is a critical angle, $\theta_c = \pi - 2\gamma$, such that for $\theta < \theta_c$ the solution in the corner has a singularity, and is locally planar otherwise. For example, the bottom left part of Figure 1 shows the solution for a domain with an angle

$\theta = 24^\circ$ that is very close to the critical angle $\theta_c = \pi - 2 \arccos 0.2 \approx 23.07^\circ$ of this fluid; the strong rise in the fluid surface as the corner is approached radially is clearly visible although it is bounded for the present case since $\theta > \theta_c$. Experimental results for similar cases but for a different fluid are, for example, shown in the images of Tim Coburn reproduced in Reference [23, Figure 5.6, p. 119].

To be more specific, the asymptotic behaviour of the solution near a corner is given as follows:

1. For angles $\theta < \theta_c$ the height at the corner is unbounded, having a $1/r$ type singularity [8, 15, 24].
2. For $\theta_c < \theta < \pi$ the solution is locally planar and bounded [25].
3. For re-entrant corners with $\pi < \theta < 2\pi - \theta_c$ the problem can admit locally planar solutions [26]. Korevaar [27] has constructed examples where the solution does not extend continuously to the vertex.
4. For re-entrant corners with $2\pi - \theta_c < \theta < 2\pi$ there are no locally planar solutions.

For more detailed studies of the behaviour, see in particular Reference [26]. We expect that these qualitative changes in the solution also affect the sensitivity with respect to model parameters. This issue is examined below.

4. THE INVERSE PROBLEM

In the inverse problem, one tries to use measurements of the fluid height $u(\mathbf{x})$ to infer the (previously unknown) values of the system parameters κ and σ . To this end, let us define by M_i the operator that extracts the i th measurement from the fluid height u . For simplicity we only consider linear operators M_i . For example, M_1 with

$$M_1 u = \int_{\Omega} u(\mathbf{x}) \, d\Omega \quad (4)$$

is the operator associated with measuring the volume of fluid under the surface (or the mean elevation, up to a constant), and M_2 with

$$M_2 u = u|_{\Gamma} \quad (5)$$

is the operator associated with measuring the fluid height along a contour such as a specified part Γ of the boundary $\partial\Omega$. Of course, in practice this might be a collection of point measurements along Γ . As these examples show, the range space of these measurement operators can be scalars or functions; i.e. more generally M_i are operators mapping u into another normed space. For later reference, let us also define the following two additional local measurement operators M_3, M_4 by

$$M_3 u = \int_{\Omega} h(\mathbf{x}) u(\mathbf{x}) \, d\Omega, \quad M_4 u = \int_{\Omega} h(\mathbf{x}) \nabla u(\mathbf{x}) \, d\Omega \quad (6)$$

where $h(\mathbf{x})$ is a weight function localized around the corner of the domain, which we have positioned at the origin for simplicity. It satisfies $h(\mathbf{x}) = 0$ for $\|\mathbf{x}\| > \tilde{r}$ and $h(\mathbf{x}) = \tilde{r}^{-2}$ for $\|\mathbf{x}\| \leq \tilde{r}$, with some small \tilde{r} chosen to be less than the capillary length scale. Here, M_3 and M_4 are related to the height and the gradient near the corner of the wedge.

Note that while we consider the measurement operators M_1, \dots, M_4 to be typical, they are actually close to what is being measured in practice by commercial devices to determine the surface tension of fluids; for example, the fluid volume measurement (4) is essentially, up to a constant, the force measured by both the Du Nouy ring and the Wilhelmy plate methods [16, 17]. However, as we will see below, a chosen measurement may, by itself, not be adequate to determine both parameters.

Let us denote by $z_i, i = 1, \dots, 4$, the values we would measure for the respective measurements listed above. For measurement operators M_1, M_3 , and M_4 these are single values, while for M_2 the measurements are all the values of the solution on Γ .

Now consider that instead of the physical system, we have a model with parameters that are yet to be determined. Let us assume that with each set of fixed values for the parameters, a solution $u(\kappa, \sigma)$ is associated. The inverse problem then involves finding those values κ, σ for which the predicted fluid height $u(\kappa, \sigma)$ has measurements $M_i u(\kappa, \sigma)$ that match the actual measured values z_i the best, i.e. for which the misfit, $M_i u(\kappa, \sigma) - z_i$, is smallest in some sense. If we elect to match *all* measurements collectively, this involves minimizing the composite misfit function

$$\mu(\kappa, \sigma) = \sum_i \alpha_i \mu_i(\kappa, \sigma), \quad \mu_i(\kappa, \sigma) = \frac{1}{2} \|M_i u(\kappa, \sigma) - z_i\|^2 \tag{7}$$

where α_i are weights scaling the individual misfits, μ_i , relative to each other. For simplicity, we have here taken the standard L^2 -norm minimization of the misfit, although more general functions are possible.

In the general case, $\mu(\kappa, \sigma)$ is a nonlinear function of its parameters, and minimizing it involves nonlinear optimization techniques. In addition, though not in the present case, $\mu(\kappa, \sigma)$ may have more than just one minimum and finding the global one requires judicious choice of a starting value or global optimization techniques.

In this study, we do not use actual measurements z_i , but rather generate them numerically; i.e. we choose a ‘target’ set of parameters κ^*, σ^* which describes the fluid we would like to investigate, and accordingly set our *synthetic* measurements $z_i = M_i \tilde{u}(\kappa^*, \sigma^*)$, where $\tilde{u}(\kappa^*, \sigma^*)$ is an accurate numerical solution to the unknown exact solution $u(\kappa^*, \sigma^*)$ of the Laplace–Young equation. Of course, in this ‘manufactured’ case, we know where the minimum of μ lies and that at the minimum $\mu(\kappa^*, \sigma^*) = 0$ holds. However, we do not make use of this knowledge other than to generate our ‘measurements’ and evaluate results.

In order to illustrate the concepts and difficulties of minimizing the misfit $\mu(\kappa, \sigma)$, Figure 2 shows the values of $\mu_i, i = 1, \dots, 4$ as a function of the parameters κ, σ . This data is generated by repeatedly computing the Galerkin finite element approximation to the Laplace–Young problem for a wide range of parameters, and evaluating the misfits of these predictions against the precomputed synthetic measurements. As can be seen, these functions are strongly nonlinear, and the minimum in each case lies inside a long narrow valley, complicating the minimization task. (Some of our computed datapoints in Figure 2 are very small or zero, as explained above. Given the logarithmic scale of the figures, they cannot be plotted and are thus omitted, leaving the surface plots ragged close to the troughs. This is an artifact of the finite resolution of our sampling mesh, while the surfaces $\mu_i(\kappa, \sigma)$ are actually smooth functions, a fact that is reflected by the smooth contours also shown.)

In fact, for the first function measuring the volume of the fluid, it is easy to show that the minimum of μ_1 is attained along an entire line rather than at a single point: using the state

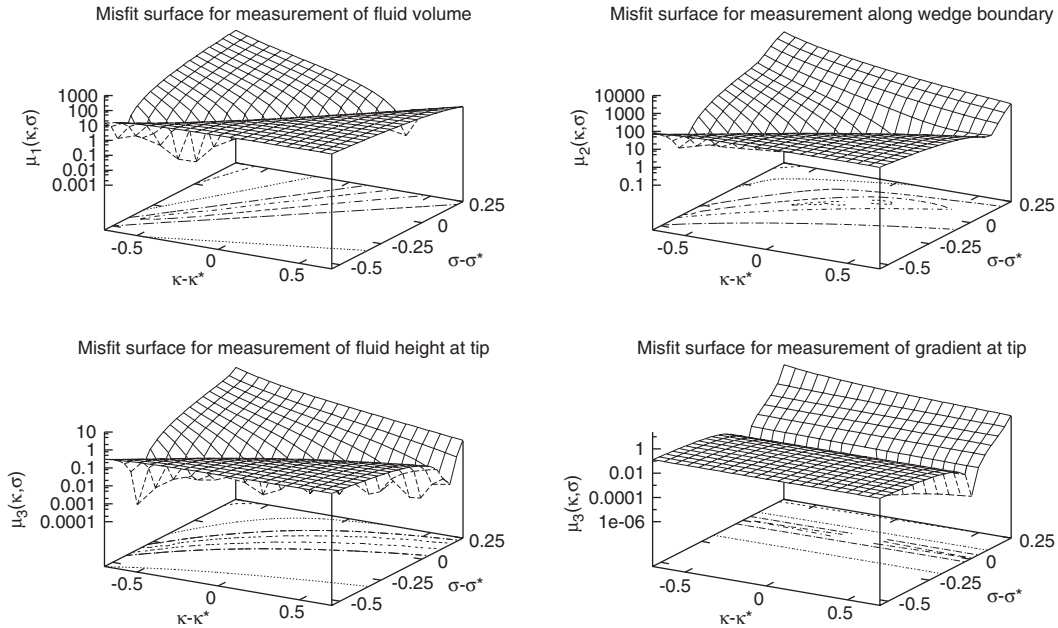


Figure 2. Surface and contour values of the misfit functions $\mu_i(\kappa, \sigma)$, $i = 1, \dots, 4$ for a domain with $R = 4, \theta = 120^\circ$ and for synthetic measurements z_i generated with $\kappa^* = 1, \sigma^* = 0.65$.

equation (2) and boundary condition (3), the volume under the fluid surface is given by

$$\begin{aligned}
 M_1 u(\kappa, \sigma) &= \int_{\Omega} u(\kappa, \sigma) = \frac{1}{\kappa} \int_{\Omega} \nabla \cdot \frac{\nabla u(\kappa, \sigma)}{\sqrt{1 + |\nabla u(\kappa, \sigma)|^2}} \\
 &= \frac{1}{\kappa} \int_{\partial\Omega} \mathbf{n} \cdot \frac{\nabla u(\kappa, \sigma)}{\sqrt{1 + |\nabla u(\kappa, \sigma)|^2}} = \frac{\sigma}{\kappa} |\partial\Omega|
 \end{aligned}$$

Thus, all combinations of parameters with fixed ratio σ/κ will yield a fluid surface with the same volume underneath. This is easily seen from the straight isolines in the left top part of the figure.

From this property, it is obvious that the single, scalar measurement of the volume is not sufficient to determine optimal values for both system parameters. This deficiency can be overcome, however, by adding other types of measurements that are sensitive to changes of parameters along the insensitive line of the first function in parameter space. Although not completely visible due to the finite resolution of the mesh we computed our results on, the second part of Figure 2 shows how a set of measurements along the wedge boundary generates a unique minimum of the corresponding function, in contrast to all other measurements that do not have an isolated minimum. In general, taking together as many measurements as possible allows for better determination of the minimum of the combined misfit surface. Ideally, one

would like to have measurements for which the troughs in the misfit surfaces are non-aligned, rather than being almost parallel as in the figure above, since in that case the sum of the functions has a more clearly marked minimum. In this respect, the fourth measurement (of the gradient at the tip of the domain) would be a good complement to the first three, a fact that would not be as obvious without computations such as the one shown.

5. SENSITIVITY

In this section, we focus on parameter sensitivity issues and how the design of the experiment (in which we generate our measurements) influences them. To introduce the concept of sensitivity, let us denote by $\hat{\kappa}, \hat{\sigma}$ the values for which the misfit function $\mu(\kappa, \sigma)$ attains its minimum. (If the measurements are generated synthetically, then of course $\hat{\kappa} = \kappa^*, \hat{\sigma} = \sigma^*$.) As mentioned in the introduction, we can take the values of the parameters as known when we compute sensitivities as they have already been identified when solving the inverse problem.

Let us assume that the M_i are differentiable operators, and that the solution $u(\kappa, \sigma)$ of the state equation depends smoothly on the values of the parameters. Note that the latter assumption is violated for parameter values for which the corner angles of the domain are close to the critical angle θ_c ; it will, however, hold for all other parameter values and we will only consider angles away from the critical angle in the examples below. Under these assumptions, we can expand the misfit function in a Taylor series around its minimum as

$$\mu(\hat{\kappa} + \delta\kappa, \hat{\sigma} + \delta\sigma) = \mu(\hat{\kappa}, \hat{\sigma}) + \frac{1}{2} \begin{pmatrix} \delta\kappa \\ \delta\sigma \end{pmatrix}^T \mathcal{S} \begin{pmatrix} \delta\kappa \\ \delta\sigma \end{pmatrix} + \text{higher order terms} \tag{8}$$

where $\mathcal{S} = \sum_i \alpha_i \mathcal{S}_i$ is the symmetric, positive definite *sensitivity matrix* composed of the *partial sensitivity matrices*

$$\mathcal{S}_i = \left(M_i \frac{\partial u(\hat{\kappa}, \hat{\sigma})}{\partial \{\kappa, \sigma\}}, M_i \frac{\partial u(\hat{\kappa}, \hat{\sigma})}{\partial \{\kappa, \sigma\}} \right) \tag{9}$$

There is no linear term since we are expanding around a minimum. As is usual in the treatment of small-residual inverse problems (see, e.g. References [2, 4]), we have omitted second order contributions to the matrices \mathcal{S}_i . Note that this omission can also be interpreted as linearizing the forward model and forming the sensitivity matrices from this linearized model. An outline of the algorithm to compute \mathcal{S}_i numerically is given in the appendix.

The importance of the sensitivity matrix lies in the following ideas: In general, no measurement is without noise. Let us assume, we have noise ε_i in the measurement z_i . The minimum of the misfit function with noise will then be of the order $\mu(\hat{\kappa}, \hat{\sigma}) + \varepsilon$ with $\varepsilon = \sum_i (\alpha_i/2) \|\varepsilon_i\|^2$. This shift in μ gives rise to perturbations $\delta\kappa, \delta\sigma$ such that, to second order in the Taylor expansion, we have

$$\frac{1}{2} \begin{pmatrix} \delta\kappa \\ \delta\sigma \end{pmatrix}^T \mathcal{S} \begin{pmatrix} \delta\kappa \\ \delta\sigma \end{pmatrix} \leq \varepsilon \tag{10}$$

For a given amount ε of random noise, the estimated parameter values may be anywhere within the region described by this relation. This is illustrated in Figure 3. The left and

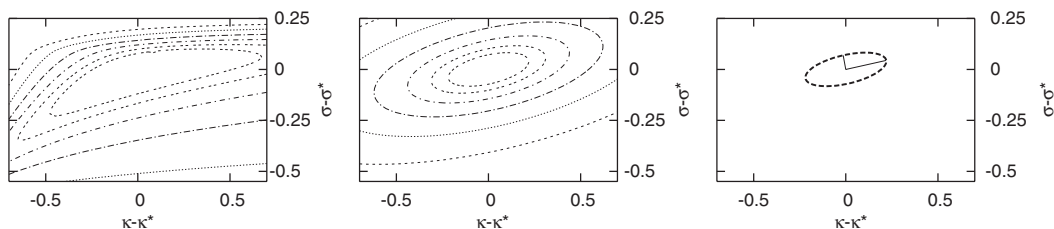


Figure 3. Left: contours of the combined misfit function $\mu(\sigma, \kappa)$. Middle: contours of its second order Taylor expansion. Right: area enclosed by (10) for a given level of measurement error ε as well as the half axes of this ellipsoid.

middle parts of the figure show contours of the original misfit function $\mu(\sigma, \kappa)$ as well as the second order Taylor expansion, respectively. The right part shows the region enclosed by solutions of (10) for a given level of measurement error ε . The half axes of this ellipsoid are given by $\sqrt{2\varepsilon/\lambda_i}$, where λ_i are the eigenvalues of the matrix \mathcal{S} . Note that if the matrix is singular, one of its eigenvalues is zero and the ellipsoid degenerates to an infinite strip in κ - σ space. This is, for example, the case for the surfaces shown in the top left, bottom left, and bottom right part of Figure 2.

Thus, the size of the elements and eigenvalues of \mathcal{S} determines the impact of small noise ε on the uncertainty $\delta\kappa, \delta\sigma$ with which we can determine the minimum $\hat{\kappa}, \hat{\sigma}$: for larger eigenvalues of \mathcal{S} , the ellipsoid in κ - σ space shown in Figure 3 is smaller, and thus a given amount of noise will have less influence on the reconstructed parameters than if the eigenvalues of \mathcal{S} are small. The eigenvalues of \mathcal{S} are therefore also called *sensitivity values*.

Obviously, large eigenvalues of \mathcal{S} are desirable. While the goal of more accurate experiments is to reduce the noise level ε , the optimal experimental design techniques discussed in this paper try to design experiments for which the eigenvalues of \mathcal{S} are larger. Both approaches reduce the size of the uncertainty ellipsoid in κ - σ space, but by complementary means.

The matrix \mathcal{S} and its eigenvalues can be given a different, statistical interpretation (see, e.g. References [2, 28]): One can assign each point (κ, σ) in parameter space a *probability density*

$$\chi(\kappa, \sigma) = C^{-1} \exp[-\mu(\sigma, \kappa)]$$

that these are the correct parameters. Here, $C = \int \int \exp[-\mu(\sigma, \kappa)] d\kappa d\sigma$ is a constant used to normalize the probability χ . Obviously, the minima of μ are the places to which we assign the highest probability. The function χ , in a sense, is a more complete description of the solution of the inverse problem than the minimum $\hat{\kappa}, \hat{\sigma}$ of μ alone.

The probability density χ can be approximated locally by a Gaussian. An elementary calculation shows that this approximation is given by

$$\chi(\hat{\kappa} + \delta\kappa, \hat{\sigma} + \delta\sigma) = \chi(\hat{\kappa}, \hat{\sigma}) \exp \left[-\frac{1}{2} \begin{pmatrix} \delta\kappa \\ \delta\sigma \end{pmatrix}^T \mathcal{S} \begin{pmatrix} \delta\kappa \\ \delta\sigma \end{pmatrix} \right] + \text{higher order terms}$$

Obviously, that part of the domain in which the probability density is more than a certain fraction of its maximum value is again given by the solutions of inequality (10) and its half widths are described by the eigenvalues of the sensitivity matrix \mathcal{S} as shown above.

6. RESULTS

In this section, we investigate how the choice of domain and measurement operator affects the sensitivity of the inverse problem. This is usually called *experimental design*. The goal here is to find an experimental setup that maximizes our ability to identify the correct values of the parameters. For example, as shown above, measuring the volume under the fluid is not suited to identifying both parameters, and other measurements are necessary to complement the identification process. On the other hand, given the nonlinear dependence of the fluid height in the corner on the angle of the pie-shaped domain, it may be that measuring with a certain wedge angle yields results that are more sensitive than measuring in domains with other values of the angle, and thus, for a given measurement error level, yields a higher accuracy for reconstructed parameters.

For clarity of exposition, we keep the description in the context of the Laplace–Young equation. However, it is important to emphasize that phenomena like degenerating identifiability of parameters for certain measurement operators, or sensitivities that depend on domain geometries and experimental setups are not unique to this particular equation. Rather, these topics are generic to all inverse problems, and the type of computational experiments shown below are a useful tool in determining parameters optimally for all parameter identification problems.

As shown in the previous section, the accuracy with which parameters can be identified is closely tied to the size of the eigenvalues of the sensitivity matrix. Therefore, in this section we first introduce a formal measure of the amount of accuracy a given experimental setup will yield, and then proceed to show how it depends on experimental design parameters such as the wedge angle and radius of the domain in which we measure, as well as the relative factors with which we weigh different experiments. The formal measure introduced before will then allow us to rank different experiments and determine which is better suited to identify parameters with the highest accuracy.

The results shown in this section were obtained by repeatedly applying the solution procedure outlined in Appendix A to obtain the sensitivity matrix, on a sequence of meshes with varying domain characteristics. All computations were performed with a program based on the Open Source finite element library deal.II [29, 30].

6.1. Eigenvalue dependence

Let us first consider how the two eigenvalues λ_i, Λ_i of the 2×2 partial sensitivity matrices \mathcal{S}_i depend on the shape of the domain in which experiments are performed. In particular, we consider the effect of varying the wedge angle θ of the pie-shaped domain in Figure 1. As explained above, larger eigenvalues are more beneficial to the identification process, since they lead to smaller uncertainties $\delta\kappa, \delta\sigma$.

Note, however, that the actual sizes of the eigenvalues are not comparable. This is due to the fact that different measurements come with different physical units and need to be

made dimensionless first. Secondly, the full sensitivity matrix $\mathcal{S} = \sum_i \alpha_i \mathcal{S}_i$ will be a linear combination of the partial matrices, weighted with the values α_i . Scaling arguments suggest that these weights should be chosen so that the weighted noise levels $\alpha_i \varepsilon_i$ are approximately of the same size. For convenience, let us assume that the noise level for each experiment is a fixed fraction η of the measured value, i.e. $\varepsilon_i = \eta \mu_i(z_i) = \eta \mu_i(M_i u(\kappa^*, \sigma^*))$; and that the noise in the different measurements is independent and uncorrelated. This corresponds to a diagonal covariance matrix for the errors with diagonal entries ε_i . Such a choice is also clearly appropriate since we are free to explore specific sensitivity ‘patterns’ in our numerical experiments. Then, the scaling

$$\alpha_i = \frac{1}{N} \frac{1}{\mu_i(M_i u(\kappa^*, \sigma^*))} \quad (11)$$

satisfies the condition that all noise levels are equal (as would any other fixed multiple of these values), where N is the number of measurements available. Here, we take the $N=4$ measurements defined in Section 4. With this particular scaling, we obtain $\varepsilon = \sum_i \alpha_i \varepsilon_i = \eta$. In our experiments we therefore also record the values of $\mu_i(M_i u(\kappa^*, \sigma^*))$ and in the following comparisons, we always consider the scaled partial sensitivity matrices $\alpha_i \mathcal{S}_i$. The results of using different scaling weights α_i are given later.

Figure 4 then shows the dependence of the eigenvalues of the scaled matrices $\alpha_i \mathcal{S}_i$ associated with the operators M_i , on the wedge angle of the domain, for a fluid with $\kappa = 1, \sigma = 0.65$. The dashed vertical line indicates the value of the critical angle $\theta_c = 81^\circ$ for this set of parameters; the model is strictly valid only to the right of this line, and results to the left would likely be meaningless due to unboundedness of the ‘real’ solution near the tip of the wedge.

We remind the reader that larger eigenvalues imply a smaller uncertainty ellipsoid. For the first measurement (the mean elevation), the smaller eigenvalue is identically zero, a fact that should not be surprising given that there are infinitely many combinations of κ and σ that yield the same fluid volume; a zero eigenvalue only indicates that the corresponding measurement is not able to distinguish between any of the parameter values that lie in the direction of the eigenvector associated with this eigenvalue, and the uncertainty ellipsoid degenerates to an infinite strip in this case. The same conclusion holds for measuring value and gradient at the tip of the wedge, since both are scalar measurements that are not able to identify both parameters at the same time. The symmetry of the domain implies that the gradient at the tip of the domain has really only one degree of freedom.

The only measurement (of those considered here) that can identify both parameters at the same time is the measurement along the wedge boundary. This is reflected by the fact that it has two associated non-zero eigenvalues. Note, however, that even though individual measurements may separately imply a zero eigenvalue, taking different measurements together may yield a matrix which has two non-zero eigenvalues, in which case we *can* identify the correct parameters. (The steps in the curve for the smaller eigenvalue are artifacts of our numerical procedure, and are explained in the appendix; the correct curve would be smooth, but within the same range of values.)

The dependence of the eigenvalues on the wedge angle is also of interest when we consider the question of designing an experiment to best identify parameters. As can be seen, some of the measurements become more sensitive to the parameters (i.e. their eigenvalues grow) as we increase the wedge angle, while the last one, M_4 , becomes less sensitive. We will investigate this in more detail in the next section.

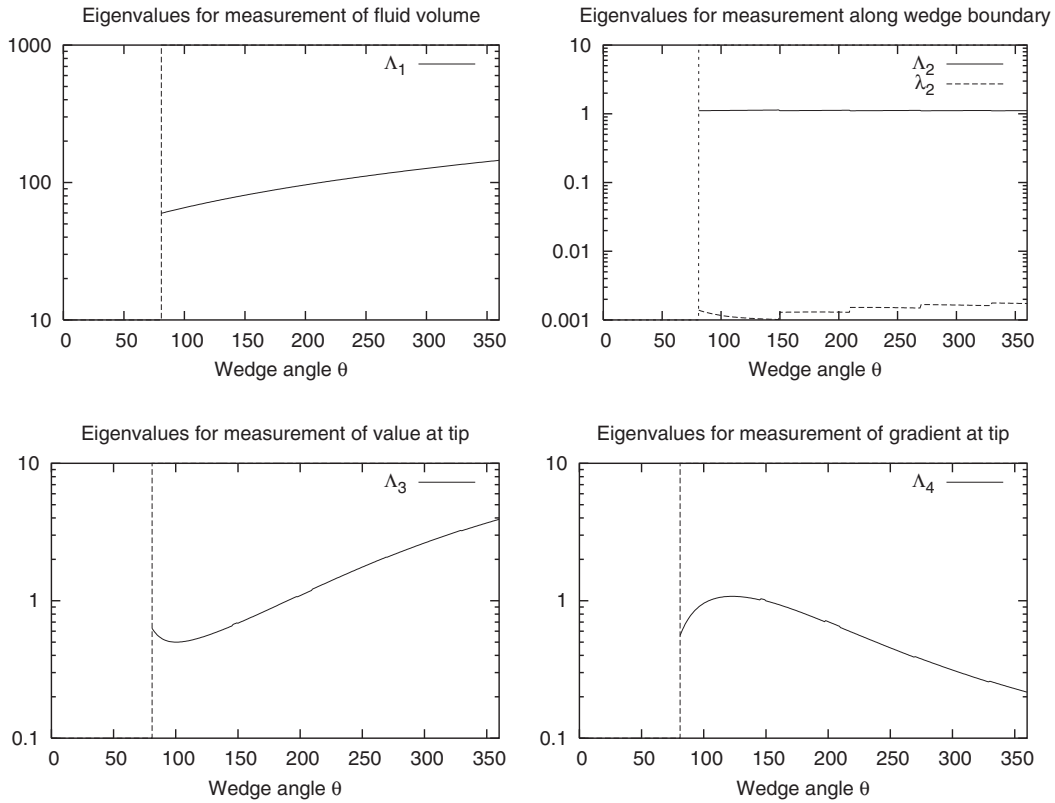


Figure 4. Dependence of the two eigenvalues λ_i, Λ_i of the scaled partial sensitivity matrices $\alpha_i \mathcal{S}_i$ on the wedge angle θ of the domain.

6.2. Measures of sensitivity

An important goal of studies like the present one is to determine the experimental setup that minimizes uncertainty in the parameters; i.e. that maximizes sensitivities [3,4]. To this end, we need to consider the sensitivity matrix \mathcal{S} as a function of experimental design variables, such as the wedge angle of the domain, or the radius.

Unfortunately, there is no single answer to the task ‘Minimize uncertainty’ (or: ‘Maximize sensitivity’), since we need to specify what we mean by ‘uncertainty’ when we are given a sensitivity matrix. The problem is of course related to the fact that there is no natural ordering on the set of symmetric positive definite matrices. Put differently: if, given a certain noise level ε , the parameter estimates may lie anywhere in the region specified by (10) and shown in the right part of Figure 3, what aspect of this ellipsoid do we want to minimize in order to reduce uncertainty?

One possible partial ordering is to consider the volume of the ellipsoid in κ - σ space in which solutions may lie. Its volume is a multiple of the product of the semi-axes of the ellipsoid, which in turn are multiples of the inverse square roots of the eigenvalues

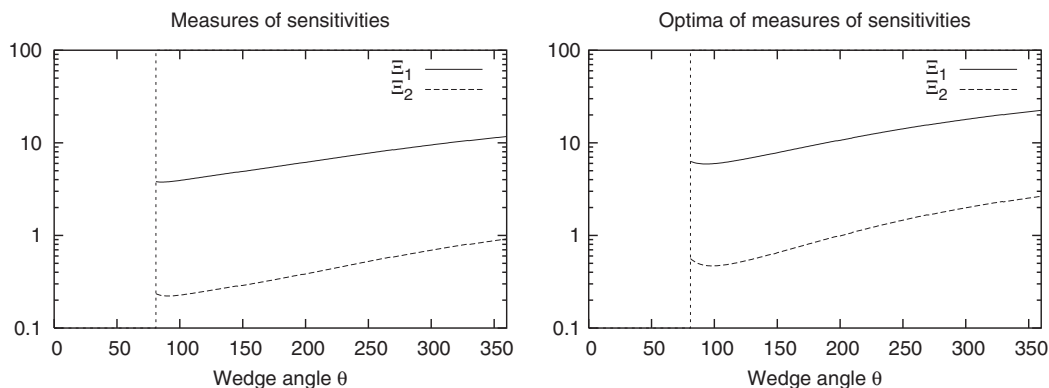


Figure 5. Measures of sensitivities computed from the sensitivity matrix when changing the angle of the domain. Left: with relative scaling (11) of the individual measurements. Right: with the sensitivity-optimizing scaling (12).

of \mathcal{S} . In this case, minimizing the volume of the ellipsoid is equivalent to maximizing the quantity

$$\Xi_1(\mathcal{S}) = \left(\prod_j \lambda_j(\mathcal{S}) \right)^{1/2}$$

where the eigenvalues $\lambda_j(\mathcal{S})$ are functions of the design variables (here the wedge angle θ), and so is then $\Xi_1(\mathcal{S})$.

Another possibility is to minimize the uncertainty in any of the eigenvector directions spanning the ellipsoid. This is equivalent to maximizing the minimal eigenvalue of \mathcal{S} ,

$$\Xi_2(\mathcal{S}) = \min_j \lambda_j(\mathcal{S})$$

Maximality of either Ξ_1 or Ξ_2 is usually referred to as the D - and E -optimality criterion of experimental design, see References [3–5, 31].

Depending on the circumstances, maximizing either Ξ_1 or Ξ_2 may be valid goals of experimental design optimization. Both quantities are plotted in the left panel of Figure 5. As can be seen, the sensitivity increases in both measures as we increase the wedge angle. The conclusion is that *the best domain (choice of θ) for the measurement is that with wedge angle close to 360° ; i.e. a circular domain that excludes a slender radial sector of vanishing small angle*. If one were to do a measurement on such a domain, one would obtain an increase in Ξ_k of a factor of 3–4, and thus parameters with an uncertainty that is approximately a factor of $\sqrt{3}$ to $\sqrt{4}$ smaller than if the domain had been chosen to have a wedge angle around 80° , in which measurements happen to yield the least accuracy. This is particularly surprising, since at first glance one might think that, close to the critical angle, measurements would be most sensitive due to the singular behaviour of the fluid height close to the tip of the wedge.

The sensitivity can also be increased by varying the relative scaling α_i of the individual measurements. One could, for example, increase the relative weight of a very sensitive

measurement with respect to the weight of an insensitive measurement. We do so by replacing the weighting (11) by the following weighting:

$$\alpha_i(\beta_i) = \frac{1}{N} \frac{\beta_i}{\mu_i(M_i u(\kappa^*, \sigma^*))} \tag{12}$$

with $\sum_i \beta_i = N, 0 \leq \beta_i \leq N$. Obviously, the original choice of weights is recovered with $\beta_i = 1$. Note that this scaling leaves the total noise level invariant: $\varepsilon = \sum_i \alpha_i \eta \mu_i(z_i) = \eta 1/N \sum_i \beta_i = \eta$. We are then interested in computing the measures of sensitivity $\Xi_k(S), k = 1, 2$, of the differently weighted sensitivity matrices $S = \sum_i \alpha_i \mathcal{S}_i$, as a function of the new weights β_i and choosing that relative weighting β_i that maximizes them:

$$\hat{\Xi}_k = \max_{\beta} \Xi_k \left(\sum_i \alpha_i(\beta_i) \mathcal{S}_i \right)$$

We obviously have that $\hat{\Xi}_k = \hat{\Xi}_k(\theta)$ is at least as large as the original sensitivity measure $\Xi_k(\theta)$ obtained with the scaling $\beta_i = 1$. Since we have only four variables β_i and since their values are confined to the interval $[0, N]$, we search for this optimum by direct, exhaustive search over all grid points of the lattice with spacing $N/100$. For problems with more measurements, it would be straightforward to solve for the best vector β using any standard optimization algorithm.

The optimal sensitivity measures are shown in the right part of Figure 5. For $\theta = 360^\circ$, the optimal weights are $\beta = (2, 0, 2, 0)$; i.e. the measurements along the boundary Γ and of the gradient at the tip are not used at all. As is apparent from the graphics, the sensitivity can be increased by another factor of 2 to 3 by selecting this better relative scaling. Such an increase in Ξ_k in turn corresponds to a reduction in uncertainty in each of the parameters by a factor of $\sqrt{2}$ to $\sqrt{3}$; i.e. an improvement of more than 30%.

A deliberate choice of both the domain in which measurements are performed, as well as of the relative weights attached to each of the measurements leads to a significant increase in the sensitivity of our experiment with respect to the parameters, and will therefore significantly decrease the uncertainty in their values.

6.3. Variation in the radius of the domain

Above, we have considered how varying the wedge angle and the relative weights of measurements affect sensitivities. Let us now investigate the effects of variations in the radius of the domain. Figures 6 and 7, respectively, show the dependence of the eigenvalues of the scaled partial sensitivity matrices and the measures of sensitivity derived from them, as a function of the domain radius for a pie-shaped domain with a fixed angle of 270° . From Figure 7, it is obvious that *measurements at very small radii are not useful, since the smaller eigenvalue goes to zero*. This again somewhat contradicts the intuition that the solution should be more sensitive to parameters if the boundary effects are more important due to the smallness of the domain.

On the other hand, the results of Figure 7 also show that beyond $R \approx 5$ one of the eigenvalues of the composite sensitivity matrix does not increase any more, while the other one does. This implies that the ellipsoid spanned by solutions of (10) contracts in one of its principal directions but remains fixed in the other direction.

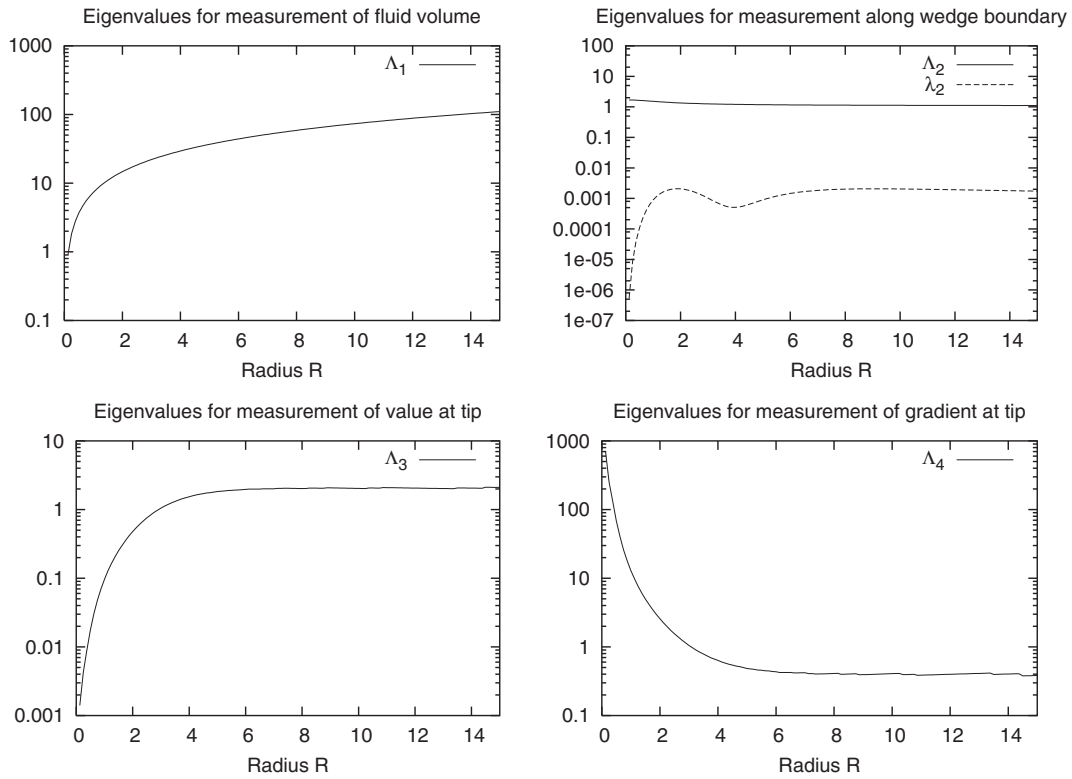


Figure 6. Dependence of the two eigenvalues λ_i, Λ_i of the scaled partial sensitivity matrices $\alpha_i \mathcal{S}_i$ on the radius R of the domain.

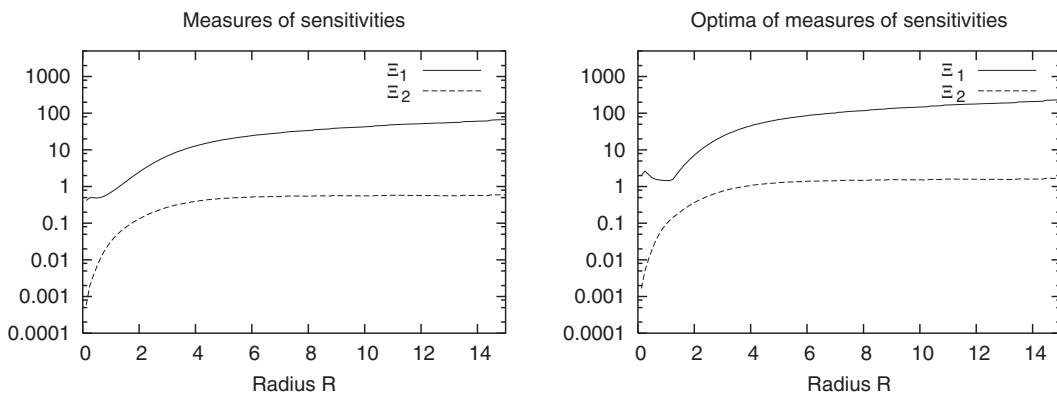


Figure 7. Measures of sensitivities computed from the sensitivity matrix when changing the radius of the domain. Left: with relative scaling (11) of the individual measurements. Right: with the sensitivity-optimizing scaling (12).

6.4. Variations in the fluid parameters

As mentioned in the introduction, the improved design for an experiment obtained with the techniques discussed in this paper depends on knowledge of the values of the parameters κ, σ . We have assumed that we know (approximate values of) these parameters from solving the inverse problem with previously available measurement data. However, sometimes a poorly chosen experiment will leave us with bad estimates of parameter values. The result of experimental design is then based on a wrong choice of linearization point. To see this, note that the sensitivity matrix \mathcal{S} depends on the parameters via Equation (9) (see also Equations (A4) and (A6) in the appendix). In this section, we investigate what kind of design we would get if we started from a significantly different initial guess. If the resulting optimal design is basically independent of the linearization point, then chances are good that the next round of measurements will result in a much better parameter estimate; if the optimal design is significantly different, then no such guarantee can be given.

In order to investigate this issue, we undertook numerical experiments for different fluid parameters. Computations for $\kappa = 1.0, \sigma = 0.2$ gave results very similar to those shown previously for $\kappa = 1.0, \sigma = 0.65$ in Figures 4 and 5. However, results for $\kappa = 0.2, \sigma = 0.2$ are different and graphed in Figures 8 and 9. In particular, note that the dependence of the sensitivities on

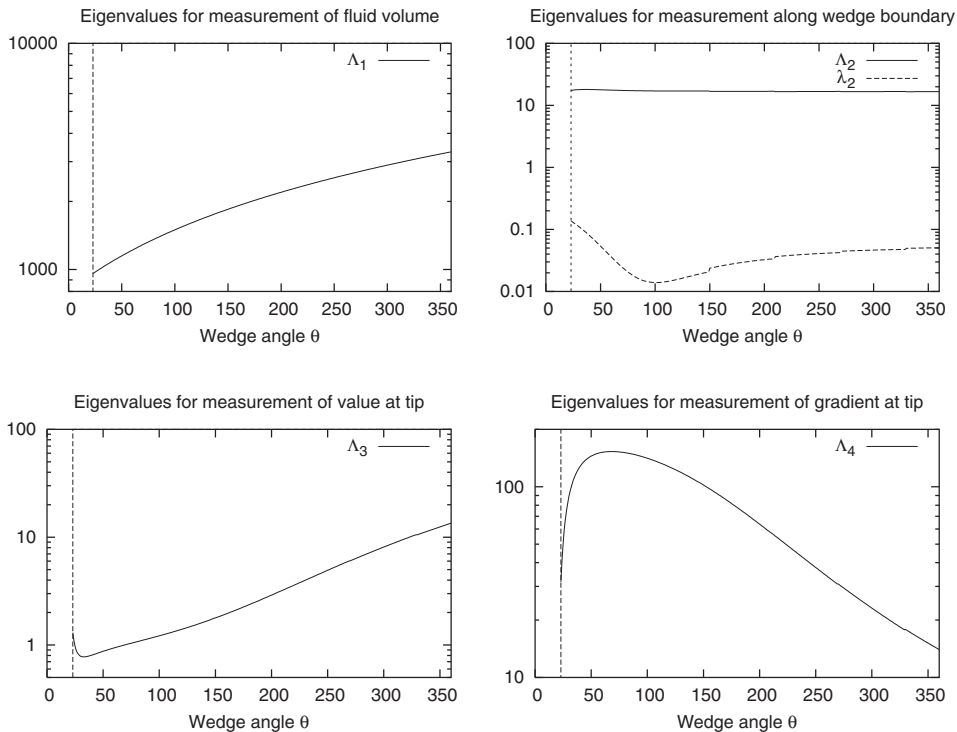


Figure 8. As in Figure 4 but for modified values $\kappa = 0.2, \sigma = 0.2$ of the fluid parameters: dependence of the two eigenvalues λ_i, Λ_i of the scaled partial sensitivity matrices $\alpha_i \mathcal{S}_i$ on the wedge angle θ of the domain.

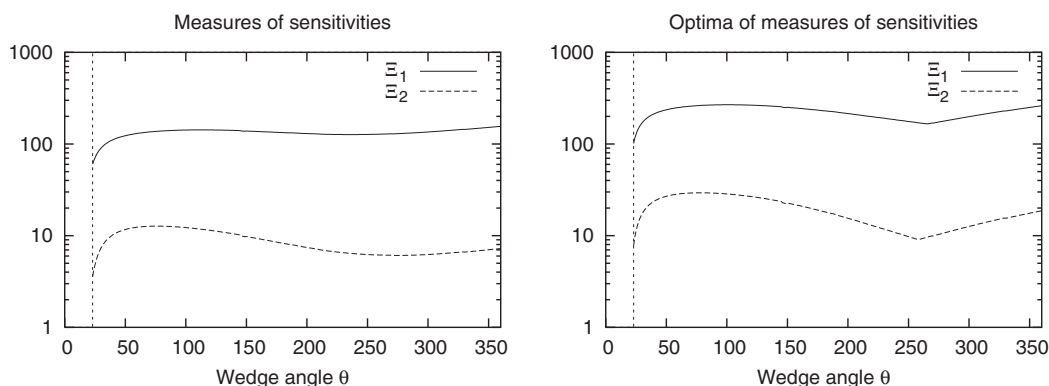


Figure 9. As in Figure 5 but for modified values $\kappa=0.2, \sigma=0.2$ of the fluid parameters: measures of sensitivities computed from the sensitivity matrix when changing the angle of the domain. Left: with relative scaling (11) of the individual measurements. Right: with the sensitivity-optimizing scaling (12).

the wedge angle is almost flat (see the left part of Figure 9), although the eigenvalues of the sensitivity matrices do indeed change. It is also noteworthy that the sensitivities for this set of parameters are significantly larger than for the previous set. An accurate identification of parameter values would thus be much simpler here.

The right part of Figure 9 shows the optimal sensitivities after optimizing for the weight factors β_i in (12). Again, some improvements are possible by properly adjusting the weights. In particular, the kink in the curves is due to the fact that for wedge angles less than 266° , the optimal weights are $\beta = (2, 0, 0, 2)$, while for angles greater than 266° the optimal weights are again $\beta = (2, 0, 2, 0)$. In other words, for smaller angles one may measure fluid volume and the gradient of the solution at the tip of the domain, while for angles larger than that the measurement of the gradient at the tip should be replaced by the measurement of the height there.

The general conclusion concerning which measurements are more or less sensitive, and for which wedge angles measurements should be done, are unchanged for both fluids, however: measuring at angles close to 360° yields the highest accuracy. Thus, even if we have only poor *a priori* information about the range in which parameters may lie, computational experiments, such as those carried out here, can determine a good experimental design for this particular equation.

7. CONCLUSIONS

Parameters enter models in various ways, and knowing the exact values of these parameters is often critical for quantitatively accurate predictions. In the context of models described by partial differential equations, the process of parameter identification is usually referred to as the *inverse problem*. We have shown in this work how numerical simulations can help guide the design of experiments in which we obtain the measurements used in the inverse problem.

In particular, we consider a class of systems that are usually characterized mathematically by nonlinear elliptic partial differential equations with system parameters. This problem class includes the Laplace–Young equation describing the interface between two fluids bounded by a container with vertical walls, which we take as a representative example for our numerical studies. The Laplace–Young equation has two parameters which are related to the surface tension and the wall contact angle, respectively.

In the context of inverse problems, the sensitivity of the solution or of quantities derived from it, to changes in these parameters is of particular interest. We are thus led to a sensitivity analysis, which is the main ingredient to evaluate experimental design issues used to optimize experimental setups and the type of measurements to be performed so that parameters can be identified with best accuracy. In other words, we wish to select an experimental design that minimizes uncertainty in the parameters. To meet this goal, we show that we need to know the eigenvalues of the sensitivity matrix. We explore this idea for a test problem in a wedge-shaped domain, and demonstrate, not surprisingly, that the two parameters occurring in the Laplace–Young equation cannot be identified uniquely when the measurement is the volume of the fluid or any other scalar quantity alone. This non-uniqueness corresponds to the presence of a zero eigenvalue of the partial sensitivity matrix. However, both parameters are identifiable if several different measurements are taken together.

In each case studied, we graph the dependence of the sensitivities as a function of the design parameters describing the domain, i.e. the wedge angle and the radius. We also consider the dependence on the relative weighting of the individual measurements. One of the interesting results is that the best domain for measurement corresponds to the wedge angle approaching 360° . This result is counterintuitive given the behaviour of the solution for this domain shape, as compared to the behaviour for other angles and especially near the critical angle. Similarly, we have shown that it is better to perform measurements in larger, rather than smaller domains, since this decreases the uncertainty in the parameters.

In addition, our results indicate that the conclusions drawn from these computations also hold if we have only poor prior estimates of the parameters. In this case, one can compute the sensitivities for the present best estimate of these parameters using the techniques shown, and thereby guide the design of an optimal experiment to obtain further measurements. These can in turn be used to improve our knowledge of parameter values. If desired, the process can then be repeated with the better values of the parameters.

In our computations, we have shown that by choosing an optimal shape of the domain we can reduce the uncertainty in the parameters by up to a factor of 2, and that a proper relative weighting of our measurements can improve this by another factor of about 1.7. If experiments are complicated or expensive to perform, such computations are therefore able to significantly improve the process of parameter identification by predicting under which conditions experiments should be conducted to obtain maximal accuracy. It is clear that this reasoning and these types of results are equally valid for a much wider class of problems than the representative case considered here.

APPENDIX A: COMPUTING THE SENSITIVITY MATRIX

Given definition (9) of the partial sensitivity matrices \mathcal{S}_i , their computation involves calculating how the solution $u(\kappa, \sigma)$ changes as a function of the parameters in the vicinity of the

minimum $\hat{\kappa}, \hat{\sigma}$. To explore this further, first note that for given values of the parameters, the solution $\hat{u} = u(\hat{\kappa}, \hat{\sigma})$ has to satisfy

$$a(\hat{\kappa}, \hat{\sigma}; \hat{u}, \varphi) = \hat{\sigma}(1, \varphi)_{\partial\Omega} \quad \forall \varphi \in V \quad (\text{A1})$$

for some appropriate space of test functions V . Treating u as a dependent variable and taking derivatives with respect to κ and σ , we obtain

$$a_{\kappa}(\hat{\kappa}, \hat{\sigma}; \hat{u}, \varphi) + a_u(\hat{\kappa}, \hat{\sigma}; \hat{u}, \varphi) \left(\frac{\partial u}{\partial \kappa} \right) = 0 \quad \forall \varphi \in V \quad (\text{A2})$$

$$a_{\sigma}(\hat{\kappa}, \hat{\sigma}; \hat{u}, \varphi) + a_u(\hat{\kappa}, \hat{\sigma}; \hat{u}, \varphi) \left(\frac{\partial u}{\partial \sigma} \right) = (1, \varphi)_{\Omega} \quad \forall \varphi \in V \quad (\text{A3})$$

where a subscript denotes a derivative of the semilinear form a with respect to the indicated argument.

In general, the exact solutions of Equations (A2)–(A3) are not known. In order to compute \mathcal{S}_i , we therefore first numerically approximate \hat{u} and then approximate the linearized operator as well as the right hand side. Using standard finite elements [32] with shape functions ψ_m , and denoting by \mathbf{y}_{κ} the vector of nodal values of a discrete version of the variation $\partial u / \partial \kappa$, we obtain the linear system

$$\mathcal{A} \mathbf{y}_{\kappa} = -\mathbf{c}_{\kappa}$$

where

$$\mathcal{A}_{mn} = a_u(\hat{\kappa}, \hat{\sigma}; \hat{u}_h, \psi_m)(\psi_n), \quad (\mathbf{c}_{\kappa})_m = a_{\kappa}(\hat{\kappa}, \hat{\sigma}; \hat{u}_h, \psi_m) \quad (\text{A4})$$

Likewise, we define $(\mathbf{c}_{\sigma})_m = a_{\sigma}(\hat{\kappa}, \hat{\sigma}; \hat{u}_h, \psi_m) - (1, \psi_m)_{\Omega} = -(1, \psi_m)_{\Omega}$; the numerical approximation to the solution $\partial u / \partial \sigma$ of Equation (A3) then has a nodal expansion \mathbf{y}_{σ} satisfying $\mathcal{A} \mathbf{y}_{\sigma} = -\mathbf{c}_{\sigma}$.

Let us combine these two solution vectors to the matrix $\mathcal{C} = [\mathbf{c}_{\kappa}, \mathbf{c}_{\sigma}]$, and define the measurement matrices, \mathcal{W}_i , as

$$(\mathcal{W}_i)_{mn} = \langle M_i \psi_m, M_i \psi_n \rangle \quad (\text{A5})$$

with $\langle \cdot, \cdot \rangle$ the inner product on the image space of operator M_i . For example, the first two of the operators M_i defined in Section 4 induce the matrices $(\mathcal{W}_1)_{mn} = (\int_{\Omega} \psi_m \, d\Omega)(\int_{\Omega} \psi_n \, d\Omega)$, and $(\mathcal{W}_2)_{mn} = \int_{\Gamma} \psi_m \psi_n \, d\Gamma$.

With these definitions, we obtain the numerical approximation for the partial sensitivity matrices

$$(\mathcal{S}_i)_h = \mathcal{C}^T \mathcal{A}^{-T} \mathcal{W}_i \mathcal{A}^{-1} \mathcal{C} \quad (\text{A6})$$

and for the full sensitivity matrix

$$\mathcal{S}_h = \mathcal{C}^T \mathcal{A}^{-T} \left(\sum_i \alpha_i \mathcal{W}_i \right) \mathcal{A}^{-1} \mathcal{C} \quad (\text{A7})$$

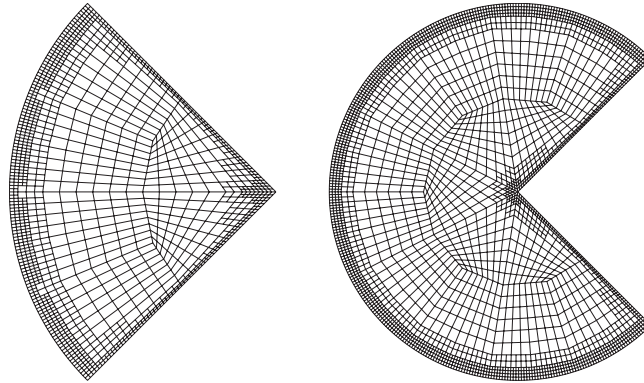


Figure A1. Typical domains and adaptively refined meshes used in the computations. Solutions on meshes similar to these but for different opening angles are shown in Figure 1.

In general, \mathcal{A} is large and sparse, and we cannot represent its inverse exactly. However, we can form $\mathcal{A}^{-1}\mathcal{C}$ for each column of \mathcal{C} separately. Since we have only two parameters, \mathcal{C} has only two columns so that this is a manageable task.

In summary, the algorithm for computing \mathcal{S}_h is as follows:

1. For given values of $\hat{\kappa}, \hat{\sigma}$, compute a finite element approximation \hat{u}_h to the solution \hat{u} of the Laplace–Young equation (A1).
2. Compute the matrices \mathcal{A} and the vectors $\mathbf{c}_p, p \in (\kappa, \sigma)$ using (A4); compute \mathcal{W}_i using (A5).
3. For each $p \in (\kappa, \sigma)$, compute the solution \mathbf{y}_p of $\mathcal{A}\mathbf{y}_p = -\mathbf{c}_p$, using a conjugate gradient method. Let \mathcal{Y} be the matrix with columns \mathbf{y}_p .
4. Compute the partial sensitivity matrices $\mathcal{S}_{h,i} = \mathcal{Y}^T \mathcal{W}_i \mathcal{Y}$ by evaluating $(\mathcal{S}_{h,i})_{pq} = \mathbf{y}_p^T \mathcal{W}_i \mathbf{y}_q$. Form $\mathcal{S}_h = \sum_i \alpha_i \mathcal{S}_{h,i}$.

For the computations in this study, we approximated \hat{u}_h using simple piecewise biquadratic finite elements on quadrilaterals, as provided by the `dea1.II` library [29, 30]. Computations were performed on meshes obtained from an initial coarse mesh of three elements by two global and three adaptive mesh refinement steps; the smallest cells therefore had a radial extent of $2^{-6}R$. Local refinement was driven by a simple smoothness criterion due to Kelly *et al.* [33]. We refined the 30% of cells with the largest errors. Typical adaptively refined meshes for two domains are shown in Figure A1. Solutions of the Laplace–Young equation on such wedge shaped domains are shown in Figure 1.

This approach ensures that a highly accurate approximation \hat{u}_h to \hat{u} is determined, and likewise for the approximations of the sensitivity factors $\partial\hat{u}/\partial\kappa, \partial\hat{u}/\partial\sigma$. It therefore implies that the experimental design goal based on the numerical approximation \mathcal{S}_h instead of the exact \mathcal{S} is not ‘masked’ by discretization error. We confirmed that the choice of a finer mesh did not have a significant effect on the results and that numerical error was therefore negligible. The only artifact of numerical errors in our computations can be seen in the smaller eigenvalues of the top right part of Figure 4: for wedge angles between 30° and 90° , we use as a coarse mesh a wedge-shaped domain composed of three quadrilaterals that are then scaled

to the correct wedge angle; for angles between 90° and 150° we scale two such basic blocks of three quadrilaterals each (see, for example, the mesh shown at the left of Figure A1); and similar for meshes between 150° and 210° , between 210° and 270° , between 270° and 330° , and beyond that. While this change in the coarse mesh at angles $90^\circ, 150^\circ, \dots$ has only a very small impact on the solution, it can be seen as an artifact in the small eigenvalue of \mathcal{S}_2 , which is some three orders of magnitude smaller than the larger one and thus more susceptible to numerical errors. These inaccuracies, however, do not affect the conclusions we draw.

As a final point of numerical interest, we remark that in the nonlinear solution of step 1 above (to be performed after each refinement step), we employed a scheme comprising five successive approximation iterations followed by 12 Newton steps, in order to combine the stability of the former with the convergence rate of the latter method. This is sufficient to solve the discretized equations to machine precision, since the Laplace–Young equation is very similar to the Helmholtz equation that can be solved with only one Newton step (since it is linear), and the additional iterations only resolve the nonlinear behaviour close to the boundary where behaviour of the solutions to the two equations differ significantly.

ACKNOWLEDGEMENTS

This study was supported in part by Postdoctoral Research Fellowships to M. Anderson and W. Bangerth from the Institute for Computational Engineering and Sciences, and by Grant 56522 from Sandia National Laboratories. The authors want to thank two anonymous referees for their careful and constructive reviews.

REFERENCES

1. Beck JV, Arnold KJ. *Parameter Estimation in Engineering and Science*. Wiley: New York, 1977.
2. Tarantola A. *Inverse Problem Theory*. Elsevier: Amsterdam, New York, 1987.
3. Atkinson AC, Donev AN. *Optimum Experimental Design*. Clarendon Press: Oxford, 1992.
4. Pázman A. *Foundations of Optimum Experimental Design*. D. Reidel Publishing Company: Dordrecht, 1986.
5. Pukelsheim F. *Optimal Design of Experiments*. Wiley: New York, 1993.
6. Carey GF. On minimal surfaces, membranes and compressible flows. In *Proceedings of the International Conference on Finite Element Methods in Engineering*, Kabaila A (ed.). Clarendon Press: Oxford, 1979; 1–17.
7. Carrier GF, Pearson CE. *Partial Differential Equations* (2nd edn). Academic Press Inc.: Boston, MA, 1988.
8. Concus P, Finn R. On a class of capillarity surfaces. *Journal d'Analyse Mathématique* 1970; **23**:65–70.
9. Bers L. Mathematical aspects of subsonic and transonic gas dynamics. *Surveys in Applied Mathematics*, vol. 3. Wiley: New York, 1958.
10. Carey GF. Variational principles for the transonic airfoil problem. *Computer Methods in Applied Mechanics and Engineering* 1978; **13**(2):129–140.
11. Chaplygin SA. On gas jets. *Scientific Memoirs of the Moscow University in Mathematical Physics* 1902; **21**:1–121 (Trans. NACA Tech. Mem. 1063, 1944).
12. Chow S-S, Carey GF. Finite element error estimates for subsonic flow. *Journal of the Australian Mathematical Society, Series B* 1987; **29**(1):88–102.
13. Dinh H, Carey GF. Approximate analysis of regularized compressible flow using the fictitious gas concept. *Nonlinear Analysis* 1987; **11**(8):861–875.
14. Morawetz CS. On the non-existence of continuous transonic flows past profiles, parts 1, 2. *Communications on Pure and Applied Mathematics* 1957; **10**:45–68.
15. Miersemann E. Asymptotic expansion at a corner for the capillary problem: the singular case. *Pacific Journal of Mathematics* 1993; **157**(1):95–107.
16. Adam NK. *The Physics and Chemistry of Surfaces*. Dover: New York, 1968.
17. Adamson AW. *Physical Chemistry of Surfaces*. Wiley: New York, 1990.
18. Gilbarg D, Trudinger NS. *Elliptic Partial Differential Equations of Second Order* (2nd edn). Springer: Heidelberg, 1983.

19. Hauksbee F. Account of the experiment on the ascent of water between two glass planes, in an hyperbolic figure. *Philosophical Transactions of the Royal Society of London* 1712; **27**:539. C. xiv, 131.
20. Hauksbee F. Some further experiments, showing the ascent of water between two glass planes in an hyperbolic curve. *Philosophical Transactions of the Royal Society of London* 1713; **28**:153. C. xiv, 131.
21. Taylor B. Concerning the ascent of water between two glass planes. *Philosophical Transactions of the Royal Society of London* 1712; **27**:538. C. xv, 131.
22. Bico J, Quéré D. Rise of liquids and bubbles in angular capillary tubes. *Journal of Colloid and Interface Science* 2002; **247**:162–166.
23. Finn R. Equilibrium capillary surfaces. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 284. Springer: New York, 1986.
24. Turkington B. Height estimates for exterior problems of capillarity type. *Pacific Journal of Mathematics* 1980; **88**:517–540.
25. Miersemann E. Asymptotic expansion at a corner for the capillary problem. *Pacific Journal of Mathematics* 1988; **134**:299–311.
26. King JR, Ockendon JR, Ockendon H. The Laplace–Young equation near a corner. *Quarterly Journal of Mechanics and Applied Mathematics* 1999; **52**(1):73–97.
27. Korevaar NJ. On the behaviour of a capillary surface at a re-entrant corner. *Pacific Journal of Mathematics* 1980; **88**(2):379–385.
28. Mosegaard K, Tarantola A. Probabilistic approach to inverse problems. *International Handbook of Earthquake & Engineering Seismology (Part A)*. Academic Press: New York, 2002; 237–265.
29. Bangerth W, Hartmann R, Kanschat G. deal.II *Differential Equations Analysis Library, Technical Reference*, 2004. <http://www.dealii.org/>
30. Bangerth W, Kanschat G. Concepts for object-oriented finite element software—the deal.II library. Preprint 99-43, SFB, October 1999.
31. John RCS, Draper NR. D-optimality of regression design: a review. *Technometrics* 1975; **17**:15–23.
32. Ciarlet PG. The finite element method for elliptic problems. *Studies in Mathematics and its Applications*, vol. 4 (1st edn). North-Holland: Amsterdam, New York, Oxford, 1978.
33. Kelly DW, Gago JPdeSR, Zienkiewicz OC, Babuška I. *A posteriori* error analysis and adaptive processes in the finite element method: part I—error analysis. *International Journal for Numerical Methods in Engineering* 1983; **19**:1593–1619.