Chapter 12
NLSE: Parameter-Based Inversion Algorithm

12.1 Introduction

Chapter 11 introduced us to the notion of an inverse problem and gave us some examples of the value of this idea to the solution of realistic industrial problems. The basic inversion algorithm described in Chap. 11 was based upon the Gauss–Newton theory of nonlinear least-squares estimation and is called NLSE in this book. In this chapter we will develop the mathematical background of this theory more fully, because this algorithm will be the foundation of inverse methods and their applications during the remainder of this book. We hope, thereby, to introduce the reader to the application of sophisticated mathematical concepts to engineering practice without introducing excessive mathematical sophistication.

We separate the discussion of inversion algorithms into two categories: (a) parameter-based, in which a few parameters are used to define the problem or anomalous region (as in Chap. 11) and (b) voxel-based, in which the anomalous region is constructed voxel-by-voxel on a grid. NLSE is the prototype of the former algorithm; the next volume in this series will introduce us to voxel-based algorithms.

12.2 NLSE: Nonlinear Least-Squares Parameter Estimation

12.2.1 Overview of the Algorithm: Nonlinear Least-Squares

Let

$$Z = g(p_1, \ldots, p_N, f) \,, \tag{12.1}$$

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."

H.A. Sabbagh et al., *Computational Electromagnetics and Model-Based Inversion*, Scientific Computation, DOI 10.1007/978-1-4419-8429-6_12,

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

$$\begin{aligned} Z_1 &= g(p_1,\ldots,p_N,f_1) \\ &\vdots \\ Z_M &= g(p_1,\ldots,p_N,f_M). \end{aligned} \tag{12.2}$$

The right-hand side of (12.2) 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 [97]. First, we decompose (12.2) 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:

$$\begin{bmatrix} R_1 \\ X_1 \\ \vdots \\ R_M \\ X_M \end{bmatrix} \approx \begin{bmatrix} R_1(p_1^{(q)},\ldots,p_N^{(q)}) \\ X_1(p_1^{(q)},\ldots,p_N^{(q)}) \\ \vdots \\ R_M(p_1^{(q)},\ldots,p_N^{(q)}) \\ X_M(p_1^{(q)},\ldots,p_N^{(q)}) \end{bmatrix} + \left[\begin{array}{ccc} \frac{\partial R_1}{\partial p_1} & \cdots & \frac{\partial R_1}{\partial p_N} \\ \frac{\partial X_1}{\partial p_1} & \cdots & \frac{\partial X_1}{\partial p_N} \\ & \vdots & \\ \frac{\partial R_M}{\partial p_1} & \cdots & \frac{\partial R_M}{\partial p_N} \\ \frac{\partial X_M}{\partial p_1} & \cdots & \frac{\partial X_M}{\partial p_N} \end{array}\right]_{(p_1^{(q)},\ldots,p_N^{(q)})} \begin{bmatrix} p_1 - p_1^{(q)} \\ \vdots \\ p_N - p_N^{(q)} \end{bmatrix}, \tag{12.3}$$

where the superscript (q) denotes the qth iteration and the partial derivatives are computed numerically by the software. The left side of (12.3) is taken to be the measured values of resistance and reactance. We rewrite (12.3) as

$$0 \approx r + Jp\ , \tag{12.4}$$

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 (12.4) 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 (12.3), 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 $\varepsilon > 0$ and a unit vector, v, the problem is to determine a sensitivity (upper) bound, σ, such that

$$\|r(x^* + \sigma v)\| \le (1+\varepsilon)\|r(x^*)\|\ . \tag{12.5}$$

We will derive a first-order estimate of σ. Equation (12.5) is equivalent to

$$\|r(x^* + \sigma v)\| - \|r(x^*)\| \leq \varepsilon \|r(x^*)\| \,. \tag{12.6}$$

The left-hand side of (12.6) can be approximated to the first order in σ by the first-order Taylor expansion:

$$\|r(x^* + \sigma v)\| - \|r(x^*)\| \approx \sigma v \cdot \nabla \|r(x^*)\| \,, \tag{12.7}$$

where ∇ is the gradient operator in N-dimensional space. We compute the gradient of the norm of the residual vector:

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

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

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

$$= e^T(x) \cdot J \,, \tag{12.8}$$

where the superscript T denotes the transpose of a matrix (or vector) and $e(x) = r(x)/\|r(x)\|$ is a unit vector. Thus, (12.7) becomes

$$\|r(x^* + \sigma v)\| - \|r(x^*)\| \approx \sigma e^T(x) \cdot J \cdot v \,. \tag{12.9}$$

The factor multiplying σ in (12.9) is the dot product of the two vectors, $e(x)$ and $J \cdot v$. Hence, its value is less than the product of the magnitude of each vector, which means that (12.9) becomes

$$\|r(x^* + \sigma v)\| - \|r(x^*)\| \leq \sigma \|J \cdot v\| \,, \tag{12.10}$$

because $e(x)$ has unit magnitude. Upon equating the right-hand side of (12.10) to the right-hand side of (12.6), we obtain the first-order estimate of σ:

$$\sigma_v = \varepsilon \left(\frac{\|r(x^*)\|}{\|J(x^*) \cdot v\|} \right) \,. \tag{12.11}$$

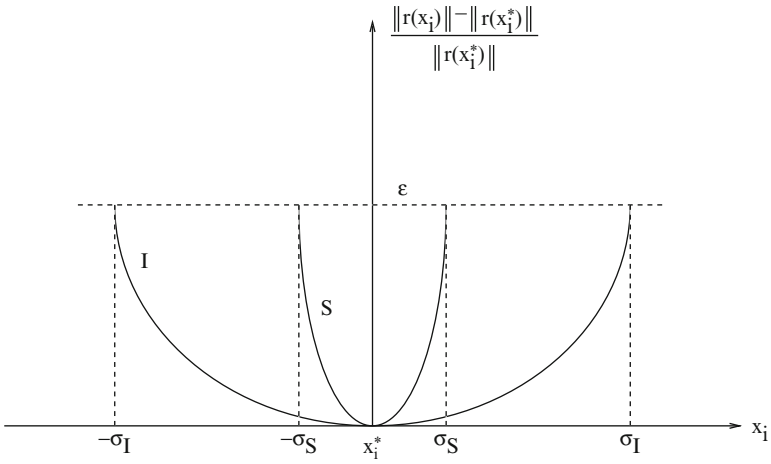

Fig. 12.1 Showing sensitivity parameters for two system responses to x_i. Response S is sensitive to x_i at x_i^*, whereas response I is not

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 (12.11) 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 $i = 1, \cdots, N$ 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. 12.1, 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 book.

If the residual norm is relatively insensitive to changes in some linear combination of the parameters, then the Jacobian matrix at the solution is nearly rank-deficient, and it may be useful to determine a set of linearly-independent parameters. The covariance matrix $(J^T J)^{-1}$ can be used for this purpose.

12.2.2 Stochastic Methods for Global Optimization

The problem defined above leads to the global optimization problem: finding the lowest minimizer of a nonlinear function of several variables that has multiple local minimizers. In the stochastic approach to global optimization, one applies a strictly descent local search procedure to a subset of a sample of starting points drawn from a uniform distribution over $\mathscr{R}$, so as to find all the local minima of $\|r\|$ that are potentially global [58–60]. One such stochastic approach, the *multilevel, single-linkage*, will, with probability one, find all relevant local minima of the objective function with the smallest possible number of local searches [59, 61]. We do not implement the multilevel, single-linkage approach in this book, but use a uniform distribution of starting points (in each coordinate direction), as in the multilevel approach. We do not reduce the sample size, however, but use the entire sample, comprising, perhaps, 500 points in each variable. We have found that, even with four variables, the procedure is so fast with modern machines, that it is quite efficient.

Hence, our global optimization algorithm starts by generating a uniform distribution of 500 points in each coordinate, and then immediately applying the Gauss–Newton iteration to each of these points. The result is 500 local minima, which are then ordered to give the smallest to largest values of the norm of the residual. The location of the smallest of these local minima is presumed to be the global minimum.

12.2.3 Computation of Function Values

NLSE is a post-processing feature of **VIC-3D®**. **VIC-3D®** is applied a priori to compute function values at certain equi-spaced values of the parameters, and these results are then stored in a table for interpolation. This speeds up the application of the algorithm, though it will require a large database for interpolation, the size, of course, depending upon the number of unknowns, and the number of precomputed function values for each unknown. The order of the interpolator is arbitrary, but we typically use first-order to fourth-order polynomial-splines for each variable, which means that we need to compute the norm of the residuals at two to five values of each variable. The derivatives in the Jacobian matrix are computed by interpolating in this table of precomputed function values.

12.2.4 Application of Statistical Communication Theory

Additional insight can be gained by appealing to the statistical theory of communication [50]. Pretend that the measured impedance vector, $\mathbf{Z} = [Z_1, \ldots, Z_M]$, of (12.2) is a random vector, along with the parameter vector, $\mathbf{p} = [p_1, \ldots, p_N]$. Think of $\mathbf{Z}$ as a received message over some communication channel, and $\mathbf{p}$ as a transmitted message. These two random vectors have a joint probability density, $\mathscr{P}_{\mathbf{Z},\mathbf{p}}(\mathbf{Z},\mathbf{p})$, which is to be maximized. That is to say, for a given $\mathbf{Z}$, we want to determine that $\mathbf{p}$ which maximizes $\mathscr{P}_{\mathbf{Z},\mathbf{p}}$.

From probability theory, we have Bayes rule:

$$\begin{aligned} \mathscr{P}_{\mathbf{Z},\mathbf{p}} &= \mathscr{P}_{\mathbf{Z}}(\mathbf{Z})\mathscr{P}(\mathbf{p}|\mathbf{Z}) \\ &= \mathscr{P}_{\mathbf{p}}(\mathbf{p})\mathscr{P}(\mathbf{Z}|\mathbf{p}) \\ &= \mathscr{P}_{\mathbf{p},\mathbf{Z}}(\mathbf{p},\mathbf{Z}) \ , \end{aligned} \tag{12.12}$$

where $\mathscr{P}_{\mathbf{Z}}(\mathbf{Z})$ and $\mathscr{P}_{\mathbf{p}}(\mathbf{p})$ are called a priori probability density functions, and $\mathscr{P}(\mathbf{p}|\mathbf{Z})$, $\mathscr{P}(\mathbf{Z}|\mathbf{p})$ are called conditional, or a posteriori, probability density functions. The variable to the right of the vertical bar in these latter functions is called the conditioning variable; for example, $\mathscr{P}(\mathbf{p}|\mathbf{Z})$ is called the a posteriori probability density for $\mathbf{p}$, conditioned on the fact that $\mathbf{Z}$ was received (or measured). This a posteriori probability density function is the object of our interest, as we shall now see.

From (12.12) we want to maximize

$$\mathscr{P}_{\mathbf{Z}}(\mathbf{Z})\mathscr{P}(\mathbf{p}|\mathbf{Z}) = \mathscr{P}_{\mathbf{p}}(\mathbf{p})\mathscr{P}(\mathbf{Z}|\mathbf{p}) \tag{12.13}$$

over $\mathbf{p}$. Because $\mathscr{P}_{\mathbf{Z}}(\mathbf{Z})$ is independent of $\mathbf{p}$, we can ignore it and maximize the a posteriori probability density

$$\mathscr{P}(\mathbf{p}|\mathbf{Z}) \propto \mathscr{P}_{\mathbf{p}}(\mathbf{p})\mathscr{P}(\mathbf{Z}|\mathbf{p}) \ . \tag{12.14}$$

That is, we want to choose that "message," $\mathbf{p}$, that is best associated with the "received signal," $\mathbf{Z}$.

If the measured $\mathbf{Z}$ of (12.2) is corrupted by "noise," of density $\mathscr{P}_{\mathbf{n}}$, then we can replace the conditional density $\mathscr{P}(\mathbf{Z}|\mathbf{p})$ by $\mathscr{P}_{\mathbf{n}}(Z_1 - g_1, Z_2 - g_2, \ldots, Z_M - g_M|\mathbf{p})$, where $g_1 = g(\mathbf{p}, f_1), \ldots, g_M = g(\mathbf{p}, f_M)$ in (12.2). Because the noise in the measurements is independent of the transmitted message, $\mathbf{p}$, we can ignore the conditioning variable, and replace (12.14) by

$$\mathscr{P}(\mathbf{p}|\mathbf{Z}) \propto \mathscr{P}_{\mathbf{p}}(\mathbf{p})\mathscr{P}_{\mathbf{n}}(\mathbf{Z} - \mathbf{g}(\mathbf{p})) \ . \tag{12.15}$$

If we have no prior knowledge of $\mathbf{p}$, or if all transmitted messages are a priori equally likely, then we can ignore $\mathscr{P}_{\mathbf{p}}(\mathbf{p})$ in (12.15) and work with the "likelihood function," $\mathscr{P}_{\mathbf{N}}(\mathbf{Z} - \mathbf{g}(\mathbf{p}))$. Maximizing the likelihood function over $\mathbf{p}$ is called "maximum likelihood estimation." We usually have some prior knowledge of $\mathbf{p}$, however, so we incorporate that knowledge in the a priori function, $\mathscr{P}_{\mathbf{p}}(\mathbf{p})$.

If $\mathbf{p}$ and $\mathbf{n}$ are jointly Gaussian processes, then we have the classical problem of communicating a Gaussian signal in Gaussian noise, which reduces to a classical least-squares problem. Typically, we work with the negative logarithm of the a posteriori density in (12.15), so we need to minimize

$$-\ln \mathscr{P}(\mathbf{p}|\mathbf{Z}) = -\ln \mathscr{P}_{\mathbf{p}}(\mathbf{p}) - \ln \mathscr{P}_{\mathbf{n}}(\mathbf{Z}-\mathbf{g}(\mathbf{p})) + f(\mathbf{Z}) \; . \tag{12.16}$$

Now, we let the M components of the noise vector, $\mathbf{n}$, be statistically independent, zero-mean, Gaussian random variables, each with variance, σ_n^2, and the N components of $\mathbf{p}$ be statistically independent Gaussian random variables, with mean values $\overline{\mathbf{p}}$, and each with variance, σ_p^2. Then, upon taking the negative logarithm of the appropriate density functions, and discarding unimportant factors, we replace (12.16) with the objective function

$$\begin{aligned} \Phi(\mathbf{p}|\mathbf{Z}) &= \frac{1}{2\sigma_n^2}\sum_{i=1}^{M}(Z_i - g_i(\mathbf{p}))^2 + \frac{1}{2\sigma_p^2}\sum_{i=1}^{N}(p_i - \overline{p}_i)^2 \\ &= \frac{|\mathbf{Z}-\mathbf{g}(\mathbf{p})|^2}{2\sigma_n^2} + \frac{|\mathbf{p}-\overline{\mathbf{p}}|^2}{2\sigma_p^2} \; , \end{aligned} \tag{12.17}$$

which is to be minimized over $\mathbf{p}$. Multiply (12.17) by $2\sigma_n^2$ to get the final expression for the objective function

$$\Phi(\mathbf{p}|\mathbf{Z}) = |\mathbf{Z}-\mathbf{g}(\mathbf{p})|^2 + \frac{\sigma_n^2}{\sigma_p^2}|\mathbf{p}-\overline{\mathbf{p}}|^2 \; , \tag{12.18}$$

where we use the same notation for the objective function.

The ratio, σ_p^2/σ_n^2, of the variances is called the signal-to-noise ratio and is usually known. In the context of least-squares problems, this ratio is known as the Levenberg–Marquardt parameter, and in mathematical inverse theory it is called the Tichonov–Miller parameter.

Setting the value of the LM-parameter determines the certainty with which we choose to assert the a priori constraint on $\mathbf{p}$. If σ_p is very small, then the LM-parameter is large, and we are more certain to impose the constraint. Unless we know the variances, we cannot determine the LM-parameter at the outset. There are numerical techniques, such as *ridge regression* [53] and *cross-validation* [62], that can be useful in selecting the LM-parameter.

Appendix

A.1 Cramer–Rao Lower Bound

The Cramer–Rao lower bound (CRLB), being the lower bound on the variance of any unbiased estimator, plays a role in statistical estimation theory [64, 65] that is similar to the sensitivity (upper) bound, σ, of (12.11). As with σ, the CRLB finds application in electromagnetic scattering and inverse problems [66].

A.1.1 Inverse Method Quality Metrics

Given the potential of inverse methods, it is important to develop a rigorous method for quantifying the performance and reliability of inversion schemes [105, 107]. Although empirical studies provide the means for evaluating the quality of NDE techniques incorporating inverse methods, opportunities also exist with inverse methods to use the model calculations with quantitative measures to evaluate key estimation performance metrics without considerable experimental burden.

In estimation theory, the Cramer–Rao Lower Bound (CRLB) provides the minimum variance that can be expected for an unbiased estimator of a set of unknown parameters. In other words, the CRLB provides a way of quantifying the inversion algorithm performance. For Gaussian noise, there is a simple inverse relationship between the CRLB and the Fisher information [65]:

$$\mathrm{var}\left(\hat{\theta}_i\right) = \left[C_{\hat{\theta}}\right]_{ii} \geq \left[I^{-1}(\theta\right]_{ii} \ , \tag{12.19}$$

where **C** is the covariance matrix, the Fisher information is defined as

$$I(\theta)_{ii} = -E\left[\frac{\partial^2 \ln f(Z;\theta)}{\partial\theta_i \partial\theta_j}\right] \ , \tag{12.20}$$

θ is the parameter being estimated, and Z is the measurement vector. Fisher information represents the amount of information contained in a measurement and depends on the derivatives of the likelihood function which is based on the forward model and the noise parameters. The variance in a measurement is inversely related to the amount of information contained in the measurement, so it is not a surprise that (12.19) shows that the variance in the measurement is greater than or equal to the inverse of the Fisher information matrix. In eddy-current NDE, the measurement is often the real and imaginary component of the impedance, $Z = [R, X]$, and the Fisher information becomes a square matrix with dimensions equal to the number of parameters being estimated.

The covariance matrix can be evaluated as a performance metric for inverse methods. First, the diagonal terms of the covariance matrix (the CRLB variances) provide a metric of sensitivity of a parameter estimated using inverse methods to measurement variation. Second, the off-diagonal terms represent the interdependence between select parameters being estimated to measurement variation. The corresponding metric is the correlation coefficient given by

$$\rho_{i,j} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}} \ . \tag{12.21}$$

These metrics can be used with parametric studies involving frequency or other probe parameters to optimize the NDE system design. As a general design rule for inverse methods, it is desirable to minimize the sensitivity to variation (the CRLB variances) and to have the correlation coefficient between the parameters being estimated approach zero.

Another tool used in numerical linear algebra for sensitivity analysis is singular value decomposition (SVD). SVD essentially provides a measure of sensitivity of measurements to perturbations in the unknown parameters [106]. To evaluate the sensitivity of an inverse problem for a set of measurements to changes in fit parameters, SVD can be applied to the Jacobian matrix where

$$\mathbf{J} = \begin{bmatrix} \frac{\partial Z_1}{\partial \theta_1} & \cdots & \frac{\partial Z_1}{\partial \theta_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial Z_m}{\partial \theta_1} & \cdots & \frac{\partial Z_m}{\partial \theta_n} \end{bmatrix} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}' \,. \tag{12.22}$$

The condition number (CN) of the matrix is defined as the ratio of the largest and smallest singular values resulting from SVD. For inversion, CN has been used to quantify the well-posedness of the inverse problem for select parameters. The ability to estimate parameters independently increases as the condition number approaches unity. It should be noted that SVD does not incorporate noise; it depends only on the noiseless relationship between the measurement output and the parameter changes.

A.1.2 Optimizing Layer Estimation Using Metrics

An inversion experiment is revisited [55] for the purpose of demonstrating estimation theory metrics [107]. In this experiment, the thickness of an AISI-304 stainless steel plate and probe lift-off were estimated. A thickness and lift-off model, similar to the one shown in Chap. 11, was used to solve the forward problem. The estimation procedure is represented in (12.23). The left side is the measured impedance, the Jacobian is simply the derivative information from the forward model, and the thickness and lift-off parameters are updated until this equation converges,

$$\begin{bmatrix} R(f,t,l) \\ X(f,t,l) \end{bmatrix} \approx \begin{bmatrix} R(f,t_0,l_0) \\ X(f,t_0,l_0) \end{bmatrix} + \begin{bmatrix} \frac{\partial R}{\partial t} & \frac{\partial R}{\partial l} \\ \frac{\partial X}{\partial t} & \frac{\partial X}{\partial l} \end{bmatrix}_{t_0,l_0} \begin{bmatrix} t - t_0 \\ l - l_0 \end{bmatrix} . \tag{12.23}$$

Four scenarios in particular are investigated. Impedance values were generated for combinations of lift-off values of 0.75 and 1.5 mm and a plate thickness values of 1.0 mm and 2.0 mm with Gaussian noise of 1 % of the impedance value added as shown in Fig. 12.2a. For each of these measurements, the NLSE algorithm is applied to estimate the thickness and lift-off simultaneously. Figure 12.2b shows the inversion results in the parameter space. Note that for high lift-off, visual inspection indicates the variance in the estimation is much greater for lift-off and likewise for the thicker plate, the variance of the estimation of thickness is greater.

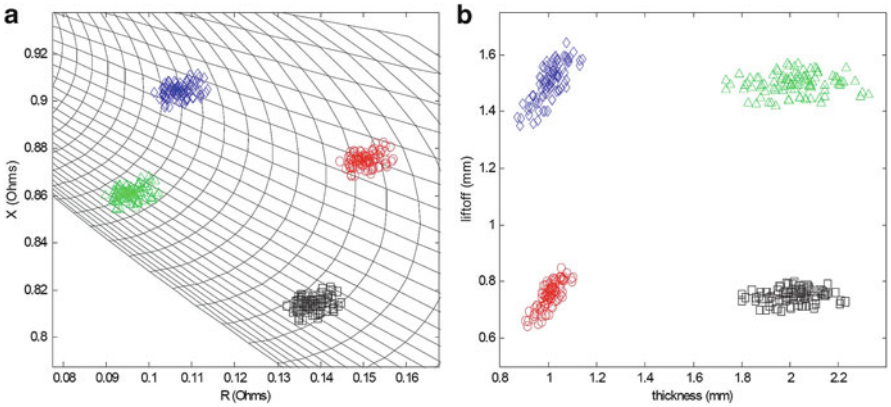

Fig. 12.2 (**a**) Distribution of source data in impedance plane, and (**b**) corresponding estimated values in lift-off-thickness parameter space

The calculations required for the CRLB involve taking numerical derivatives of the impedance changes with respect to the parameter changes from the forward model. These calculations thus require far less computational expense with respect to Monte-Carlo simulation. Following (12.20), the Fisher information for this particular case is given by:

$$\mathbf{I} = \begin{bmatrix} J_{11}^2 + J_{21}^2 & J_{12}J_{11} + J_{22}J_{21} \\ J_{11}J_{12} + J_{21}J_{22} & J_{12}^2 + J_{22}^2 \end{bmatrix} . \tag{12.24}$$

The covariance matrix is then calculated from the Fisher information (by (12.19)):

$$\mathbf{C} = \sigma^2 \mathbf{I}^{-1} . \tag{12.25}$$

The Jacobian is also decomposed into its singular values and singular vectors in the form of the right-hand side of (12.22). The ratio of the smallest to largest singular values provides the condition number.

Figure 12.3 shows the CRLB of the estimation of the thickness and lift-off of a 1 mm thick plate and 1 mm lift-off for multiple frequencies. The agreement between the CRLB and the Monte-Carlo approach is quite good. This analysis demonstrates that there is an optimal frequency to achieve highest accuracy in the estimation of thickness. Estimating conductivity and thickness simultaneously is typically more ill-conditioned than estimating thickness and lift-off simultaneously. The CRLB for conductivity and thickness estimation along with the condition number and correlation number as a function of frequency are all displayed in Fig. 12.4. The behavior of the CRLB as a function of frequency for estimating conductivity and thickness simultaneously follows a similar trend and this is expected since the impedance changes due to conductivity and thickness are similar. The condition number reaches a maximum around 95 kHz which implies that selectivity is good and the correlation is zero at this frequency which further confirms that point.

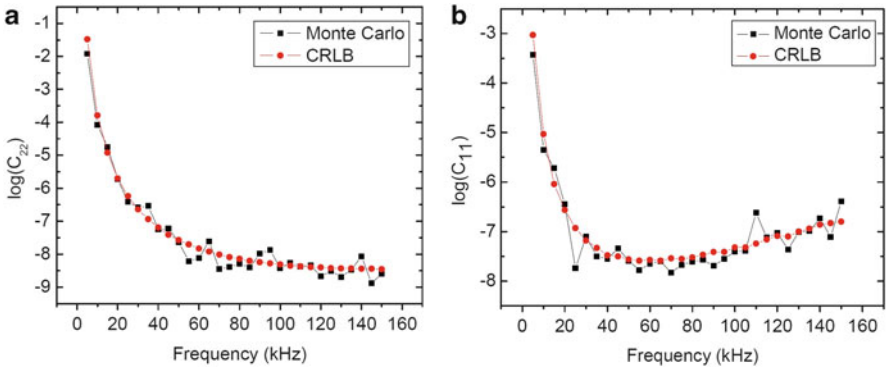

Fig. 12.3 Comparison of variance with varying frequency using CRLB and Monte Carlo methods for estimating (**a**) lift-off and (**b**) thickness, respectively

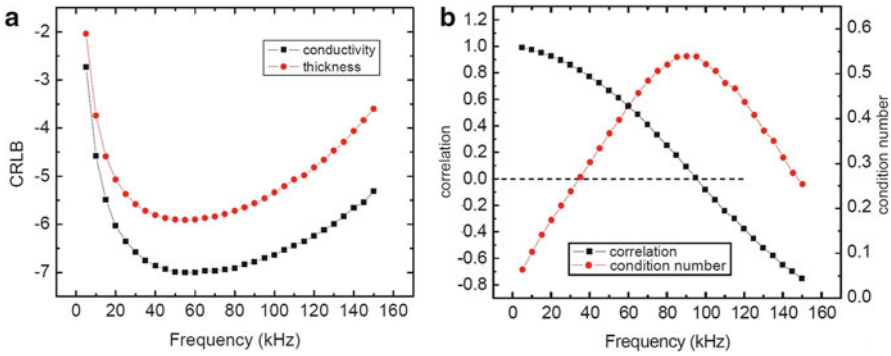

Fig. 12.4 Comparison of inversion metrics with varying frequency: (**a**) CRLB variance for thickness and conductivity estimation and (**b**) correlation and condition number

A.1.3 Two Examples from Chap. 11

Figure 11.5 of Chap. 11 shows the variation with frequency of the derivative of impedance with respect to thickness for Test Case 1 (shown in Fig. 11.4). There is a relatively broad peak in this derivative extending from roughly 20 to 50 kHz. Thus, we would expect that the optimum range of frequencies to be used for determining the thickness should lie in this range and that was essentially confirmed when frequencies in the range of 20–40 kHz were found to give good results.

Figure 12.5(top) illustrates the CRLB thickness response for this same test case, and we see that there is a relatively broad minimum centered near 50 kHz, but generally including the same frequency range chosen in executing the test case. This confirms the consistency between the CRLB result and the information contained in the first derivative.

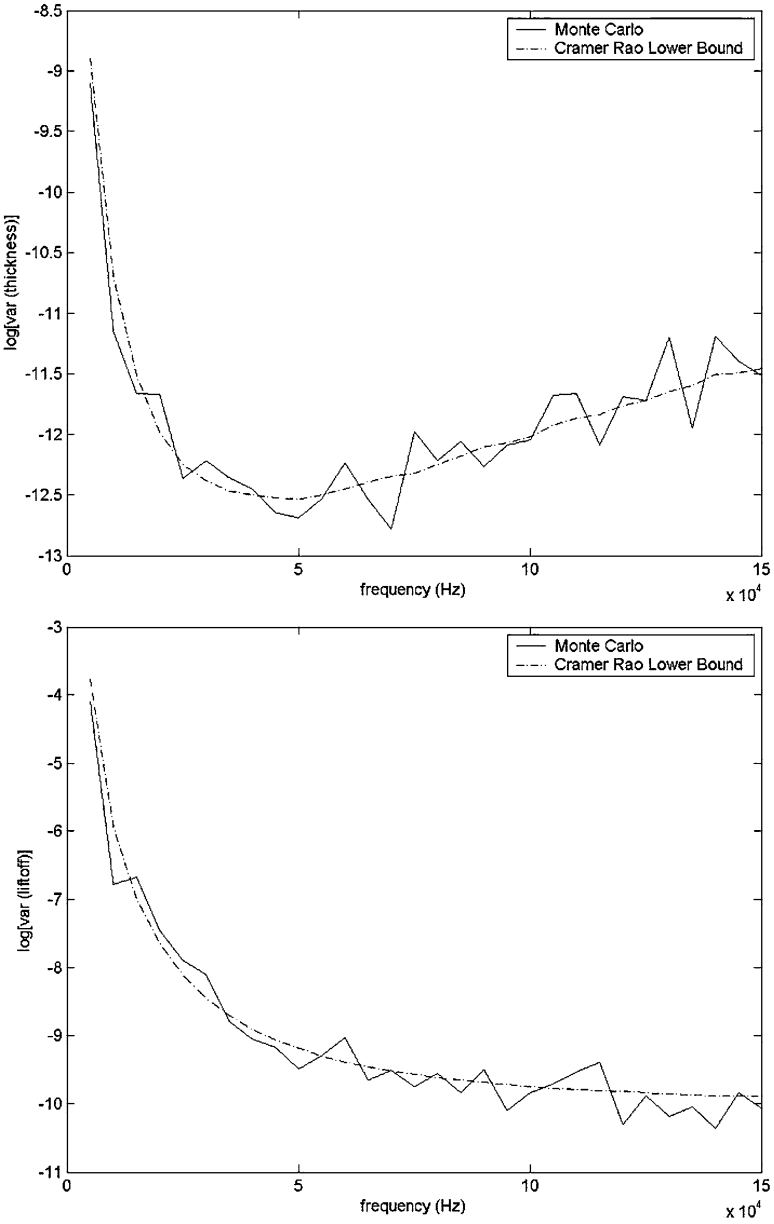

Fig. 12.5 Comparison of variance with varying frequency for Test Case 1 in Chap. 11, using CRLB and Monte Carlo methods for estimating (*top*) thickness and (*bottom*) lift-off, respectively

As for the lift-off problem in Test Case 1, we see from Fig. 11.6 that the derivative with respect to lift-off continues to increase with frequency even beyond 100 kHz, which corresponds to what we might expect for the CRLB associated with lift-off,

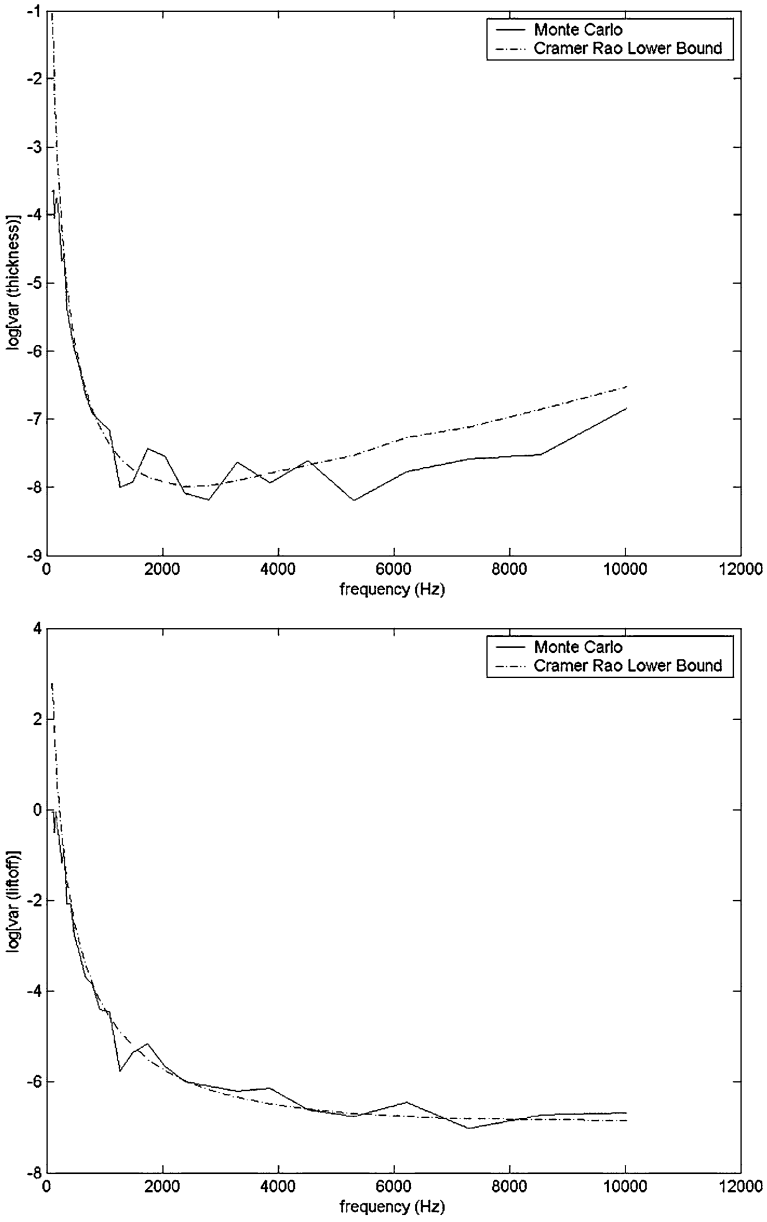

Fig. 12.6 Comparison of variance with varying frequency for Test Case 2 in Chap. 11, using CRLB and Monte Carlo methods for estimating (*top*) thickness and (*bottom*) lift-off, respectively

as shown in the bottom part of Fig. 12.5. In the latter figure, we see that the CRLB never does achieve a true minimum, even out to 150 kHz, but the decrease is leveling off at this frequency.

Similar results are obtained for Test Case 2 of Chap. 11, except that the frequencies are much lower (see Figs. 11.9 and 11.10). The derivative with respect to thickness for a 1.0 mm-thick brass plate has a maximum in the vicinity of 2 kHz, and the derivative with respect to lift-off peaks at approximately 10 kHz, for a nominal lift-off of 2.0 mm. These results are consistent with the CRLB results shown in Fig. 12.6. In particular, note that the CRLB for lift-off is virtually flat beyond 8 kHz, which corresponds with the levelling off of the peak derivative shown in Fig. 11.10.

A.2 Selected Bibliography of Inverse Problems in Eddy-Current NDE

- H. A. Sabbagh and L. D. Sabbagh, Development of a system to invert eddy-current data and reconstruct flaws, Final Report: Contract No. N60921-81-C-0302, Naval Surface Weapons Center, White Oak Labs, Silver Springs, MD (1982).
- H. A. Sabbagh and L. D. Sabbagh, Inversion of eddy-current data and the reconstruction of flaws using multifrequencies, Final Report: Contract No. N60921-82-C-0139, Naval Surface Weapons Center, White Oak Labs, Silver Springs, MD (1983).
- B. A. Auld, G. Mcfetridge, M. Riaziat, and S. Jefferies, Improved probe-flaw interaction modeling, inversion processing, and surface roughness clutter, Review of Progress in QNDE, Vol 4A (1985).
- B. A. Auld, S. Jefferies, J. C. Moulder, and J. C. Gerlitz, Semi-elliptical surface flaw EC interaction and inversion: theory, Review of Progress in QNDE, Vol 5A (1986).
- H. A. Sabbagh and L. D. Sabbagh, An eddy-current model for three-dimensional inversion, IEEE Transactions on Magnetics (1986).
- H. A. Sabbagh, L. D. Sabbagh, Verification of an eddy-current flaw inversion algorithm, IEEE Transactions on Magnetics (1986).
- L. Udpa and S. Udpa, Solution of inverse problems in eddy-current nondestructive evaluation (NDE), Journal of Nondestructive Evaluation (1988).
- L. D. Sabbagh, and H. A. Sabbagh, Eddy-current modeling and flaw reconstruction, Journal of Nondestructive Evaluation (1988).
- H. A. Sabbagh, L. D. Sabbagh, and T. M. Roberts, An eddy-current model and algorithm for three-dimensional nondestructive evaluation of advanced composites, IEEE Transaction on Magnetics (1988).
- S. M. Nair and J. H. Rose, Reconstruction of three-dimensional conductivity variations from eddy current (electromagnetic induction) data, Inverse Problems (1990).
- J. C. Moulder, E. Uzal, and J. H. Rose, Thickness and conductivity of metallic layers from eddy current measurements, Rev. Sci. Instrum (1992).

- S. J. Norton and J. R. Bowler, Theory of eddy current inversion, J. Appl. Phys (1993).
- O. Baltzersen, Model-based inversion of plate thickness and lift-off from eddy current probe coil measurments, Materials Evaluation (1993).
- N. J. Goldfine, Magnetometers for improved materials characterization in aerospace applications, Materials evaluation (1993).
- H. A. Sabbagh and R. G. Lautzenheiser, Inverse problems in electromagnetic nondestructive evaluation, International Journal of Applied Electromagnetics in Materials (1993).
- J. R. Bowler, S. Norton, D. J. Harrison, Eddy-current interaction with an ideal crack. II. The inverse problem, J. Appl. Phys. (1994).
- S. K. Burke, Eddy-current inversion in the thin-skin limit: Determination of depth and opening for a long crack, J. Appl. Phys. (1994).
- C. Tai, J. H. Rose, and J. C. Moulder, Thickness and conductivity of metallic layers from pulsed eddy-current measurments, Rev. Sci. Instrum. (1996).
- L. Udpa and S. Udpa, Application of signal processing and pattern recognition techniques to inverse problems in NDE, Int J. of Applied Electromagnetics and Mechanics (1997).
- Z. Chen and K. Miya, ECT inversion using a knowledge-based forward solver, Journal of Nondestructive Evaluation (1998).
- H. T. Banks et al, Nondestructive evaluation using a reduced-order computational methodology, Inverse Problems (2000).
- H. T. Banks et al, Real time computational algorithms for eddy-current-based damage detection, Inverse Problems (2002).
- W. Yin, S. J. Dickinson, and A. J. Peyton, Imaging the continuous conductivity profile within layered metal structures using inductance spectroscopy, IEEE Sensors Journal (2005).
- W. Yin and A. J. Peyton, Thickness measurement of non-magnetic plates using multi-frequency eddy current sensors, NDT&E International (2007).