Chapter 6
Stochastic Inverse Problems: Models and Metrics

6.1 Introducing the Problem

Over the past 2 years, we have been developing a theory of uncertainty quantification and propagation that is computationally feasible with large numbers of unknowns. We have applied it to a problem of characterizing the eddy-current response of a shot-peened surface, where the surface is modeled as a one-dimensional random conductivity field with a known covariance function. We are currently extending the model to more general materials characterization problems, such as modeling two-dimensional random anisotropic grain noise in titanium alloys. In this case, we assume the existence of a (two-dimensional) covariance function for the random distribution of Euler angles that define the orientation of each crystallite within the material.

With this background, we want to develop a theory of stochastic inverse problems for more traditional eddy-current NDE flaw characterization and sizing. Instead of a random material, we assume that the flaw can be characterized as a random process. That this is a reasonable approach is suggested by reference to Fig. 6.1, which shows the typical shape of fatigue-crack growth progression in cold-worked fastener holes. Clearly, the ensemble of cracks cannot be modeled by a simple canonical shape with three parameters, length, width, height, so we will need to invoke a stochastic model for analyzing such cracks.

With such a stochastic model, we can draw parallels between 'probability of detection' (POD) and 'likelihood of inversion' (LOI). In the former, we are given a flaw, and ask ourselves, 'Can we detect it, and what are the metrics that measure our success?' In the latter, we are given data, and ask ourselves, 'Can we associate a flaw with them, and what are the metrics that measure our success?'
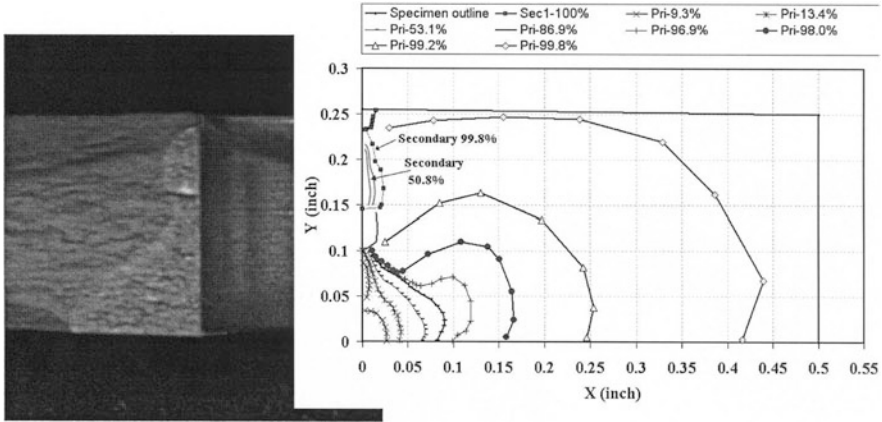
Fig. 6.1 Typical shape of fatigue-crack growth progression in cold-worked fastener holes. Images courtesty of D. Forsyth, TRI/Austin

The stochastic model will be described later, but first we will develop some background tools that are currently resident in **VIC-3D®**, and will be the basis of our stochastic computational model.

6.2 NLSE: Nonlinear Least-Squares Parameter Estimation

Let

$$Z = g(p_1, \ldots, p_N, f) , \qquad (6.1)$$

where $p_1, \ldots, p_N$ are the N parameters of interest, and f is a control parameter at which the impedance, Z, is measured. f can be frequency, scan-position, lift-off, etc. It is, of course, known; it is not one of the parameters to be determined. To be explicit during our initial discussion of the theory, we will call f 'frequency.'

In order to determine $p_1, \ldots, p_N$, we measure Z at M frequencies, $f_1, \ldots, f_M$, where $M > N$:

$$Z_1 = g(p_1, \ldots, p_N, f_1)$$

$$\vdots$$

$$Z_M = g(p_1, \ldots, p_N, f_M) . \qquad (6.2)$$

The right-hand side of (6.2) is computed by applying the volume-integral code to a model of the problem, usually at a discrete number of values of the vector, **p**, forming a multidimensional interpolation grid.

Because the problem is nonlinear, we use a Gauss-Newton iteration scheme to perform the inversion. First, we decompose (6.2) into its real and imaginary parts, thereby doubling the number of equations (we assume the $p_1, \ldots, p_N$ are real). Then we use the linear approximation to the resistance, R_i, and reactance, X_i, at the ith frequency:

$$
\begin{bmatrix} R_1 \\ X_1 \\ \vdots \\ R_M \\ X_M \end{bmatrix} \approx \begin{bmatrix} R_1(p_1^{(q)}, \ldots, p_N^{(q)}) \\ X_1(p_1^{(q)}, \ldots, p_N^{(q)}) \\ \vdots \\ R_M(p_1^{(q)}, \ldots, p_N^{(q)}) \\ X_M(p_1^{(q)}, \ldots, p_N^{(q)}) \end{bmatrix} + \begin{bmatrix} \dfrac{\partial R_1}{\partial p_1} & \cdots & \dfrac{\partial R_1}{\partial p_N} \\ \dfrac{\partial X_1}{\partial p_1} & \cdots & \dfrac{\partial X_1}{\partial p_N} \\ & \vdots & \\ \dfrac{\partial R_M}{\partial p_1} & \cdots & \dfrac{\partial R_M}{\partial p_N} \\ \dfrac{\partial X_M}{\partial p_1} & \cdots & \dfrac{\partial X_M}{\partial p_N} \end{bmatrix}_{(p_1^{(q)}, \ldots, p_N^{(q)})} \begin{bmatrix} p_1 - p_1^{(q)} \\ \vdots \\ p_N - p_N^{(q)} \end{bmatrix}, \quad (6.3)
$$

where the superscript (q) denotes the qth iteration, and the partial derivatives are computed numerically by the software. The left side of (6.3) is taken to be the measured values of resistance and reactance. We rewrite (6.3) as

$$ 0 \approx r + Jp , \qquad (6.4) $$

where r is the $2M$-vector of residuals, J is the $2M \times N$ Jacobian matrix of derivatives, and p is the N-dimensional correction vector. Equation (6.4) is solved in a least-squares manner starting with an initial value, $(x_1^{(0)}, \ldots, x_N^{(0)})$, for the vector of unknowns, and then continuing by replacing the initial vector with the updated vector $(x_1^{(q)}, \ldots, x_N^{(q)})$ that is obtained from (6.3), until convergence occurs.

We are interested in determining a bound for the sensitivity of the residual norm to changes in some linear combination of the parameters. Given an $\epsilon > 0$ and a unit vector, v, the problem is to determine a sensitivity (upper) bound, σ, such that

$$ \|r(x^* + \sigma v)\| \le (1 + \epsilon)\|r(x^*)\| . \qquad (6.5) $$

We will derive an estimate of σ. Equation (6.5) is equivalent to

$$ \|r(x^* + \sigma v)\| - \|r(x^*)\| \le \epsilon \|r(x^*)\| . \qquad (6.6) $$

The left-hand side of (6.6) can be approximated to the second order in σ by the second-order Taylor expansion:

$$\|r(x^* + \sigma v)\| - \|r(x^*)\| \approx \sigma v \cdot \nabla \|r(x^*)\| + \frac{\sigma^2}{2} \sum_{i,j} \frac{\partial^2 \|r(x)\|}{\partial x_j \partial x_i}|_{x^*} v_i v_j , \quad (6.7)$$

where ∇ is the gradient operator in $N-$dimensional space. Even though the gradient vanishes at the minimum point, we will compute it to get the algebra started:

$$\nabla \|r(x)\| = \nabla \left[f_1^2(x) + f_2^2(x) + \cdots + f_{2M}^2(x) \right]^{1/2}$$

$$= \frac{1}{\|r(x)\|} \begin{bmatrix} f_1 \dfrac{\partial f_1}{\partial x_1} + \cdots + f_{2M} \dfrac{\partial f_{2M}}{\partial x_1} \\ \vdots \\ f_1 \dfrac{\partial f_1}{\partial x_N} + \cdots + f_{2M} \dfrac{\partial f_{2M}}{\partial x_N} \end{bmatrix}^T$$

$$= \frac{r(x)^T}{\|r(x)\|} \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_N} \\ & \vdots & \\ \dfrac{\partial f_{2M}}{\partial x_1} & \cdots & \dfrac{\partial f_{2M}}{\partial x_N} \end{bmatrix}$$

$$= e^T(x) \cdot J , \quad (6.8)$$

where the superscript T denotes the transpose of a matrix (or vector), and $e(x) = r(x)/\|r(x)\|$ is a unit vector.

The second derivative that we want is the gradient of (6.8):

$$\nabla \nabla \|r(x)\| = -\frac{\nabla \|r(x)\|}{\|r(x)\|^2} \begin{bmatrix} f_1 \dfrac{\partial f_1}{\partial x_1} + \cdots + f_{2M} \dfrac{\partial f_{2M}}{\partial x_1} \\ \vdots \\ f_1 \dfrac{\partial f_1}{\partial x_N} + \cdots + f_{2M} \dfrac{\partial f_{2M}}{\partial x_N} \end{bmatrix}^T$$

$$+ \frac{1}{\|r(x)\|} \nabla \begin{bmatrix} f_1 \dfrac{\partial f_1}{\partial x_1} + \cdots + f_{2M} \dfrac{\partial f_{2M}}{\partial x_1} \\ \vdots \\ f_1 \dfrac{\partial f_1}{\partial x_N} + \cdots + f_{2M} \dfrac{\partial f_{2M}}{\partial x_N} \end{bmatrix}^T . \quad (6.9)$$

Before going further, we can immediately drop the first term in (6.9) because the gradient of the norm vanishes at the solution x^*. Thus, (6.9) becomes, using index notation,

$$\nabla\nabla\|r(x)\| = \frac{1}{\|r(x)\|}\frac{\partial}{\partial x_j}\left[f_1\frac{\partial f_1}{\partial x_i} + \cdots + f_{2M}\frac{\partial f_{2M}}{\partial x_i}\right], \quad i,j = 1,\ldots,N$$

$$= \frac{1}{\|r(x)\|}\sum_\alpha\left[\frac{\partial f_\alpha}{\partial x_j}\frac{\partial f_\alpha}{\partial x_i} + f_\alpha\frac{\partial^2 f_\alpha}{\partial x_j\partial x_i}\right], \quad \alpha = 1,\ldots,2M. \quad (6.10)$$

Following [88, page 523], we discard the second-derivative term in (6.10) by arguing that the residual vector for a good model fit should be small, which would make the second derivative term small. Furthermore, it is likely that the residual vector should have terms that are uncorrelated with each other and with the model, thus tending to cancel the second derivative terms when summed over α. We will call (6.10) the first-order curvature tensor, Γ_{ij}, of the mapping (or deformation) of the parameter space, $\{x_i\}$, into the model-measurement space. If we call the ith column of the Jacobian matrix, c_i, then it follows from (6.10) that

$$\Gamma_{ij}(x^*) = \frac{c_i(x^*)\cdot c_j(x^*)}{\|r(x^*)\|}, \quad (6.11)$$

where we are ignoring the second-derivative term in (6.10).

Digression on Computing $\Gamma_{ij}(x^*)$ We can use the MINPACK code that is already in NLSE to compute $c_i(x^*)\cdot c_j(x^*)$. The computation of the diagonal elements is already available as the 'self sensitivities,' so that leaves the off-diagonal elements. Consider $\|c_i(x^*)/\sqrt{2} + c_j(x^*)/\sqrt{2}\|^2 = \dfrac{\|c_i(x^*)\|^2}{2} + c_i(x^*)\cdot c_j(x^*) + \dfrac{\|c_j(x^*)\|^2}{2}$. Hence, it follows that $c_i(x^*)\cdot c_j(x^*) = \|c_i(x^*)/\sqrt{2} + c_j(x^*)/\sqrt{2}\|^2 - \dfrac{\|c_i(x^*)\|^2 + \|c_j(x^*)\|^2}{2}$, where the right-hand side is already calculable using MINPACK in NLSE.

Substituting this result into (6.7) yields an upper bound for the quadratic term:

$$\sigma^2\sum_{i,j}\frac{\partial^2\|r(x)\|}{\partial x_j\partial x_i}|_{x^*}v_i v_j = \frac{\sigma^2}{\|r(x^*)\|}\sum_\alpha\left[\sum_{i,j}\frac{\partial f_\alpha}{\partial x_i}v_i\frac{\partial f_\alpha}{\partial x_j}v_j\right]_{x^*}$$

$$= \frac{\sigma^2}{\|r(x^*)\|}(J(x^*)\cdot v)\cdot(J(x^*)\cdot v)$$

$$= \frac{\sigma^2}{\|r(x^*)\|}\|J(x^*)\cdot v\|^2, \quad (6.12)$$

and if we equate this to the right-hand side of (6.6), we get the final result

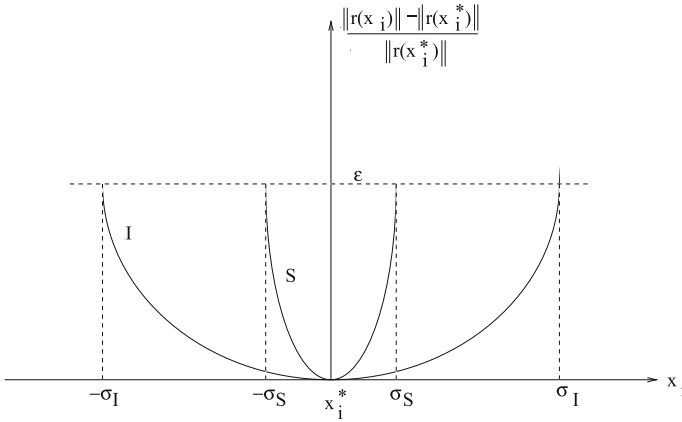$$\sigma_v = \epsilon^{1/2}\left(\frac{\|r(x^*)\|}{\|J(x^*)\cdot v\|}\right). \quad (6.13)$$

Fig. 6.2 Showing sensitivity parameters for two system responses to x_i. Response S is sensitive to x_i at x_i^*, whereas response I is not

We will call this the 'first-order' approximation, in the sense that we have truncated the Taylor series expansion with the first nonzero term, and have ignored the second-derivative terms in (6.10). This is the expression that is stated, but not derived, in [77].

Note that if $\|J(x^*) \cdot v\|$ is small compared to $\|r(x^*)\|$, then σ is large and the residual norm is insensitive to changes in the linear combination of the parameters specified by v. If $v = e_i$, the ith column of the $N \times N$ identity matrix, then (6.13) produces σ_i, the sensitivity bound for the ith parameter. Since σ_i will vary in size with the magnitude of x_i^*, it is better to compare the ratios σ_i/x_i^* for $i = 1, \ldots, N$ before drawing conclusions about the fitness of a solution.

The importance of these results is that we now have metrics for the inversion process: $\Phi = \|r(x^*)\|$, the norm of the residual vector at the solution, tells us how good the fit is between the model data and measured data. The smaller this number the better, of course, but the 'smallness' depends upon the experimental setup and the accuracy of the model to fit the experiment. Heuristic judgement based on experience will help in determining the quality of the solution for a given Φ.

The sensitivity coefficient, σ, is more subtle, but just as important. It, too, should be small, but, again, the quality of the 'smallness' will be determined by heuristics based upon the problem. If σ is large in some sense, it suggests that the solution is relatively independent of that parameter, so that we cannot reasonably accept the value assigned to that parameter as being meaningful, as suggested in Fig. 6.2, which shows a system, S, for which the system is sensitive to variable, x_i, at the solution point, x_i^*, and another system, I, for which the system is insensitive to x_i.

An example occurs when one uses a high-frequency excitation, with its attendant small skin depth, to interrogate a deep-seated flaw. The flaw will be relatively invisible to the probe at this frequency, and whatever value is given for its parameters

will be highly suspect. When this occurs we will either choose a new parameter to characterize the flaw, or acquire data at a lower frequency.

These metrics are not available to us in the current inspection method, in which analog instruments acquire data that are then interpreted by humans using hardware standards. The opportunity to use these metrics is a significant advantage to the model-based inversion paradigm that we propose in this paper.

6.3   Confidence Levels: Stochastic Global Optimization

We can extend the previous results to obtain a statistical measure of confidence in the solution. Referring to Fig. 6.2, we have the probability relation

$$\text{Prob}[x_i^* - \sigma_v v \le x_i \le x_i^* + \sigma_v v] = \text{Prob}\left[\frac{\|r(x_i)\| - \|r(x_i^*)\|}{\|r(x_i^*)\|} \le \epsilon\right]. \qquad (6.14)$$

Arguing that $\dfrac{\|r(x_i)\| - \|r(x_i^*)\|}{\|r(x_i^*)\|}$ is a random variable allows us to transform the inverse methods of [111] into the realm of 'stochastic inverse problems.'

This approach is based on the current 'Multi-Level Single Linkage' algorithm that is used in NLSE to reach the global minimum with probability one [21, 78, 89, 94], and also fits our concept of 'stochastic inversion.' Furthermore, it allows us to use prior knowledge of the unknown parameters. Let the model parameters, $\{x_n\}$, be a set of independent random variables, each uniformly distributed over its known range of values. We'll sample the parameter space by choosing, say, 500 points randomly, in accordance with the distribution function of each parameter, and compute the norm of the residual vector at each of the points, as in the first step of NLSE. In NLSE, these points are trial initial points for the minimization algorithm, (6.3), and the lowest of the resulting 500 minima is guaranteed to be the global minimum with unit probability [21, 78, 89, 94].[1]

The random variable, $\dfrac{\|r(x_i)\| - \|r(x_i^*)\|}{\|r(x_i^*)\|}$, in (6.14) is a continuous function of $\{x_i\}$ defined on a compact set (the 'prior feasible set'), so it achieves a finite maximum on that set. This maximum, if it could be determined with probability one, is precisely ϵ in (6.14), and when this is substituted into the transfer function, (6.13), we would have determined the confidence level, σ_v, with unit probability. Later we will relax any claims of unit probability in determining ϵ, but we are permitted to

---

[1] The Multi-Level Single Linkage method guarantees that the global minimum will be found within a finite number of iterations with probability one, given a sufficiently large sample size of trial points. Numerical experiments with model and laboratory data for a variety of inverse problems over many years [111] suggest that 500 trial points yield a reliable estimate of the global minimum for problems with the number of variables that we are considering.
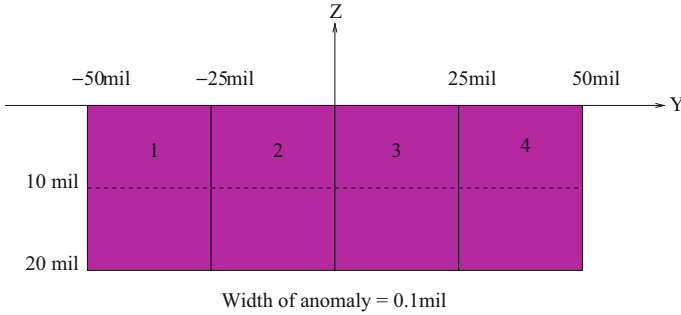
Fig. 6.3 Showing the configuration of the one-dimensional pulse basis functions for parametrizing a complex flaw. The nodes are located at depths of 0, 10, and 20 mils

make a strong statement about the confidence level, because in this formulation of a stochastic inverse problem, we are assuming prior statistical constraints of the unknown parameters, $\{x_n\}$. This approach is quite 'Bayesian', in the sense that we are combining prior information on the random variables with a likelihood estimation (which follows from the least-squares inversion process) to get posterior information on the variables.

Example: A Complex 'Flaw' The configuration of the problem is shown in Fig. 6.3. The expansion of the flaw in the $(Y, Z)-$plane is given by

$$f(y, z) = \sum_{i=1}^{4} \alpha_i \pi_i^{(1)}(y) \pi^{(1)}(z) , \tag{6.15}$$

where $\pi^{(1)}$ is a unit pulse function, and the expansion coefficients, $\{\alpha_i\}_{i=1}^{4}$, determine the magnitude of $\pi^{(1)}(z)$. These coefficients are the unknown degrees of freedom of the problem, and will be modeled as independent random variables with a uniform distribution over the range $[0, 20]$. They will be determined by inversion of the data, which are impedances measured by a probe that is scanned over $-100 \leq Y \leq 100$, $X = 0$. It should be understood that this formalism fixes the resolution of the flaw in the $Y-$direction to be 25 mils, as well as the width of the flaw in the $X-$direction to be 0.1 mil. These numbers are arbitrary, of course, and can be changed to suit the problem. Furthermore, with the four blocks arranged as shown, this configuration will be best suited for modeling and reconstructing midbore, throughwall, and corner bolt-hole cracks.

Figure 6.4 illustrates a complex flaw extending over the entire range in Y. We will use the output of a **VIC-3D®** model of this flaw to serve as the input data for inversion. To illustrate the inversion process and the importance of the 'surrogate' interpolation table for the $\{\alpha_i\}$, we will perform a numerical experiment in which the table has successively two, three and four nodes per dimension. In the first case, the
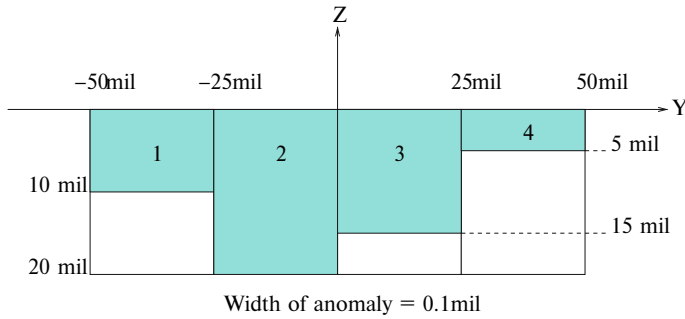
Width of anomaly = 0.1mil

Fig. 6.4 Showing the configuration of a complex flaw extending over the entire range of Fig. 6.3

Table 6.1 Results for the example problem vs. number of nodes per dimension

# Nodes	$\|r(\alpha^*)\|$	α_1/sensit	α_2/sensit	α_3/sensit	α_4/sensit	# Points
2	0.319(−1)	11.47/4.12	20.0/3.4	16.80/3.58	4.64/4.46	272
3	0.405(−2)	10.18/0.53	19.86/0.698	15.87/0.558	3.9/0.444	226
4	0.159(−2)	11.19/0.125	20.11/0.216	15.56/0.212	6.06/0.1918	255

nodes are at [0, 20] , in the second, they are at [0, 10, 20], as in Fig. 6.3, and in the third, [0, 7, 14, 21] (in this case, we assume a uniform distribution of the variables over the range [0, 21]). Thus, the first table comprises $2^4 = 16$ nodes, the second $3^4 = 81$ nodes, and the last $4^4 = 256$ nodes. A blending function for each node is computed by **VIC-3D®**. We quickly see the 'curse of dimensionality' occurring. This curse will be obviated through the use of sparse-grid interpolation techniques to reduce the computational burden of building the new table.

The results of the experiment are shown in Table 6.1. The column labeled '# Points' lists the number of the original 500 global starting points that are attracted to the global minimum. These results show that increasing the number of nodes per dimension yields improvements in reducing the norm of the residuals, Φ, and the sensitivity coefficients of each variable. Figure 6.5 illustrates the results of Table 6.1, and clearly indicates that increasing the number of nodes beyond 4 will have little effect on the norm of the residuals, r, and only a slight reduction in the various sensitivity coefficients, sensit_i.

We ran NLSE four times, effectively sampling the $\{\alpha_i\}$ space 2000 times, yielding values of $\|r(\alpha)\|_{\max} = 0.2545, \ 0.2689, \ 0.2351,$ and 0.265. The inverted results of each of these runs were identical to those tabulated in Table 6.1, as we expected, since the algorithm in NLSE ensures convergence to the global minimum with probability one. Hence, using the data of the bottom row of Table 6.1 we have

$$\frac{\|r(\alpha)\|_{\max} - \|r(\alpha^*)\|}{\|r(\alpha^*)\|} = \frac{0.2689 - 0.00159}{0.00159} = 168.12 = \epsilon \,, \tag{6.16}$$
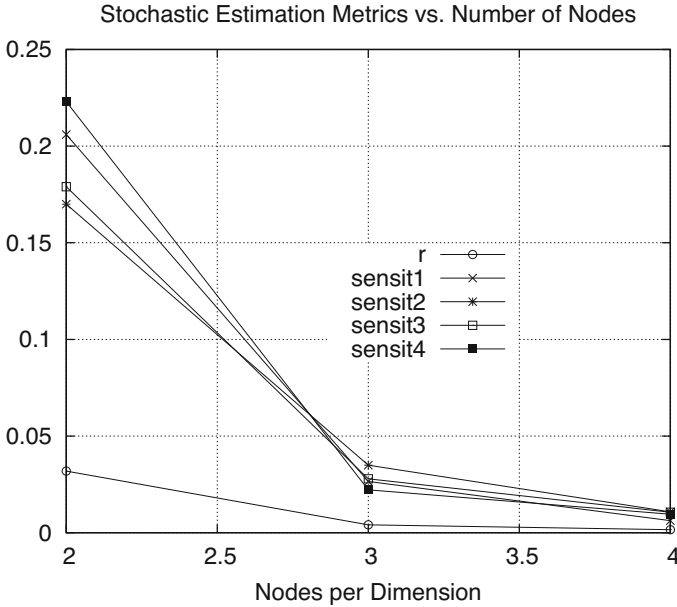
Fig. 6.5 The results of Table 6.1 indicate that increasing the number of nodes beyond 4 will have little effect on the norm of the residuals, r, and only a slight reduction in the various sensitivity coefficients, sensit(i). The sensitivity coefficients have been scaled downward by a factor of 20 to make them commensurate with the norm of the residuals for clarity of visualization

and when this is substituted into (6.13), along with the sensitivity coefficients tabulated in the bottom row of Table 6.1, we get the parameters of the confidence intervals to be $\sigma_1 = 1.62$, $\sigma_2 = 2.8$, $\sigma_3 = 2.75$, $\sigma_4 = 2.49$. These effectively define the posterior distribution of the $\{\alpha_i\}$, which is certainly much different than the prior distribution.

We summarize the results for α_i by claiming that we are 'certain' that $\alpha_i - \sigma_i \leq \alpha_i \leq \alpha_i + \sigma_i$, with the most likely value being α_i^*. In the case where one of the posterior limits on α_i exceeds the prior limit, we reject it in favor of the prior limit, because if the crack actually exceeded the prior limit, the inversion process would have been constrained at the prior limit of the interpolation table. For example, $17.31 \leq \alpha_2 \leq 21$, rather than $17.31 \leq \alpha_2 \leq 22.91$.

The Chebyshev Inequality We can improve the calculation of the confidence level, and even make its definition more precise in our example, by resorting to the Chebyshev inequality [63], which states that, if Z is a random variable, then, for every $\xi > 0$,

$$P[|Z| \geq \xi] \leq \frac{\text{VAR}(Z)}{\xi^2} = \text{MAX UNCERTAINTY}(\xi)$$

Table 6.2 Results at 100 Hz–1 kHz for conductivity and permeability

Trial	$\|r(x)\|_{max}$	$\|r(x^*)\|$	σ/sensit	μ/sensit
1	0.2503	0.188(−2)	1.372(6)/2.32(4)	68.18/0.1504
2	0.2509	0.188(−2)	1.372(6)/2.32(4)	68.18/0.1504
3	0.2521	0.188(−2)	1.372(6)/2.32(4)	68.18/0.1504
4	0.2552	0.188(−2)	1.372(6)/2.32(4)	68.18/0.1504
5	0.2525	0.188(−2)	1.372(6)/2.32(4)	68.18/0.1504

$$\text{MINIMUM CERTAINTY}(\xi) = 1 - \frac{\text{VAR}(Z)}{\xi^2} , \tag{6.17}$$

where ξ is the threshold or decision boundary for determining the confidence interval. For example, if we want to be at least 95% confident in our assertion of the probability of the first equality in (6.17), then $1 - \dfrac{\text{VAR}(Z)}{\xi^2} = 0.95$, which implies that $\xi = \left(\dfrac{\text{VAR}(Z)}{0.05}\right)^{1/2}$.

To apply this theorem to our problem, we define $Z = \|r(\alpha)\|_{max} - \overline{\|r(\alpha)\|}_{max}$, where $\|r(\alpha)\|_{max}$ is a random variable whose sample value is the output of the following 'experiment': run a 500-sample trial, as in the Multi-Level Single Linkage algorithm, and choose the largest result for $\|r(\alpha)\|_{max}$. Repeat the experiment for the second sample, and so on. We have already given an example of this, with the result after four trials that $\{\|r(\alpha)\|_{max}\} = \{0.2545, 0.2689, 0.2351, 0.265\}$, from which follow $\overline{\|r(\alpha)\|}_{max} = 0.2559$, $\text{VAR}(Z) = 0.0001716$, and $\xi = (0.0001716/0.05)^{1/2} = 0.0586$ for 95% confidence level.

From the Chebyshev inequality we have, therefore, $\|r(\alpha)\|_{max} = 0.2559 + 0.0586 = 0.3145$. This replaces $\|r(\alpha)\|_{max} = 0.2689$ in (6.16), so that the 95% upper bound is given by

$$\frac{\|r(\alpha)\|_{max} - \|r(\alpha^*)\|}{\|r(\alpha^*)\|} = \frac{0.3145 - 0.00159}{0.00159} = 196.8 = \epsilon . \tag{6.18}$$

The new values for the parameters corresponding to the 95% confidence interval are $\{\sigma_1 = 1.75, \sigma_2 = 3.03, \sigma_3 = 2.97, \sigma_4 = 2.69\}$. The confidence intervals for the four variables are, therefore: $\alpha_1 : [9.44, 12.94]$, $\alpha_2 : [17.08, 21]$, $\alpha_3 : [12.59, 18.53]$, $\alpha_4 : [3.37, 8.75]$.

Joint Measurement of Conductivity and Magnetic Permeability We have taken impedance measurements over the frequency range of 100Hz–1 kHz of a ferritic heat-exchanger tube, with the intention of jointly determining the conductivity and relative magnetic permeability of the tube. The interpolation table had the following nodal values: $\sigma : 1.0 \times 10^6, 1.2 \times 10^6, 1.4 \times 10^6, 1.6 \times 10^6, 1.8 \times 10^6$; $\mu :$ 50, 60, 70, 80, 90. We ran five trials of NLSE with the following results (Table 6.2):

Table 6.3 Inverted results
for the width of a crack

Trial	$\|r(x)\|_{\max}$	$\|r(x^*)\|$	W/sensit
1	0.7979(−1)	0.2495(−1)	0.0793/0.0296
2	0.7990(−1)	0.2495(−1)	0.0793/0.0296
3	0.8024(−1)	0.2495(−1)	0.0793/0.0296
4	0.7941(−1)	0.2495(−1)	0.0793/0.0296
5	0.7999(−1)	0.2495(−1)	0.0793/0.0296

Following the procedure described above with respect to the Chebyshev inequality, we calculate a value of $\epsilon = 137$, which yields $\sigma_{\text{cond}} = 0.272 \times 10^6$, and $\sigma_\mu = 1.76$. Hence, we can say that the most likely value of the conductivity is 1.372×10^6, with a 95% confidence interval of $[1.1 \times 10^6, \ 1.644 \times 10^6]$. For the permeability we get even tighter results; the most likely value is 68.18, with a 95% confidence interval of $[66.42, \ 69.94]$.

The fact that the permeability is well defined at these low frequencies has been validated by use of the Cramer-Rao Lower Bound (CRLB), [111, pp. 407–410], where it is also shown that the optimum frequency for estimating conductivity is 6.0 kHz.

Estimation of Width of a Long, Thin Crack We are given data at 200 kHz for a crack in a bolt-hole. The data were obtained by a splitD probe with ferrite cores, and the crack was 100 mils long and 18 mils deep. The objective was to determine the width of the crack. The problem is described in greater detail in [111, Section 6.6].

The interpolation table for the width has nodes at 0, 0.125 mils, and 0.25 mils. The inverted results after 5 trials are shown in Table 6.3.

These results yield a value of $\epsilon = 2.31$ and $\sigma_W = 0.045$. The most likely value of W is 0.0793 mils, and the 95% confidence interval is $[0.0343, \ 0.1243]$. We should note that in these two examples, the confidence interval calculation becomes more precise with an increase in the number of nodes in the interpolation table, as indicated earlier.

6.4 Summary

We summarize the algorithm and process here.

1. $\dfrac{\|r(x)\|_{\max} - \|r(x^*)\|}{\|r(x^*)\|} = \text{OBJ}(x)$ is a random variable.
2. $\|r(x^*)\|$ and the Jacobian, $J(x^*)$, are determined with prob $\to 1$ (Stochastic Global Optimization via MLSL).
3. The set $\{x(\epsilon)\} \ni \text{OBJ}(x) \leq \epsilon$ is the 'posterior feasible set at level ϵ'.
4. If OBJ(x) is parabolic (ellipsoidal in N-space), then the set $\{x(\epsilon)\}$ is called the 'first-order posterior feasible set at level ϵ'.
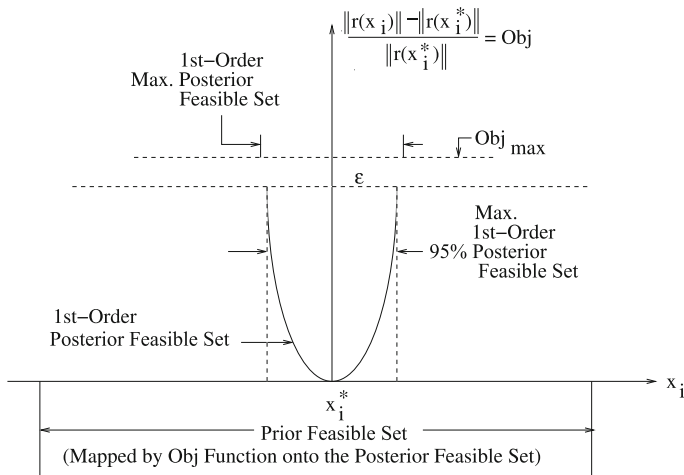
Fig. 6.6 Illustrating the interpretation and calculation of confidence intervals

5. $\sigma_v = \epsilon^{1/2}\left(\dfrac{\|r(x^*)\|}{\|J(x^*)\cdot v\|}\right)$ is a mapping from the 'prior feasible set' to the 'first-order posterior feasible set at level ϵ'.

6. If we choose ϵ to be at the 95% confidence level, as with the Chebyshev Inequality, then the measure of $\{x^* - \sigma \leq x \leq x^* + \sigma\}$ is at least 95% that of the maximum first-order posterior feasible set, and x^* is the most likely value of x.

Figure 6.6 illustrates the algorithm.

A 2D Example The results just given are for the situation in which each parameter is tested separately, while the others are fixed at the solution point. Now, we must consider the general case in which the totality of variables are considered jointly. This means operating in four-dimensional space. The tools that we have already set up allow us to do that with no additional expense, except for a minor enhancement to the NLSE code in **VIC-3D®**. Equation (6.13) is valid for arbitrary orientations of the unit vector, v, and ϵ has already been computed using the entire four-dimensional random parameter space in the MLSL stochastic global optimization algorithm.

Consider the 2D example shown in Fig. 6.7, which is the projection onto the (x_1, x_2)-plane of the four-dimensional hyperellipsoid associated with the complex flaw example described earlier. Using NLSE, we compute the joint sensitivity associated with the unit vector, $v = [0.5, 0.5, 0.5, 0.5]$ to be 0.129. Then, using $\epsilon = 196.8$, as before, we compute $\sigma_{0.5,0.5,0.5,0.5} = 1.81$ from (6.13) for the 95%-confidence region for this combination of variables. It should be understood
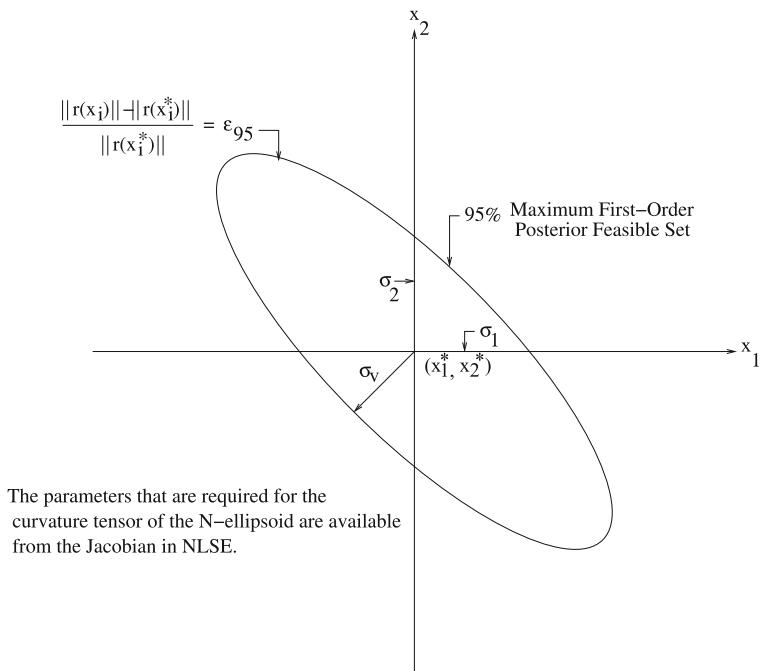
Fig. 6.7 A 2D ellipsoidal confidence region. The origin, (x_1^*, x_2^*), is the solution of the optimization problem

that NLSE already gives us the information to generate the entire N-dimensional hyperellipsoid for a given problem. This would allow us to analytically calculate such things as the volume of the ellipsoid, or cross-sectional areas, etc.