ACOUSTIC SIGNAL MODELLING AND ITS APPLICATION TO THE CALIBRATION OF UNDERWATER ELECTROACOUSTIC TRANSDUCERS IN REVERBERANT LABORATORY TANKS

P M Harris (2) and S P Robinson (1)

(1) Centre for Mechanical and Acoustical Metrology, National Physical Laboratory, UK
(2) Centre for Information Systems Engineering, National Physical Laboratory, UK

1. INTRODUCTION

In the calibration of underwater electroacoustic transducers, the stepped-sinusoidal signals that are commonly used are contaminated by transients due to the resonant behaviour of the devices, as well as noise. In addition, when calibrations are undertaken in reverberant laboratory tanks of finite size, reflections from the tank boundaries often arrive before the steady-state response of the transducer can be observed directly.

The approach considered here to such calibration is to model the free-time response of the device by a function consisting of a sum of complex exponential terms which are used to describe both the steady-state and resonant behaviour of the device. In this paper, we review two classes of method: linear prediction methods (such as Prony's method and its variants [4, 10]), and nonlinear least-squares methods.

Linear prediction methods are shown to be statistically biased and inefficient. This observation motivates the development of a linear prediction method in which proper account is taken of the error structure in the data. A nonlinear least-squares algorithm is also described based on a safeguarded Gauss-Newton algorithm that uses regularisation to address the ill-conditioning that is a property of the underlying problem. Results are presented of using these methods to analyse data obtained from measurements of real devices with the aim of applying them to transducer calibration.

2. MODELS

In order to achieve good *prediction* from the model fitted to the data (which is important when direct measurement of the steady-state is limited), we consider here *physical* models. Assuming that the system behaves as a linear damped harmonic oscillator, we model the observed output $y(t)$ of the system by a function consisting of a sum of p complex exponential terms:

$$y(t) = \sum_{k=1}^{p} \alpha_k e^{\beta_k t}. \quad (1)$$

The residues α_k and poles β_k are complex-valued and, in order for $y(t)$ to be real-valued, occur in complex conjugate pairs. An alternative way of writing the model that uses real-valued parameters only is

$$y(t) = A_1 \sin\{2\pi f_1 t + \phi_1\} + \sum_{k=2}^{1+n_r} A_k e^{d_k t} \sin\{2\pi f_k t + \phi_k\}, \quad (2)$$

where we use here a single undamped sinusoid to represent the steady-state behaviour and n_r damped sinusoids to represent the resonant behaviour. The parameters in (2) are the amplitudes A_k, frequencies f_k, damping factors d_k and phase angles ϕ_k.

SIGNAL MODELLING IN REVERBERANT LABORATORY TANKS

In some circumstances we may wish to include *a priori* knowledge about the system output. How this is done depends on the model used and the approach chosen to fit the model to the data. For example, if we know accurately the frequency of the signal driving the system, this fixes the frequency parameter f_1 for the undamped component. Furthermore, if we have determined elsewhere the frequency and Q-factor for a resonance of the system, this gives us information about the frequency and damping factor parameters for one of the damped components. This information can be incorporated as above by replacing, for example, f_2 and d_2 by their known values. Alternatively, if these values arise from measurement and are known with an associated uncertainty, the information can be incorporated in the estimation problem by including additional observation equations. The latter approach is discussed in Section 3.

3. ESTIMATION

We are concerned with fitting models $y(t)$ of the forms discussed in Section 2 to data $\{(t_i, y_i): i = 1, \ldots, m\}$ measured at equally spaced times t_i. If ε_i is the measurement error for the ith data value y_i, and the values t_i are known accurately, we wish to determine the function $y(t)$ that satisfies the *observation equations*

$$y_i = y(t_i) + \varepsilon_i, \qquad i = 1, \ldots, m. \tag{3}$$

If we assume the errors ε_i are uncorrelated samples from a Gaussian probability distribution with mean zero and standard deviation σ, unbiased and efficient estimates of the parameters defining $y(t)$ are obtained by solving

$$\text{minimise} \quad \sum_{i=1}^{m} \{y_i - y(t_i)\}^2 \tag{4}$$

with respect to the parameters of $y(t)$. The residuals $e_i = y_i - y(t_i)$ evaluated at the solution provide estimates of the errors ε_i, and an estimate of σ is given by the root-mean-square error s where

$$s = \sqrt{\frac{1}{m-n} \sum\nolimits_{i=1}^{m} e_i^2}, \tag{5}$$

and n is the number of parameters defining $y(t)$.

If *a priori* knowledge of any of the model parameters is available, this information can be represented by additional observation equations, and the estimation problem is modified accordingly. For example, if f_r and d_r are the frequency and damping factor for a resonance of the system, the estimation problem becomes

$$\text{minimise} \quad \sum_{i=1}^{m} \{y_i - y(t_i)\}^2 + u^2\{f_r - f_2\}^2 + v^2\{d_r - d_2\}^2, \tag{6}$$

where u and v are "weights" that are used to reflect the relative accuracy between the measured data and the *a priori* knowledge.

Although we assume the device behaves linearly, the model $y(t)$ is a nonlinear function of its parameters, and problems (4) and (6) are *nonlinear least-squares* problems. Algorithms for solving these problems are discussed in Section 5. Furthermore, because of the particular type of model considered, we will see that these problems are difficult to solve (Section 6).

There is a large body of work in the literature concerned with *linear prediction* methods, also known as *Prony* methods [4, 10]. These methods have the following advantages: (a) they involve the solution of (simpler) linear problems, and (b) they are direct methods, i.e., unlike algorithms for the nonlinear problem, they do not require that initial estimates of the parameters are provided. However, because the methods do not reflect properly the error structure associated with the data, they do not provide unbiased and efficient estimates of the model parameters. Nevertheless, linear prediction methods are important because they provide a means of generating starting estimates for algorithms for the nonlinear problem. Linear prediction methods are discussed in Section 4.

4. LINEAR ESTIMATION METHODS

Linear estimation methods are based on the observation that the model values $y(t_i)$ satisfy a set of recurrence equations, with undetermined parameters $\delta_1, \delta_2, \ldots, \delta_{p+1}$:

$$\delta_1 y(t_i) + \delta_2 y(t_{i+1}) + \ldots + \delta_{p+1} y(t_{i+p}) = 0, \qquad i = 1, \ldots, m-p. \tag{7}$$

Replacing the values $y(t_i)$ by the measured values y_i, the recurrence equations are no longer satisfied exactly, but instead

$$\delta_1 y_i + \delta_2 y_{i+1} + \ldots + \delta_{p+1} y_{i+p} = e_i, \qquad i = 1, \ldots, m-p. \tag{8}$$

These constitute a set of linear equations for the parameters $\delta_1, \delta_2, \ldots, \delta_{p+1}$. If we set $\delta_{p+1} = 1$ (note the parameters are determined up to a scale factor), we determine estimates of the parameters $\mathbf{d} = (\delta_1, \delta_2, \ldots, \delta_p)^T$ by solving the least-squares problem

$$\text{minimise} \quad \mathbf{e}^T \mathbf{e}, \qquad \mathbf{e}^T \mathbf{e} = \sum_{i=1}^{m-p} e_i^2, \tag{9}$$

with respect to $\mathbf{d}$. The poles β_k for the system are then recovered from the parameters $\delta_1, \delta_2, \ldots, \delta_{p+1}$. Finally, the residues α_k are obtained by fitting the model (1), now regarded as a function of α_k only, to the data: this is another linear least-squares problem.

The procedure described above is the least-squares Prony method described in [4, 10, 12, 13]. It is noted in [11] that although the method is consistent as $\sigma \to 0$, it is inconsistent as $m \to \infty$. (In other words, in the presence of noise, the method gives estimates that do not converge to the true solution as the number of sampled points increases.) Consequently, the method is only useful for low noise levels regardless of how many measurements are made. This has led to the development of a number of variations on the basic method. For example, the use of every lth sample to satisfy the recurrence equations instead of adjacent samples [4]; the use of a large number p of poles with forward and backward prediction to help in distinguishing between "true" system poles and poles associated with measurement noise [10]; the use of a large number p of poles with the singular value decomposition [7] or complete orthogonal factorization [3]. There has also been work concerned with including *a priori* information. In [10], this is done by applying a filter to the data to remove known system poles before applying a Prony method; in [3, 6], recurrence equations relating samples of the input *and* output signals are used to model the system's transfer function.

It is a property of least-squares estimation that we expect solving (9) to provide unbiased and efficient estimates of the parameters $\mathbf{d}$ in the case that the errors e_i are samples of random variables that are independent and identically distributed. Suppose that ε_i is the error associated with the ith measurement as in (3). It is reasonable to assume that the errors ε_i are samples of random variables that are independent and identically distributed, i.e., their variances and covariances are, respectively,

$$E(\varepsilon_i^2) = \sigma^2 \qquad \text{and} \qquad E(\varepsilon_i \varepsilon_j) = 0, \qquad i, j = 1, \ldots, m, \quad i \neq j. \tag{10}$$

We show below that this property is *not* inherited by the errors e_i, and consequently the least-squares Prony method cannot provide unbiased and efficient estimates of the parameters $\mathbf{d}$.

Substituting (3) into (8), we obtain

$$e_i = \sum_{k=1}^{p+1} \delta_k \{ y(t_{i-1+k}) + \varepsilon_{i-1+k} \} = \sum_{k=1}^{p+1} \delta_k \varepsilon_{i-1+k}. \tag{11}$$

It follows that the variance of e_i is given by

$$E(e_i^2) = \sum_{k=1}^{p+1} \delta_k^2 E(\varepsilon_{i-1+k}^2) = \sigma^2 \sum_{k=1}^{p+1} \delta_k^2, \tag{12}$$

and the covariance of e_i with e_{i+j}, $j = 1, \ldots, p$, is given by

$$E(e_i e_{i+j}) = \sum_{k=1}^{p+1-j} \delta_{j+k}\delta_k E(\varepsilon_{i+j-1+k}^2) = \sigma^2 \sum_{k=1}^{p+1-j} \delta_{j+k}\delta_k . \tag{13}$$

Since we do not expect in general the right hand side of (13) to be zero, we conclude that the errors e_i are not independent.

To obtain unbiased and efficient estimates of the parameters **d**, we solve the *weighted* least-squares problem

$$\text{minimise} \quad \mathbf{e}^{\mathrm{T}} V^{-1} \mathbf{e}, \tag{14}$$

with respect to **d**, where V is the covariance matrix for the errors e_i defined by (12) and (13). The problem posed in (14) correctly accounts for the error structure for the original measured data. However, because the elements of V depend on the unknown parameters **d**, it is necessary to apply an iterative scheme in which at each iteration V is formed using the estimates of **d** obtained at the previous iteration. The procedure is begun by setting V to be the identity matrix which corresponds to the original least-squares Prony method. The covariance matrix V needs to be further modified if the data in the above analysis comes from the filtering of the original measured data to remove known components.

5. NONLINEAR ESTIMATION METHODS

The estimation problems (4) and (6) given in Section 3 are nonlinear least-squares problems. Standard algorithms exist (see, for example, [8]) for solving this type of problem including the Gauss-Newton and full-Newton methods. These are iterative methods that at each iteration take a step towards the minimum by solving a linear least-squares problem. The algorithms differ in the amount of information that needs to be supplied about the model: the Gauss-Newton method requires that the first derivatives of the model with respect to its parameters are available, whereas in a full-Newton method the second derivatives are also required.

In addition to choosing an algorithm specific to least-squares problems, we can exploit structure in the model. The parameters naturally separate into those (the poles or equivalently frequencies and damping factors) that appear nonlinearly in the model, and those (the residues or equivalently amplitudes and phases) that appear linearly. The use of variable projection methods as described in [1, 2, 3, 5, 9] exploits this structure in the solution of the nonlinear least-squares problem. Yet another approach is presented in [11], involving a reparametrisation of the model and the solution of a nonlinear eigenvalue problem.

It is usual to use the solution from a linear prediction method to provide starting estimates for solving the nonlinear estimation problem. In Section 6 we note that in order to obtain good fits using linear prediction methods, it may be necessary to choose a model for which the number n_r of damped sinusoidal components exceeds what is believed to be the number of "true" resonances. We also illustrate that the nonlinear estimation problem (4) is inherently ill-conditioned, i.e., large changes in the frequency and damping factor parameters may produce small changes in the residual sum of squares function that we wish to minimise. These facts can make it difficult for standard algorithms to converge satisfactorily to a solution.

We have chosen to use a Gauss-Newton algorithm, safe-guarded with a line search algorithm, with *a priori* information incorporated as described in Section 3, equation (6). Furthermore, if additional damped sinusoidal components are used to define the initial model fit, these are either explicitly removed prior to applying the Gauss-Newton algorithm or *regularisation* is used to replace the estimation problem to be solved by one that is better conditioned.

SIGNAL MODELLING IN REVERBERANT LABORATORY TANKS

6. RESULTS

In order to apply these methods to a calibration, an ITC1001 transducer (resonance frequency: 18 kHz) was driven with discrete-frequency tone-burst signals in the range 5 to 30 kHz in steps of 1 kHz. The acoustic signal was detected using a calibrated Reson TC4034 hydrophone (resonance frequency: 350 kHz) and the waveform was acquired using a signal analyser (12-bit ADC sampling at 10 MHz). The calibrated hydrophone and measuring equipment were chosen such that the system had no pole close to the frequency range of interest. From the steady-state receive voltage, projector drive voltage, hydrophone sensitivity and projector-hydrophone separation distance, the transmitting voltage response of the projector was derived.

For the waveform analysis described below, two time-windows were applied, one of approximately two cycles of the resonance frequency and one of approximately four cycles. For each drive frequency and each time-window, the prediction methods were used to derive estimates of the steady-state receive voltage. For comparison, a free-field calibration was also performed in the 5.5 m diameter, 5 m deep test tank at NPL where the echo-free time was sufficient to allow direct measurement of the steady-state voltage.

In Figure 1 we show the recorded waveforms corresponding to drive voltages with frequencies of, respectively, 2 kHz and 20 kHz. We can identify in these figures the various "phases" in the response: the turn-on of the device followed immediately by an oscillation at its resonant frequency of approximately 18 kHz. After between three and four cycles of this oscillation, the resonant behaviour is sufficiently damped to observe the steady-state response of the device which takes the form of an undamped oscillation at the frequency of the drive voltage. It is the amplitude of this steady-state behaviour that we wish to estimate. Finally, the response is contaminated by a combination of the turn-off of the device and the arrival of the first reflections. These two effects are most clearly identified in the response to the 20 kHz drive voltage.

The "crosses" below each graph identify particular time-points used in the following analyses. The cross on the left defines a time close to, but exceeding, the turn-on of the device which we subsequently use to define the time-origin for the modelled signals (see Figure 3). The second cross defines the end of the first time-window comprising $m = 56$ points and (approximately) two cycles of the resonant response. The first and third crosses define the second time-window comprising $m = 112$ points and (approximately) four cycles of the resonant response. Finally, the first and fourth crosses define a time-window that contains a sufficient part of the steady-state for all the responses to allow direct measurement of the voltage responses to be made.

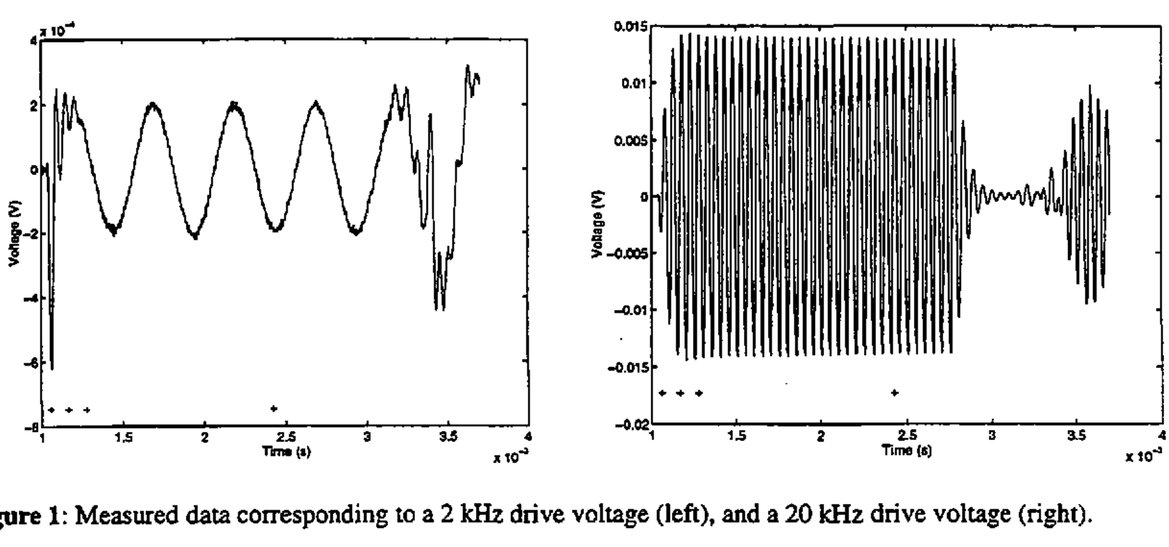

Figure 1: Measured data corresponding to a 2 kHz drive voltage (left), and a 20 kHz drive voltage (right).

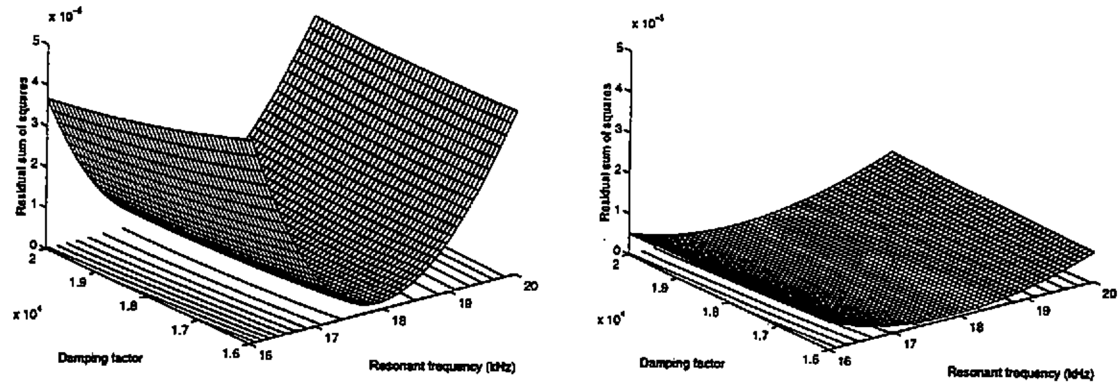

Figure 2: Residual sum of squares as a function of resonant frequency and damping factor computed for windows of the data comprising m =56 points (left), and m = 112 points (right).

Figure 2 shows the residual sum of squares surface (as a function of resonant frequency and damping) for the response of the device to a 20 kHz drive voltage restricted to, respectively, the first 112 points, and the first 56 points. We note how, for both cases, the residual sum of squares value is insensitive to changes in the resonant damping and, for the shorter time-window, it is insensitive to changes in both parameters. Furthermore, the minimum itself occurs at a different position in the two cases. This illustrates the inherent ill-conditioning of the problem of finding a minimum of the residual sum of squares function, and suggests that the problem of finding accurately the parameters defining this minimum is inherently ill-posed, and thus *any* numerical procedure can be expected to experience difficulties.

In Figure 3 we show the solutions returned by a selection of the methods considered in this work. The methods are applied using data corresponding to a 5 kHz drive voltage, and restricted to a time-window comprising m = 56 points that is delimited by the time-origin as defined earlier and the vertical dashed line. We also show the measured data and the computed model evaluated beyond this window, thus allowing us to quantify the ability of each method to estimate the steady-state amplitude.

The results of four methods are presented in Figure 3. The first is Prony's method with a model composed of a single undamped sinusoidal component with a fixed frequency of 5 kHz and a single damped sinusoidal component of unknown frequency representing the resonant behaviour. It is clear that the data within the time-window considered is fitted very poorly, and the method fails completely to estimate the resonant behaviour. The second is Prony's method with the same model but accounting for the data errors in the manner described in Section 4. This time we obtain good information about the resonant behaviour and the steady-state amplitude. The third method is Prony's method using a model now with five damped sinusoidal components and a strategy that uses the singular value decomposition to select a solution from within a suitable subspace of the model space. The data within the time-window is not fitted as well, but a good estimate of the steady-state amplitude is obtained. Finally, the fourth figure shows the computed solution for the nonlinear estimation problem using the second fit to provide starting estimates.

Finally, we present in Figure 4 the transmission voltage responses for the ITC1001 projector calculated using estimates of the voltage response at each frequency obtained by (i) examining directly the steady-state response, (ii) modelling the measurements from a time-window comprising m = 112 points, and (iii) modelling the measurements from a time-window comprising m = 56 points. In each case, the decision to accept an estimate of the steady-state amplitude for a particular frequency was made as objective as possible. For example, the decision was based on how well the data within the time-window considered was fitted; a fitted model was only accepted if it satisfied the convergence criteria implemented for the non-linear estimation problem and, wherever possible, the model was chosen to reflect our belief that only one resonance is present in the system.

SIGNAL MODELLING IN REVERBERANT LABORATORY TANKS

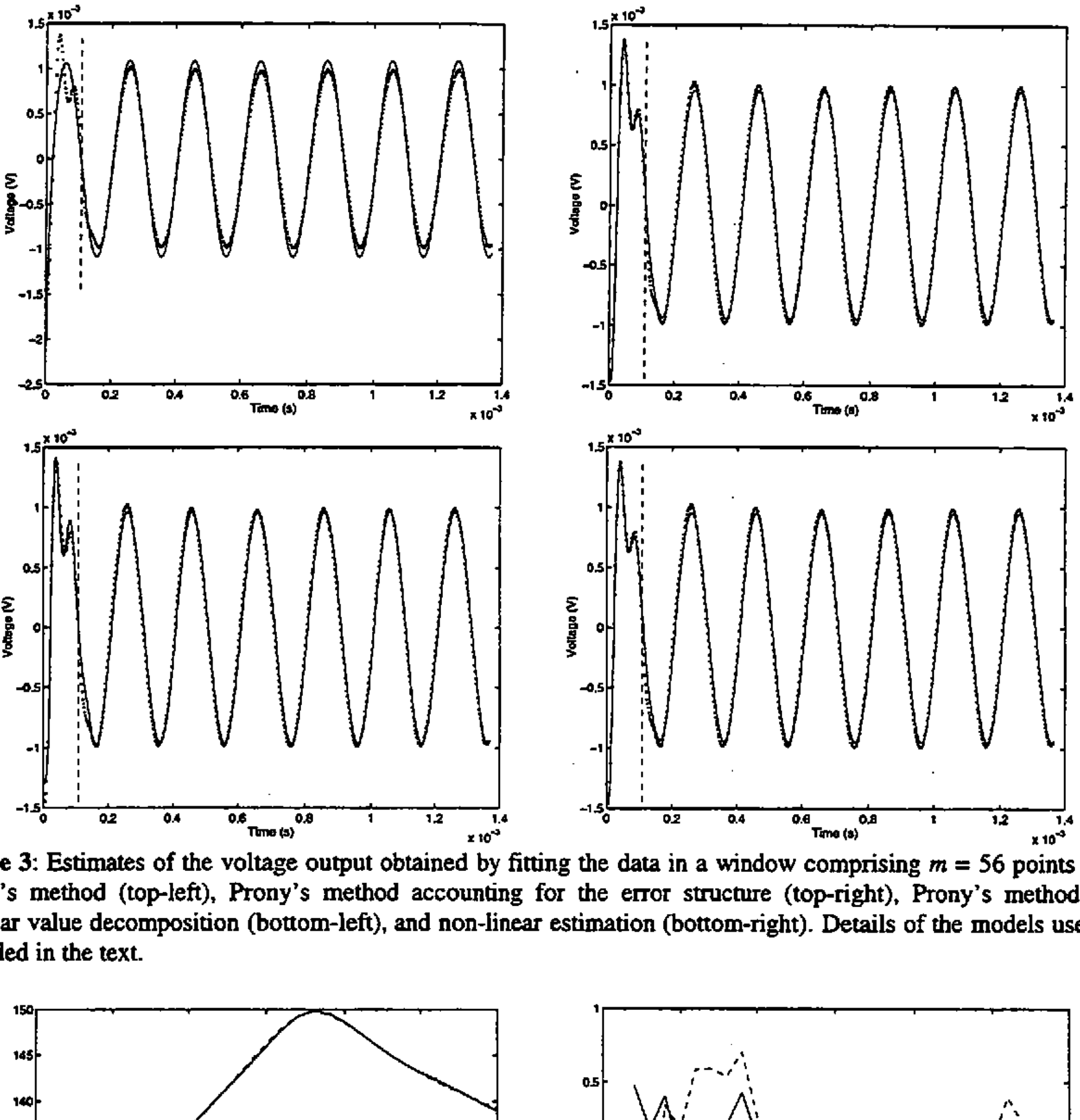

Figure 3: Estimates of the voltage output obtained by fitting the data in a window comprising m = 56 points using Prony's method (top-left), Prony's method accounting for the error structure (top-right), Prony's method with singular value decomposition (bottom-left), and non-linear estimation (bottom-right). Details of the models used are provided in the text.

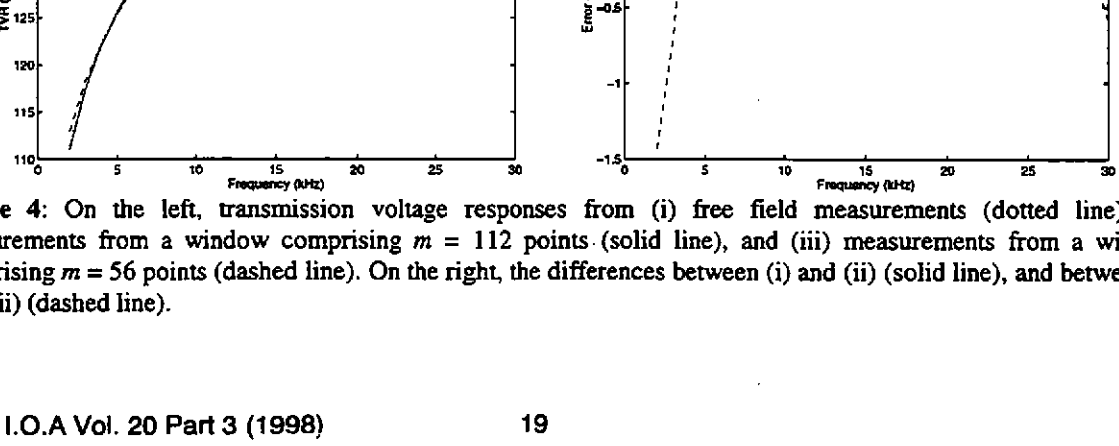

Figure 4: On the left, transmission voltage responses from (i) free field measurements (dotted line), (ii) measurements from a window comprising m = 112 points (solid line), and (iii) measurements from a window comprising m = 56 points (dashed line). On the right, the differences between (i) and (ii) (solid line), and between (i) and (iii) (dashed line).

7. CONCLUSIONS

A comparison has been made between linear prediction methods and nonlinear estimation methods for modelling the free-time signal from an acoustic transducer in order to predict the steady-state response where this cannot be observed directly. Linear prediction methods have been shown to be statistically biased and inefficient, and this has lead to the development of an approach in which proper account is taken of the error structure in the data. A nonlinear least-squares algorithm has also been described. Results have been presented of using these methods for estimating the transmitting voltage response of an 18 kHz projector using time-windows of only a few cycles of the resonance frequency. Agreement with the transmitting voltage response measured under free-field steady-state conditions is achieved to within 0.5 dB at all but the extreme of the frequency range.

8. ACKNOWLEDGEMENTS

The authors would like to thank Mr Geraint Green for undertaking some of the measurements reported here, and to acknowledge the support and funding of the National Measurement System Policy Unit of the UK Department of Trade and Industry.

9. REFERENCES

[1] AINSLEIGH, P.L. and GEORGE, J.D. Variable projection methods with application to sums of exponentials in white noise. *NRL Memorandum Report 6643*, 1990.

[2] AINSLEIGH, P.L. and GEORGE, J.D. Full Newton and constraint methods for semilinear signal modeling. *IEEE Transactions on Signal Processing*, 1993, **41**, 1689-1692.

[3] AINSLEIGH, P.L. and GEORGE, J.D. Signal modeling in reverberant environments with application to underwater electroacoustic transducer calibration. *J. Acoust. Soc. Am.* 1995, **98**, 270-279.

[4] BEATTY, L.G., GEORGE, J.D and ROBINSON, A.Z. Use of the complex exponential expansion as a signal representation for underwater acoustic calibration. *J. Acoust. Soc. Am.*, 1978, **63**, 1782-1794.

[5] GEORGE, J.D. and AINSLEIGH, P.L. A constrained Newton algorithm for maximum likelihood modeling of sums of exponentials in noise. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1990, 5, 2443-2446.

[6] GEORGE, D.J. and GOODMAN, D.M. Application of the output error system identification method to the calibration of underwater acoustic transducers. *Underwater Acoustic Data Processing*, Kluwer Academic Publishers, 1989, 39-46.

[7] GEORGE, J.D., JAIN, V.K. and AINSLEIGH, P.L. Estimating steady-state response of a resonant transducer in a reverberant underwater environment. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1992, 5, 1988, 2737-2740.

[8] GILL, P.E., MURRAY, W. and WRIGHT, M.H. *Practical Optimization*. Academic Press, 1981.

[9] GOLUB, G.H. and PEREYRA, V. The differentiation of pseudoinverses and nonlinear least squares problems whose variables separate. *SIAM J. Numer. Anal.*, 1973, **10**.

[10] MARPLE, S.L. *Digital Spectral Analysis with Applications*. Prentice Hall, Englewood Cliffs, New Jersey. 1987.

[11] OSBORNE, M.R. and SMYTH, G.K. A modified Prony algorithm for exponential fitting. *SIAM J. Sci. Comput.*, 1995, **16**, 119-138.

[12] TRIVETT, D.H. and ROBINSON, A.Z. Modified Prony method approach to echo-reduction measurements. *J. Acoust. Soc. Am.*, 1981, **70**, 1166-1175.

[13] TUFTS, D.W. and KUMARESAN, R. Estimation of frequencies of Multiple Sinusoids: Making linear prediction perform like maximum likelihood. *Proceedings of the IEEE*, 1982, **70**, 975-989.