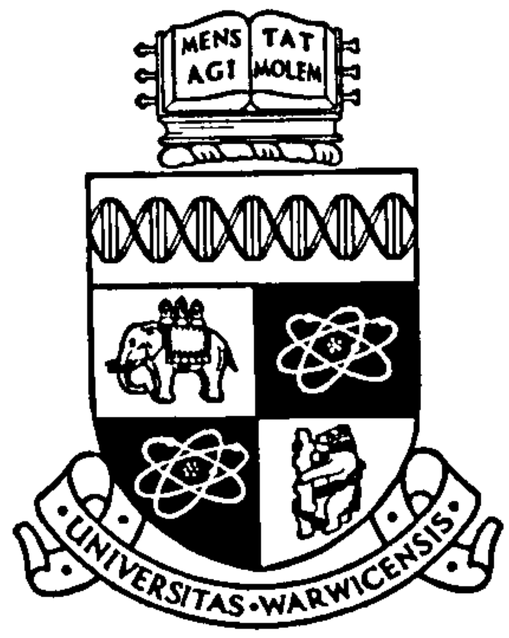

The Design of Periodic Excitations for Dynamic System Identification

Anthony Sean McCormack

University of Warwick

Department of Engineering

Submitted in accordance with the requirements of the Doctor of Philosophy degree

March 1995

To Mum, Dad, Caroline and Emer

Contents

List of Figures

List of Tables

Acknowledgements

Special thanks are extended to Professors Keith Godfrey and John Flower, for providing the impetus for this work, and offering guidance and assistance during its completion.

Financial support by the British Science and Engineering Research Council (now the British Engineering and Physical Sciences Research Council), under Grant GR/H 14267, is also gratefully acknowledged.

I am indebted to my Parents, my Sister Caroline, and Emer for much understanding and encouragement.

Declaration

The following material has been included in the thesis.

Chapter 1 :

A. S. McCormack, K. R. Godfrey and J. O. Flower. The design of multilevel multiharmonic signals for system identification. *IEE Proceedings Part D*, 1995. Accepted for publication.

A.S. McCormack, K. R. Godfrey and J. O. Flower. The design of multilevel multiharmonic signals. In *IEE International Conference "Control 94"*, pages 520-525, University of Warwick, 1994.

A.S. McCormack, K. R. Godfrey and J. O. Flower. The detection of and compensation for nonlinear effects using periodic input signals. In *IEE International Conference "Control 94"*, pages 297-302, University of Warwick, 1994.

Chapter 2 :

K. R. Godfrey, A. S. McCormack and J. O. Flower. Applying system identification using commercially available software and hardware. In *10th IFAC Symposium on System Identification*, pages 3.119–3.124, Copenhagen, 1994.

K. R. Godfrey, A. S. McCormack and J. O. Flower. Applying system identification using commercially available software and hardware. *Control Engineering Practice*, 1995. Accepted for publication.

Chapters 3 and 4 :

J. O. Flower, A. S. McCormack and K. R. Godfrey. Frequency domain identification of processes with output drift. In *IEEE Instrumentation and Measurement Technology Conference "IMTC/93"*, pages 252-257, Irvine. 1993.

J. O. Flower, A. S. McCormack and K. R. Godfrey. Reference phases of low peak factor multiharmonic signals. *Electronics Letters*, 29(2): 144-145, 1993.

A. S. McCormack, J. O. Flower and K. R. Godfrey. The suppression of drift and transient effects for frequency-domain identification. *IEEE Transactions on Instrumentation and Measurement*, 43(2): 232-237, 1994.

Chapter 5 :

A. S. McCormack, K. R. Godfrey and J. O. Flower. Periodic signal designs for nonlinear system identification. *IEEE Instrumentation and Measurement Technology Conference "IMTC/95"*, Waltham, Massachusetts, 24-26 April, 1995, pp. 265-270.

Summary

System identification techniques are developed for modelling linear and nonlinear systems. The main results of the work are concerned with the design and utilisation of periodic perturbation signals in general areas of time- and frequency-domain system identification. A design strategy is given for a new class of perturbation signals, together with examples of their use in system identification applications. Signal processing procedures are developed for the practical treatment of drift disturbances and transient effects, and also for the detection of nonlinear contributions to the measurement data. The techniques rely completely on the periodicity of the excitation, and so the advantageous properties of periodic input signals are considered in detail. The use of periodic excitations in discrete- and continuous-time nonlinear system identification is also reported, with the identification methods illustrating the worth of frequency-domain measurements in this area. An automatic tuning procedure for PID controllers is also developed, which illustrates an application of system identification techniques to control problems.

Introduction to the Thesis

The problems encountered when characterising a system using measurements of its input and output behaviour have led to many decades of research effort. Solving these problems continues to be a fundamental task in a broad range of engineering research activities. The work reported here covers a wide area of this subject, but makes the underlying assumption that a specific class of external stimuli, namely a periodic signal, can be applied to the system during the identification experiment. The postulation of periodic excitations and their favourable properties is by no means a novel topic, and indeed their application among practitioners of identification is reasonably well established. However more recent advances and directions in periodic excitation design have either been acknowledged with minimal interest by identification aficionados, or have been inexorably linked to non-parametric identification, even by practitioners active in the field.

While it is true that periodic signal design and frequency-domain identification form an excellent partnership, and also that frequency-domain identification relies on period excitations, the utilisation of periodic excitations certainly does not enforce the use of frequency-domain system identification. The coupling of these two components of system identification is often natural since the estimation of the non-parametric frequency response function and/or the coefficients of a Laplace transfer function is the most obvious utilisation of periodic data. However, comparing the situation with the case when pseudo-random signals are used, impulse response estimation via cross-correlation is just a conveniently simple (although very important) means of identification. In this work, it is shown that periodic excitations find many areas of application in system identification, sometimes allowing considerable simplifications to be made in the problem formulation, and always allowing measurements to be performed at increased signal-to-noise ratios.

The techniques developed during the course of this research and reported in Chapters 1 through 7 are directed at a series of practical problems commonly encountered in applying system identification techniques. Apart from the studies made in Chapters 5 and 6, the end product of the identification strategy is a linear model description, and to this end, identification algorithms now available as commercially available software are utilised. Using these established techniques, it is believed that the procedure of selecting a linear model from within a family of such models is something that can be achieved with a high degree of confidence, when good quality, representative input and output data is used. This use of established identification algorithms, (both time- and frequency-domain techniques) has naturally led to a more objective view of the benefits of periodic signal design.

Individual chapters describe a distinct area of application and, as such, may be read quite independently of all other material. This said, it is easy to envisage a single application which would encapsulate all of the problems considered in each of the chapters. All the identification techniques, proposed throughout the thesis, may be used in a coherent manner, and it is this that binds the work together. Frequent use is made of simulation and experimental examples, both to illustrate the concepts, and the accuracy of the proposed methods. A summary of the contents of each chapter now follows.

The opening chapter examines the design of perturbation signals which can approximate an arbitrary Fourier specification, and have a time-domain realization containing only a small number of amplitude levels. This chapter also serves as introductory material, since the complete design of the excitation is treated, while in examining the properties of the excitation, the main classes of multiharmonic signal are compared. Along with the selection of the power spectrum and the design of the multilevel signal, the suppression of energy from sets of harmonics is also considered. This finds immediate application when the system under investigation is

likely to exhibit nonlinear effects (whether for detection or modelling), but it is seen in Chapters 3 and 4, that suppressing harmonics from the excitation spectrum, can also facilitate the treatment of drift and transient effects. This chapter also includes material on the assessment of optimal power spectra and the use of the coherence function when used with periodic input signals.

Chapter 2 describes several applications of identification techniques that have made it to the consumer market. After a discussion of the various identification strategies treated in the chapter, several application examples are given. Each application involves a comparison of the results achieved with the different identification methods. From the comparison of the software tools, it is seen that there is little difference between the accuracy of the results obtained with quite different parameter estimation methods. Therefore, a sounder basis for comparison is the intended use of the model, and the experimental situation under which the measurements are performed. The applications involving the hardware systems illustrate the necessity to pay attention to details other than the parameter estimation, when a model is to be estimated from real data. The results from one of the hardware systems are wide-open to misinterpretation, due to the lack of appropriate structure detection facilities.

The influence of drift disturbances is addressed in Chapter 3. This form of disturbance can be the dominant source of error in system identification, tending mainly to distort low frequency measurements, and so it is essential to remove its effects before carrying out any system parameter estimation. Drift estimation schemes, which are proposed as preprocessing methods, are shown to be applicable to the commonly encountered forms of drift, i.e. a disturbance which can be modelled using a low-order polynomial. The advantage of using a periodic excitation in this instance is that an arbitrarily complex drift disturbance may be modelled during the estimation of the Fourier coefficients. If an excitation with a specific type of power

spectrum can be employed, a particularly simple drift estimator results. Although the drift estimation is carried out in the frequency-domain, it is not necessarily assumed that the system parameter estimation is carried out likewise. The accuracy of the estimation procedures is assessed throughout the chapter, and it is shown that the simplified estimator performs well when compared to the more complex coupled estimation procedure. Along with the assumption of a periodic excitation is the implicit assumption that no transient effects are present in the estimation data. Identification in the presence of drift when this is not the case is treated in Chapter 4. However, a signal design procedure is given in Chapter 3, for situations where the transient problem manifests itself in the form of a linear drift, i.e. applications involving type 2 systems.

Signal design and processing procedures which can be used to diminish the effects of initial transients are reported in Chapter 4. This material is included mainly to alleviate the problem of wasted experimental time when using periodic signals with frequency-domain identification. The techniques are therefore most relevant in application areas where the increased experimental time involved in waiting for transient effects to decay, before acquiring a full period of steady-state data, is a noticeable disadvantage of using periodic excitations. The first solution proposed in Chapter 4, is applicable to systems that have a single long time constant (known *a priori*), while the second method is applicable to all stable systems. The main result of developing these techniques is that the practitioner will have the confidence to process the first period of the response to the excitation in the normal way, e.g. spectral analysis or parametric model identification. One of the methods uses an alternative algorithm instead of direct use of the discrete/fast Fourier transform. System identification using a single period of the excitation may suffer if significant external disturbances corrupt the measurements. However the accuracy of the proposed truncation method is shown to be comparable to using the DFT with

steady-state records, in all but the harshest tests of the algorithm. The coupling of the truncation method to the drift estimation technique reported in Chapter 3 is discussed and illustrated using a realistic identification example.

In Chapter 5, an assessment is made of the performance of several periodic signal designs, when they are used in estimating the parameters of a nonlinear model. To obtain an objective assessment, discrete-time identification methods and models are considered. Three different types of excitation power spectrum are used, with a random excitation included for comparison purposes. It is seen that if a nonlinear model is to be estimated, it is advantageous to apply a periodic signal with an extremely rich harmonic content. This type of signal can be easily designed using multi-frequency signal design algorithms, already developed for linear system identification.

The first part of Chapter 6 discusses both the impact of using periodic excitations when nonlinearities are present in the system, and common ways to control the resulting errors when a linear model is to be estimated. The remainder of Chapter 6 develops signal design procedures which are relevant only to specific methods of frequency-domain identification. This material is included to highlight the areas where periodic excitations simplify the identification of certain models of nonlinear systems. The identification methods, discussed in Chapter 6, are relevant to the three main areas of nonlinear identification: functional series expansions; block-structured systems; and parametric methods. The reliance of all the techniques on frequency-domain measurements (and admittedly in some cases on some very peculiar signals) is kept mainly to explore the potential of nonlinear system identification in the frequency-domain. The advantages, disadvantages and limitations of using frequency-domain measurements in this application are discussed. The identification methods are illustrated using both simulated and experimentally measured data.

Chapter 7 is an application of system identification methods for control system design. The consideration given to the PID controller is owing to its almost universal acceptance in industry, and also the renewed attention it is receiving in the literature. The automatic tuning of the PID controller parameters generally involves the estimation of a limited part of the frequency response of the system, followed by the application of a simple tuning rule, many of which have been postulated in the literature. The application of multi-frequency signal techniques for the task of frequency response estimation is immediately apparent. However, the work reported in Chapter 7 covers only the bare skeleton of the research needed to produce a robust autotune PID controller system, i.e. attention is only paid to the signal design and frequency response estimation method. However these alone are enough to illustrate the efficacy of the proposed solution to the autotune problem. The rules considered for calculating the PID parameters, include all those that are known to the author. However, a rapprochement is necessary between the broadband identification results and the tuning of the controller, since at the moment the tuning is based on knowledge of the system dynamics at a single frequency.

Chapter 8 summarises the main results reported throughout the thesis and discusses areas of future research.

Chapter 1

The Design of Multilevel Multiharmonic Signals for System Identification

1.1 Abstract

The chapter begins with a review of the existing multifrequency signals available for linear system identification in the frequency-domain. The motivation for an excitation with a specified Fourier amplitude spectrum, yet having a small number of signal levels, is then outlined. An algorithm is given for the design of multilevel signals, which realises an arbitrarily defined set of Fourier specifications. It is shown that the properties of the multilevel signals make them a strong candidate for linear system identification, as well as for the detection of nonlinear effects.

1.2 Introduction

The selection of the power spectrum (Goodwin and Payne, 1977) and the constraints on the form of the excitation (Godfrey, 1993c) are two important issues which need to be considered during the design of an experiment for system identification. For linear system identification, periodic excitation signals offer many advantages (Schoukens, Pintelon and Van hamme, 1994), with design algorithms available to optimise the power spectrum (Van den Eijnde and Schoukens, 1991), and to produce good time-domain realisations (Schroeder, 1970; Van den Bos and Krol, 1979; Van der Ouderaa, Schoukens and Renneboog, 1988a; Guillaume, Schoukens, Pintelon and Kollár, 1991). The complete design of the excitation is usually a two-step process (Schoukens, Guillaume and Pintelon, 1993a), involving power spectrum selection/optimisation, followed by the optimisation of an arbitrary realisation of this spectrum, taking into account the inevitable amplitude constraints present on the input.

The two classes of signal most commonly used to realise the specified power spectrum are the multisine (sum of harmonics (SOH)), and the discrete-interval

binary signal (DIBS). This chapter discusses the properties of multilevel (m-level) signals which are shown to offer an intermediate level of approximation to the desired power spectrum, when compared to the DIBS and multisine signals. These are useful when a signal with a large number of amplitude levels cannot be accurately applied to the system under investigation (Chen, Kerlin and Fry, 1972; Van den Bos and Krol, 1979; Franck and Rake, 1985). In these situations the properties of the DIBS excitation, although excellent in terms of the realisation of the desired power spectrum, may be unsuitable from a practical (Harris, 1993), or theoretical (Nowak and Van Veen, 1994) standpoint. The algorithm proposed here to design the m-level signals is similar to that developed for the design of DIBS signals (Van den Bos and Krol, 1979), and the clipping method used for the design of multisines (Van der Ouderaa et al., 1988a).

Due to the strong coupling between the nature of the desired power spectrum and the ability of the design algorithms to realise it, simple guidelines for the selection of the power spectrum are given in Section 1.3, that are appropriate in situations where approximate knowledge of the system dynamics is available. The results achieved for the design of DIBS, m-level and multisine excitations, for an illustrative system, are compared in Section 1.3.1, and on a more general note, the ability of the m-level signal design algorithm to match the spectral properties achieved with DIBS and multisine signals, is considered in Section 1.4.1.

The final part of the chapter discusses the extension of the design algorithm to include additional constraints on the power spectrum of the excitation. Suppression of harmonics from the power spectrum (a useful property of multisines), is shown to be possible with the m-level signals in a limited, but useful, way. This can be useful for the estimation of drift effects (McCormack, Flower and Godfrey, 1994a), and the detection of nonlinear effects (McCormack, Godfrey and Flower, 1994b). It is possible to suppress harmonics from the power spectrum of multilevel pseudorandom

signals, derived from maximum length sequences (Barker, 1993), but due to the nature of the codes, stringent conditions are enforced on the shape of the power spectrum, and also on the number of levels necessary in the signal.

1.3 Power Spectrum Synthesis

Before detailing the m-level design procedure, the selection of the power spectrum is discussed. For practical identification, the lack of good *a priori* information about the system under investigation can sometimes be a problem, meaning that optimisation of the input signal is rarely carried out (Norton, 1986). However, in this section, knowledge of the system parameters is assumed, so that an assessment can be made of the coupling between the power spectrum and time-domain optimisation procedures.

The example that is used for illustration throughout Section 1.3.1 is the same as that used by Schoukens and Pintelon (1991, section 4.3.5), with the emphasis here on the m-level signals.

1.3.1 D-Optimal Designs

Experimental design which minimises the determinant of the estimated parameter covariance matrix, (D-optimal designs), is considered here. As is usual, the covariance matrix is approximated by the Cramér-Rao lower bound, for which an explicit expression is given by Schoukens and Pintelon (1991). The design algorithm used is due to Van den Eijnde and Schoukens (1991), and is implemented in Kollár (1994). The quantity examined here is the response dispersion function, given by,

$$\nu(\chi, \jmath\omega) = \text{trace}\{[\text{Fi}(\chi)]^{-1}\text{fi}(\jmath\omega)\} \tag{1.1}$$

where,

$$\chi(\jmath\omega) = (|X(\jmath\omega_1)|^2 \cdots |X(\jmath\omega_K)|^2) \text{ with } \sum_{k=1}^{K} |X(\jmath\omega_k)|^2 = \mathcal{P}. \tag{1.2}$$

$\mathrm{Fi}(\chi)$ and $\mathrm{fi}(\jmath\omega)$ are the Fisher information matrices corresponding to the spectra $\chi(\jmath\omega)$ and $X(\jmath\omega)$, respectively, with $\mathcal{P}$ representing the signal power, and $|X(\jmath\omega)|^2 = \mathcal{P}$. Here $\mathrm{Fi}(\chi)$ will be calculated from the spectrum that has been realised with each time-domain realisation.

If the system can be represented as an output-error model, with disturbing noise variance $\sigma_y^2(\jmath\omega)$, the corresponding model uncertainty, $\sigma_H(\jmath\omega)$, is given by,

$$\sigma_H(\jmath\omega) = \sqrt{\frac{\nu(\chi, \jmath\omega)\sigma_y^2(\jmath\omega)}{\mathcal{P}}} \tag{1.3}$$

and so to simplify the considerations, $\sigma_y^2(\jmath\omega)$ is assumed to be equal to unity for all frequencies. The design procedure generates, at each iteration, an improvement on the previous power spectrum, assuming equal power for each design. This is never the case if multiharmonic signals are used with an amplitude constraint present on the system input, and so to assess the actual improvement at each iteration, the power spectrum, matching the amplitude constraint, must be produced. This is a rather tedious task since the respective design algorithms use an iterative procedure to ensure that a near-optimal signal is found. To combat this, the approach adopted here is to use Schroeder phases (Schroeder, 1970), to design each signal. For the DIBS and m-level signals this corresponds to mapping the Schroeder multisine to the amplitude levels of both signals. This one-step procedure gives a good indication of the final value of the uncertainty (assuming correct prior knowledge), since the results from Schroeder phasing are usually indicative of the results which will be produced with the computationally intensive iterative design procedures.

The system that is used for illustrative purposes (see also Kollár (1994)) is a band-pass system with fourth-order numerator, $\alpha_2 s^2 \cdots \alpha_4 s^4$, and sixth-order de-

nominator, $\beta_0 s^0 \cdots \beta_6 s^6$, with the system coefficients given by ((Schoukens and Pintelon, 1991, pp. 179)),

$$(\alpha_2, \alpha_3, \alpha_4) = (8.973\text{e-}10,\ 5.155\text{e-}12,\ 3.2010\text{e-}17)$$

$$(\beta_0 \cdots \beta_6) = (1, 2.5017\text{e-}4,\ 3.5869\text{e-}7,\ 5.5550\text{e-}11,$$

$$3.36031\text{e-}14,\ 2.5351\text{e-}18,\ 1.0131\text{e-}21). \qquad (1.4)$$

An optimal design was carried out for a three-level signal (levels: -1 0 1) designed to have power concentrated in the first 50 harmonics of the fundamental (20Hz). The results are given in Figure 1.1, which show the model uncertainty produced after 1, 2, 3, 10 and 100 iterations of the design algorithm. Also shown are the results produced with the Schroeder phased DIBS and multisine excitations. The determinants (D), of the corresponding covariance matrices are shown in Table 1.1, together with the respective crest factors (CF). The spectra obtained after 100 iterations of the

Iterations	*Binary*		*3 Level*		*Multisine*	
	CF	D	CF	D	CF	D
0	1.09	104.247	1.175	375.28	1.713	$0.037\text{x}10^7$
1	1.103	0.235	1.215	1.219	2.069	$0.0018\text{x}10^7$
2	1.114	0.16	1.151	0.266	2.001	$0.0006\text{x}10^7$
3	1.11	0.139	1.144	0.231	1.962	$0.0003\text{x}10^7$
10	1.108	0.13	1.20	0.453	2.137	$0.0012\text{x}10^7$
100	1.126	0.16	1.245	0.914	3.183	$1.19\text{x}10^7$

Table 1.1: Crest factor and determinant of the covariance matrix for Schroeder phased binary, three-level and multisine excitations. Iteration 0 corresponds to a uniform power spectrum

optimal design algorithm, using the one-step design procedure for the excitations, are shown in Figure 1.2, with the corresponding dispersion functions shown in Figure 1.3. Note that $\max\{\nu(\chi, \jmath\omega)\} \leq$ number of estimated parameters corresponds to D-optimality. In Figures 1.2 and 1.3, the power in the desired harmonics has

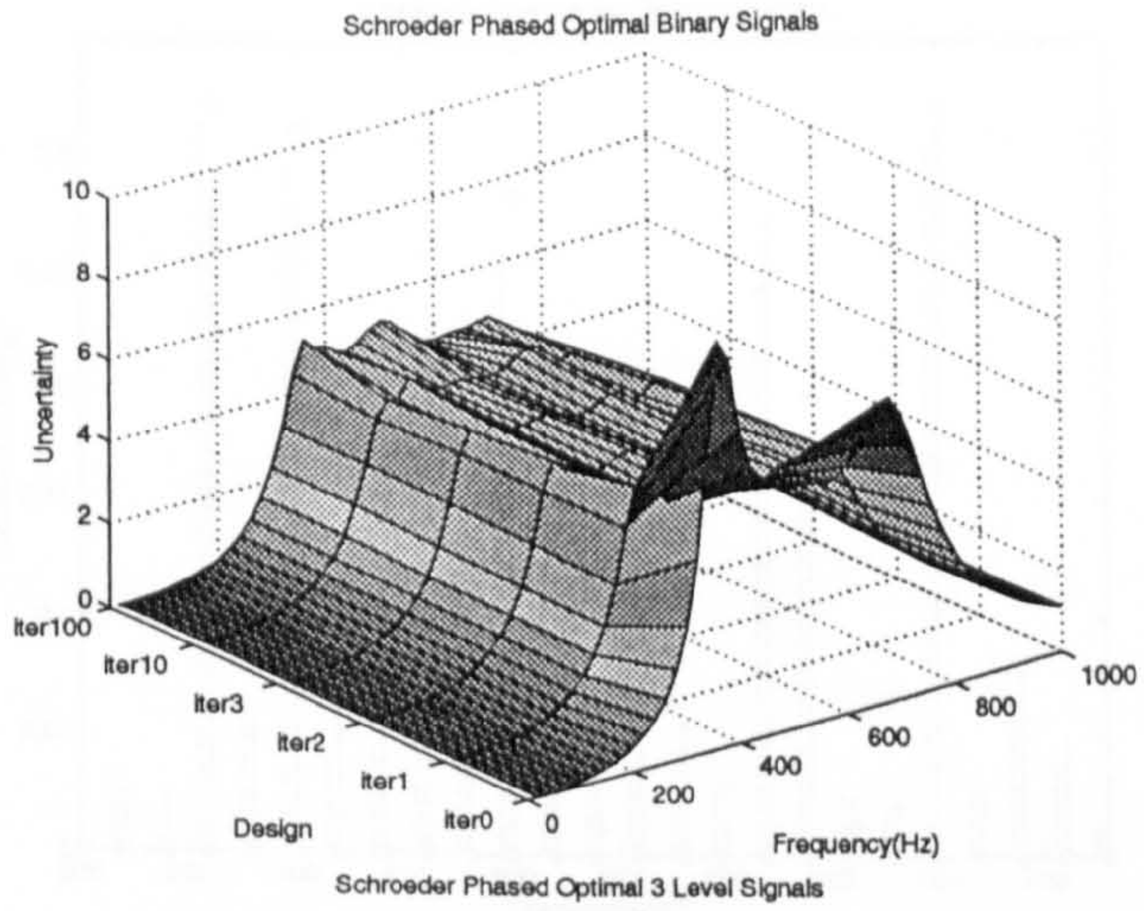

Figure 1.1: Uncertainty of the frequency response functions obtained using binary, three-level and multisine excitation signals.

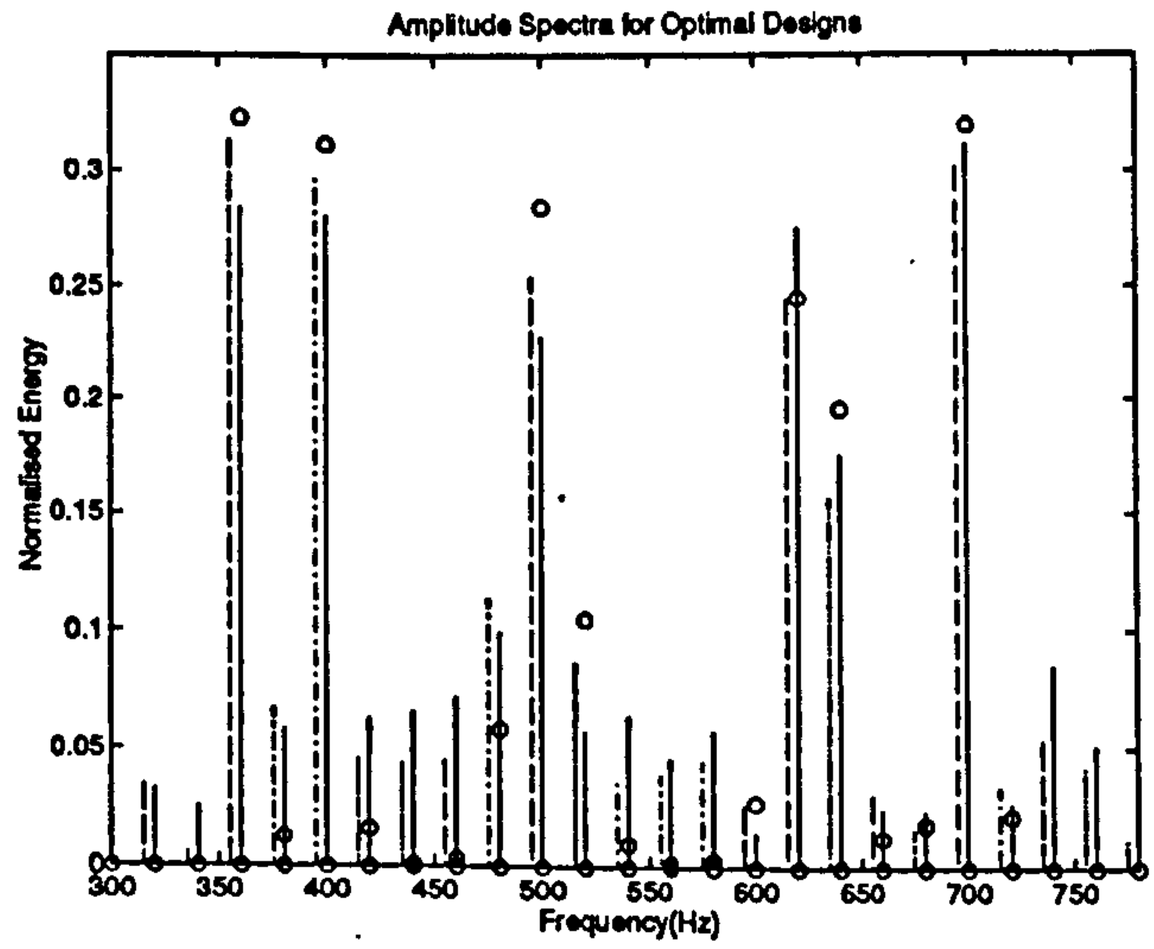

Figure 1.2: Absolute value of the normalised amplitude spectrum obtained after 100 iterations of the optimisation algorithm: binary (full), three-level (dashed) and multisine (circles). The frequency range has been limited to the band at which the desired spectra is non-zero

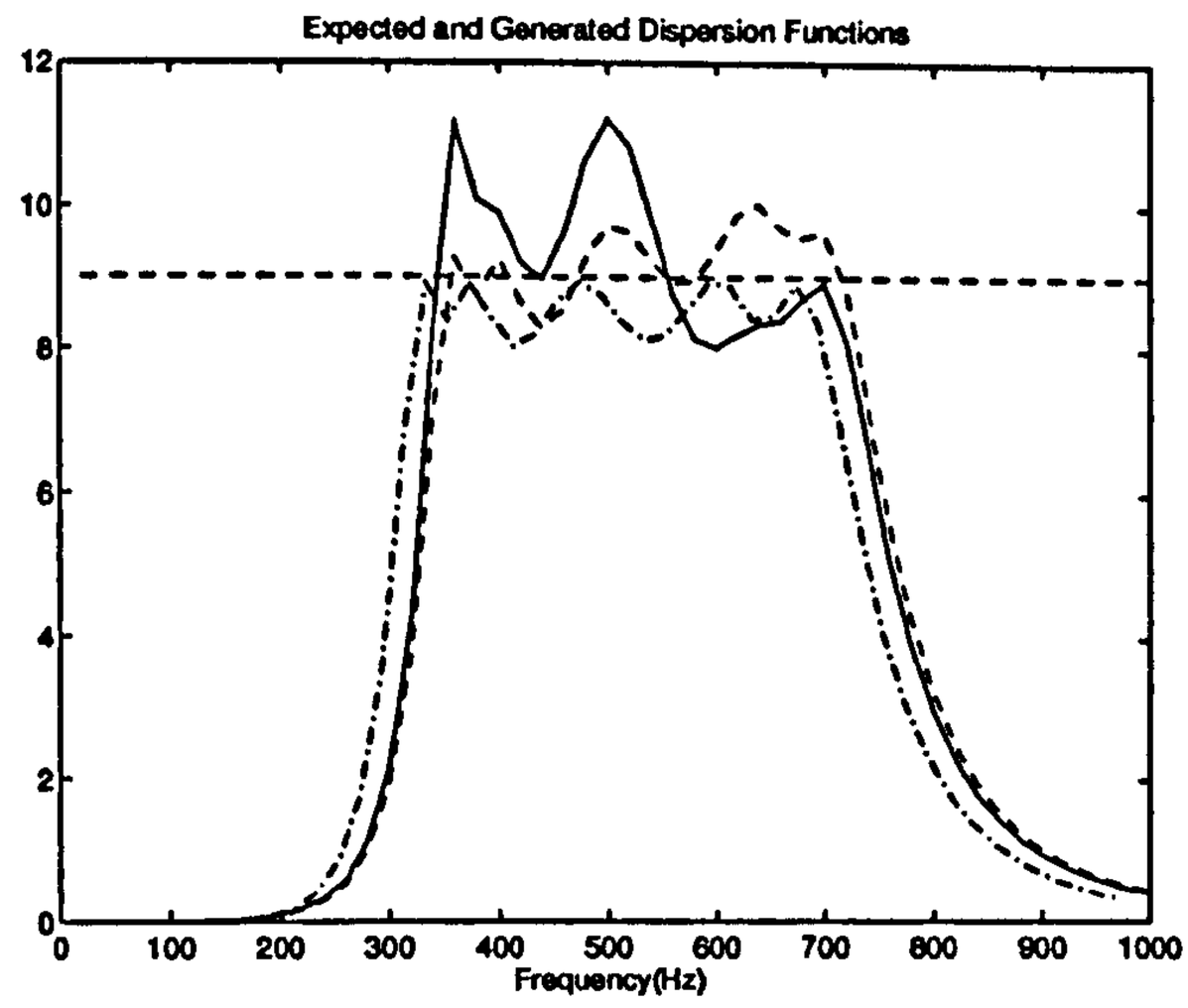

Figure 1.3: Dispersion function after 100 iterations of the optimisation algorithm: binary (full), three-level (dashed) and multisine (dash-dot). The minimum of the dispersion function over the frequency grid is also shown.

been normalised to aid visualisation of the "fit" of each realisation to the desired spectrum. This normalisation is not present in Figure 1.1.

The results given in Table 1.1 show a strong correspondence among the three signal types with, as stated by Schoukens et al. (1993a), the designs produced after the initial iterations of the algorithm yielding the most significant decrease in the value of the determinant. However, this has been obtained in this work by simple use of the Schroeder design formula, thus drastically reducing the computational time. The other interesting property displayed by the DIBS and m-level signals is their ability to realise the desired spectra (across 50 harmonics) to a high level of accuracy. The results achieved with the uniform power spectrum, and that produced after one iteration of the design algorithm need careful attention, since these are typical of the spectra which will most commonly be used in the majority of cases. The former spectrum usually results from a heuristic basis, where good coverage of the system dynamic range is desired, while the latter simply corresponds to concentrating the signal power crudely around the dominant modes of the system.

The importance of the power realized in the desired harmonics can be seen by comparing the results of Figure 1.1 and Figure 1.3. Even though the multisine spectrum is the perfect "fit" to the desired spectrum (Figure 1.3), the lower crest factors of the DIBS and m-level signals produce the best overall results (the DIBS and m-level signals with uniform spectra produce better results than the "optimum" multisine). Schroeder phased multisines are used purely to assess the relative improvement achieved for each design. It is acknowledged that the actual results for each spectra will be improved with the use of iterative design procedures, however the relative improvements will remain the same. The inclusion of snow during the design of the multisine will give crest factors as low as the DIBS and m-level signals.

Hence in terms of the accuracy of the experiment, the spectra realised by the m-level signals can give results which are better than those obtained with the mul-

tisine signal, and similar to those obtained with both the DIBS and multisine with snow signals. This means that the relevant question guiding the selection of the time-domain properties of the signal, is whether a bandwidth-limited or piecewise-constant excitation signal is more desirable for the identification (Schoukens et al., 1994). Given that the m-level signals possess these properties, the specific nature of the system should guide the selection of the most suitable excitation, e.g. the nature of the input transducer, and/or the required measure of the systems nonlinear characteristics.

1.4 Time-Domain Signal Design

The algorithm used to approximate the desired Fourier amplitude spectrum, C_{xk}, ($k \in K$ the set of harmonics) is based on the time-frequency domain method used for conventional multiharmonic signal design (Van den Bos and Krol, 1979; Van der Ouderaa et al., 1988a). The harmonic phases are used to maximise the minimum relative Fourier amplitude, E_{min}, realised in the m-level signal

$$E_{min} = \min\{\frac{|C_{mk}|}{|C_{xk}|}\}.E_{des} \tag{1.5}$$

where C_{mk} is the realised Fourier coefficients, and E_{des} is the total energy specified in the design. The normalisation by E_{des} allows direct comparison of E_{min} to that produced with a DIBS excitation with levels 1 and -1. The mapping in the time-domain is an M-level quantizer with symmetric bands of size Q_i, where,

$$Q_i = \left|\frac{B_i}{q^{i-1}} - \frac{B_{i+1}}{q^i}\right| \; ; i = \begin{cases} 1, \cdots, M/2, & M \text{ even} \\ 1, \cdots, (M-1)/2, & M \text{ odd}, \end{cases} \tag{1.6}$$

where B_i, $i = 1 \cdots M + 1$, result from dividing the input range into M equal segments. The convergence of this algorithm for binary signals has been proven by Van den Bos and Krol (1979). According to the experience of the author, when the (fixed) mapping in the time-domain is selected using eqn. (1.6), convergence is achieved in a small number of iterations. The parameter q needs to be selected for each set of specifications. The approach adopted here is to carry out a series of designs initialised with Schroeder phases, with the value of q chosen as the one which yields the maximum value of E_{min}. This value can then be used to carry out a multiple start procedure, with random harmonic phases used each time as initial conditions. The effect of q on a design requiring a uniform spectrum in harmonics 1 to 10, is shown in Figure 1.4. This procedure was found to give good results for

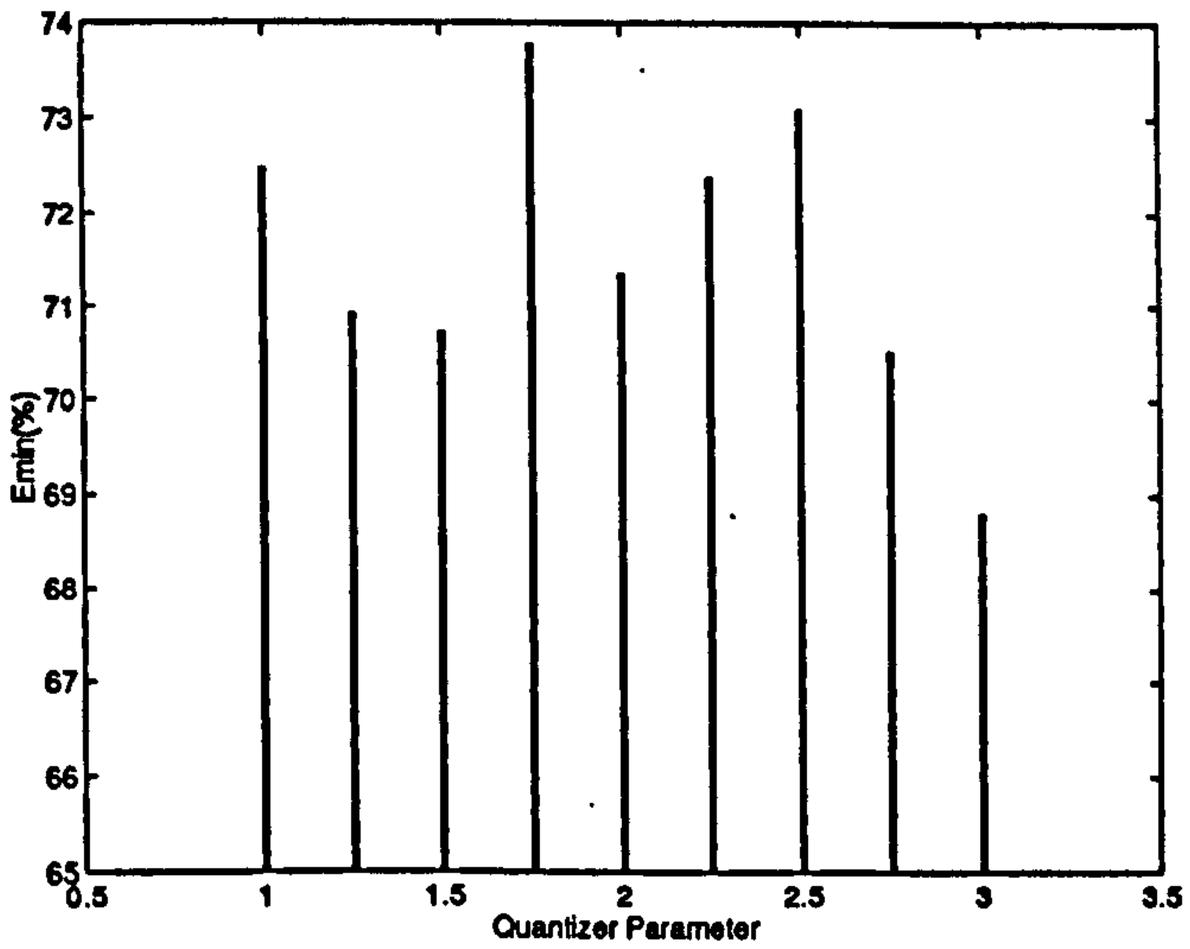

Figure 1.4: Minimum relative harmonic amplitude with changing quantizer characteristics for a three-level signal with uniform amplitude spectrum

all the spectra examined, as detailed in Section 1.4.1.

1.4.1 Comparison of Multifrequency Signal Designs

The results achieved with the m-level design algorithm are now compared with the characteristics of both the DIBS and multisine signals. The multisine signals used for each design are zero-order-hold multisines (ZOHSOH). These usually give better results than the bandwidth-limited designs since the time-domain signal properties need only be optimised at the desired sampling instants. The benefits of ZOHSOH signals are greatest when the minimum number of samples is used to realize the spectrum, i.e. with an over-sampling factor of 1. Three- and four-level signals are considered here, and the properties given are the results obtained after 100 randomised starts of the phase optimisation. The same procedure is adopted for the DIBS signal, whereas each multisine signal is the result after 2000 iterations of the minimisation algorithm with Schroeder's initial conditions. The DIBS and multisine signals were designed using the Frequency Domain System Identification Toolbox (Kollár, 1994). The multisine signals with snow (SOHS) were designed with the l_p norm (Guillaume et al., 1991) being minimised for $p = 2, 4, 6, ..., 200$, starting at $p = 2$ with Schroeder phases. The snow was allowed in all non-excitation frequencies up to the Nyquist frequency. The spectra considered for the comparison are as follows (the oversampling factor for each signal is given after the set of harmonics)

Spectrum 1 Uniform Amplitudes in Harmonics 1 to 10 (6.4).

Spectrum 2 Linearly Increasing Amplitudes in Harmonics 1 to 10 (6.4).

Spectrum 3 Uniform Amplitudes in Harmonics 1,2,4,8,16,32,64 (4).

Spectrum 4 Linearly Increasing Amplitudes in Harmonics 1,2,4,8,16,32,64 (4).

Spectrum 5 Amplitudes 1,1,1,2,3,4 in Harmonics 1,2,4,8,16,32 (2).

The results are summarised in Table 1.2, with the designated spectrum in each case being denoted by S. The signal power, x_P, is quoted for the signals after

S	*Signal*	P_{use} (%)	E_{min} (%)	x_P	q
1	2 Level	79.67	83.18	1	$\gg 1$
	3 Level	**90.72**	**81.3**	**0.83**	**1.75**
	4 Level	**94.54**	**79.42**	**0.77**	**1.25**
	ZOHSOH	100	66.76 $\forall k$	0.45	-
	SOHS	80.06	85.78 $\forall k$	0.92	-
2	2 Level	79.1	80.6	1	$\gg 1$
	3 Level	**89.36**	**80.3**	**0.88**	**2.5**
	4 Level	**91.6**	**79.95**	**0.82**	**1.75**
	ZOHSOH	100	63.38 $\forall k$	0.4	-
	SOHS	81.85	87.21 $\forall k$	0.93	-
3	2 Level	68.3	80.3	1	$\gg 1$
	3 Level	**81.39**	**75.74**	**0.76**	**2.25**
	4 Level	**79.85**	**76.92**	**0.77**	**2.75**
	ZOHSOH	100	46.08 $\forall k$	0.21	-
	SOHS	72.78	83.16 $\forall k$	0.95	-
4	2 Level	71.0	77.6	1	$\gg 1$
	3 Level	**84.29**	**74.57**	**0.72**	**2**
	4 Level	**86.70**	**73.54**	**0.7**	**2**
	ZOHSOH	100	43.5 $\forall k$	0.19	-
	SOHS	72.13	83.22 $\forall k$	0.96	-
5	2 Level	62.7	74.4	1	$\gg 1$
	3 Level	**83.05**	**73.55**	**0.73**	**2.25**
	4 Level	**87.32**	**72.71**	**0.65**	**1.75**
	ZOHSOH	100	49.1 $\forall k$	0.24	-
	SOHS	71.89	80.41 $\forall k$	0.90	-

Table 1.2: Comparison between the properties of the different multifrequency signals

their maximum absolute value has been scaled to unity, while P_{use} represents the percentage of the signal power realized in the desired harmonics. From the table it can be seen that, generally, the characteristics of the DIBS and the SOHS signals are similar in terms of the figures for x_P and P_{use}, while E_{min} is consistently greater for the SOHS signals. Comparing the properties of the three- and four-level signals to those of the DIBS signal, and assuming that a SOHS signal cannot be applied to the system, the values for E_{min} are seen to be comparable for each of the signal designs. This is realized in the m-level signals with as little as two thirds of the power of the DIBS signal, giving a significant reduction in the amount of wasted power being applied to the system.

Hence for *linear* system identification, with no concerns other than the distribution and absolute value of the power realized in the perturbation signal, the m-level signals have very useful properties. Attention is now given to an added advantage of using more than two signal levels in the approximation of the desired Fourier coefficients: harmonic suppression.

1.5 Harmonic Suppression

It is often desirable to leave gaps within the perturbing power spectrum so that energy in the response signal, at non-excited frequencies, can be used to gain information about the system other than the linear behaviour, e.g. for the treatment of drift effects (McCormack et al., 1994a). This problem was originally introduced by Barker (1993), where multilevel pseudorandom sequences were considered. The restrictions of the pseudorandom codes meant that a high number of levels had to be used to suppress any useful harmonics, e.g. a seven-level sequence was needed to suppress multiples of harmonics two and three from the spectrum. Also the designer is offered very little control over the power spectrum of the signal. Therefore this

concept is developed here for the m-level signals. Clearly the limited flexibility of the DIBS signal in this regard makes it unattractive for these applications.

Formally stated, we wish to suppress all multiples of the prime set, $P = \{p_1, \cdots, p_n\}$ from the m-level power spectrum. To enable this, the following equality must hold for the m-level signal, $x_m(r)$:

$$\sum_{n=0}^{p-1} x_m(\tau + \frac{nN}{p}) = 0, \qquad \tau = 0, \cdots, \frac{N}{p} - 1, \quad \forall p \in P. \tag{1.7}$$

with $\min(M) = 3$ if $p_l \geq 3$. The signal length, N, must be an integer multiple of all the members of $\{P\}$, and so the values of N that correspond to 2^n, n an integer, are generally not available. This rules out the use of the radix-2 FFT algorithm, but due to the availability of mixed-radix FFT algorithms, e.g. the chirp-z transform algorithm (Rabiner, Schafer and Rader, 1969), this is not deemed a particular problem. The most common P is $\{2, p_s\}$ with the most useful p_s being 3. This is because values other than 3, or subsequent members of P, cause very few extra harmonics to be suppressed. For example, $P = \{2, 3, 5\}$ causes only three extra harmonics to be suppressed than $P = \{2, 3\}$, when harmonic 53 is considered to be the maximum. For this reason, consideration is limited here to $P = p_s$ and $P = \{2, p_s\}$. To suppress all multiples of harmonics two and three from an m-level spectrum, the signal time-series must obey the usual inverse-repeat property, as well as,

$$x_m(\tau) + x_m(\tau + \frac{N}{3}) + x_m(\tau + \frac{2N}{3}) = 0, \qquad \tau = 0, \cdots, \frac{N}{3} - 1, \tag{1.8}$$

with N an integer multiple of six. The implementation of this can be carried out on an existing m-level signal, but it is found that the resulting value of E_{min} can drop to around 30%, which is clearly unsatisfactory. To get around this problem the signal must be adjusted to satisfy eqn. (1.7) at each iteration of the phase optimisation.

For $P = 3$, it is required to construct, at each iteration,

$$x_{sup}(\tau) = \sum_{n=0}^{2} x_m(\tau + \frac{nN}{3}), \qquad \tau = 0, \cdots, \frac{N}{3} - 1, \tag{1.9}$$

and for $x_{sup}(\tau) \neq 0$, modify $x_m(\tau)$, $x_m(\tau + N/3)$ or $x_m(\tau + 2N/3)$, such that

$$x_{sup}(\tau) = 0 \qquad \forall \tau \tag{1.10}$$

and

$$x_{sup}(r) \leq \max(|V|) \qquad r = 0 \cdots N - 1 \tag{1.11}$$

where V equals the maximum allowable value of the m-level signal. For $P = \{2, 3\}$, the modifications to $x_m(r)$ must not alter its inverse-repeat structure, but because $x_{sup}(\tau)$ is inverse-repeat, for an inverse-repeat $x_m(r)$, this is possible.

As an example, Figure 1.5 shows the results achieved with this algorithm on a specification for a three-level signal containing uniform amplitudes in harmonics 1, 5, 11, 17, 37, 67, 131. For this signal $E_{min} = 78.24\%$, and $P_{use} = 80.34\%$. The spectrum shown in Figure 1.6 has the same specification, but has all multiples of $P = \{2, 3\}$ suppressed. The signal has $E_{min} = 64.82\%$, and $P_{use} = 74.75\%$. The value for E_{min} has clearly been affected by the suppression technique, but compared to the value which would be expected with a multisine signal, the properties of the signal are still attractive.

1.6 Detection of Nonlinear Effects

The estimate of the frequency response function (FRF) gives an excellent means for validating any parametric model estimated from system input/output measurements. A common method for validating the FRF estimate is the coherence function

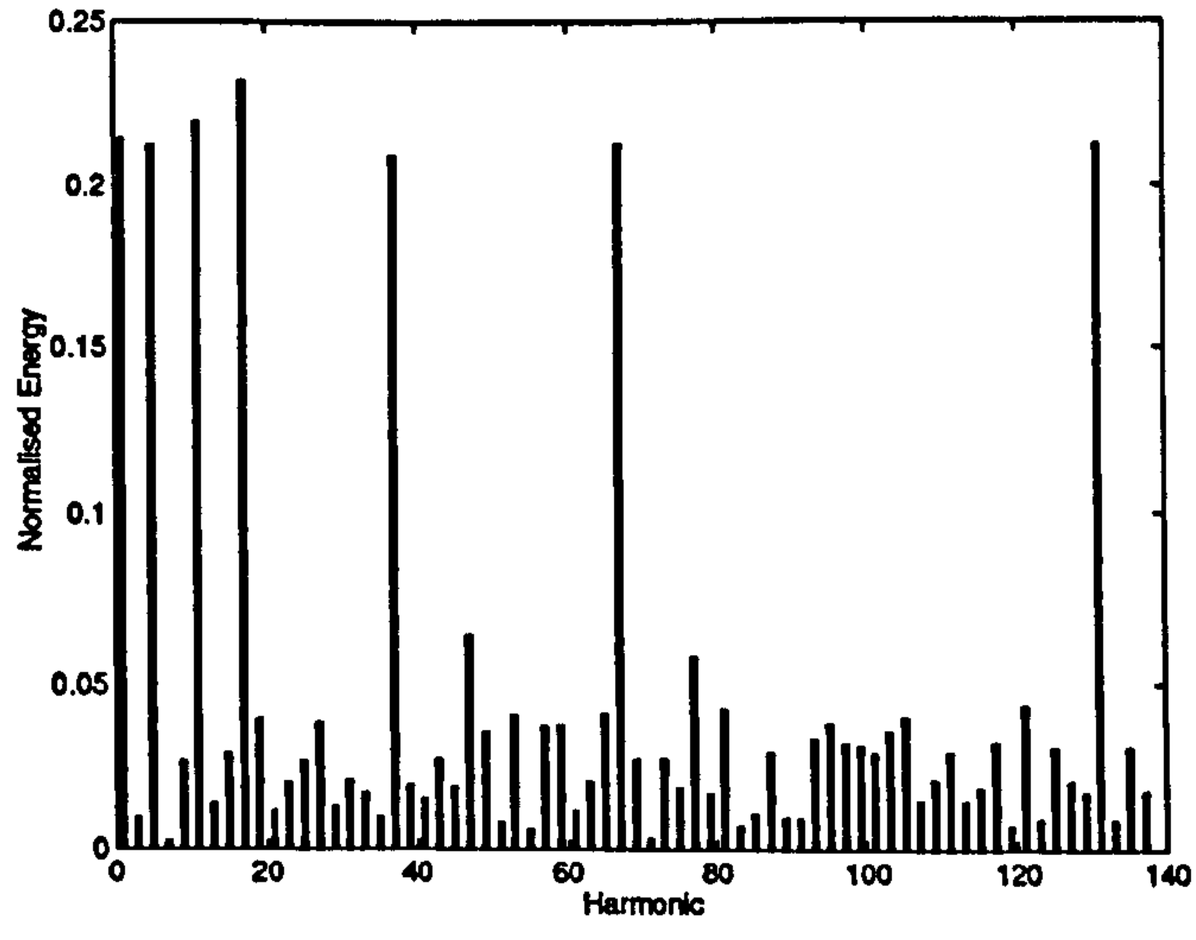

Figure 1.5: Normalised amplitude spectrum of three-level signal specified to have equal energy in harmonics 1,5,11,17,37,67,131

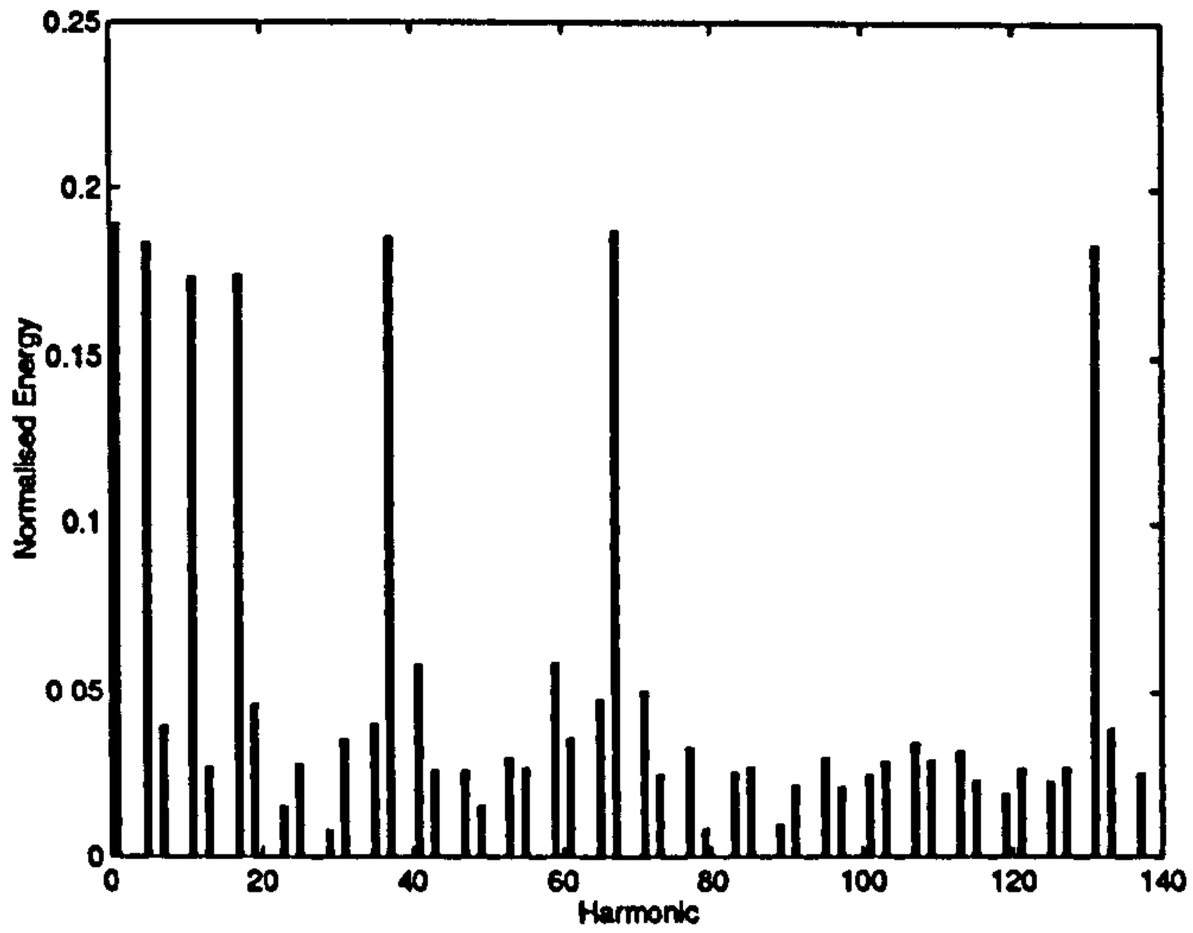

Figure 1.6: Normalised amplitude spectrum of three-level signal specified to have all multiples of harmonics two and three suppressed, with equal energy in harmonics 1,5,11,17,37,67,131

(Bendat and Piersol, 1980), which arose originally in applications involving stochastic time records. The (squared) coherence function between two signals, $x(t)$ and $y(t)$, calculated at the excitation harmonics, is defined as:

$$\gamma_{xy}^2(k) = \frac{|S_{xy}(k)|^2}{S_{xx}(k).S_{yy}(k)}. \tag{1.12}$$

For a periodic signal, measured over D periods of the excitation, the cross-spectral terms are calculated as

$$S_{uv}(k) = \frac{1}{D}\sum_{d=1}^{D} U_d^*(k)V_d(k) \tag{1.13}$$

where the variables $U_d(k)$ and $V_d(k)$ represent the discrete Fourier transform of the d^{th} period of $u(t)$ and $v(t)$ respectively, and the asterisk denotes the complex conjugate. For the case where D is equal to 1, the value of $\gamma_{xy}^2(k)$ reduces to

$$\tilde{\gamma}_{xy}^2(k) = \frac{X^*(k).Y^*(k).X(k).Y(k)}{X^*(k).X(k).Y^*(k).Y(k)}. \tag{1.14}$$

giving $\tilde{\gamma}_{xy}^2(k) = 1 \; \forall \; k \in K$. Now for the case when there is no noise present on the input and output records, it is clear that estimates of $S_{xy}(k)$, $S_{xx}(k)$ and $S_{yy}(k)$ in eqn. (1.12) can be calculated independently of the value of D. Hence the coherence function always reduces to the calculation given in eqn. (1.14). This leads to the conclusion that $\gamma_{xy}^2(k)$ will be unity even if $y(t)$ is derived from a purely nonlinear function of $x(t)$, a fact that can easily be overlooked. In situations where only the output signal is disturbed by noise, eqn. (1.12) reduces to

$$\bar{\gamma}_{xy}^2(k) = \frac{|Y(k)|^2}{S_{yy}(k)}, \tag{1.15}$$

which gives the same result as for the noise-free case regarding the detection of nonlinear effects *at the excitation frequencies.* The expression given in eqn. (1.15)

does offer, however, a very useful means of assessing the periodicity of the energy measured at non-excited frequencies in the output signal, since the calculation is independent of the input spectrum. This feature can be used with an m-level signal as the excitation, with the harmonic suppression implemented, to assess the linearity of the system under investigation.

If a significant amount of coherent output energy is detected in the output signal, the linear representation of the system may need to be augmented with a characterisation of the nonlinear contribution. The identification method used to obtain structural knowledge and parameter estimates for the system is at this point a choice to be made by the user. If the nonlinearity can be represented by a low-order polynomial function of the input or output signal, frequency-domain identification schemes can be easily formulated (Evans, Rees and Jones, 1992; McCormack et al., 1994b). In the absence of structural knowledge and/or in the presence of more general nonlinear behaviour, a strong possibility is offered by the nonlinear rational model representation (Billings and Zhu, 1991; Billings and Zhu, 1994). Although discrete-time estimators are not as dependent upon the characteristics of the excitation signal (particularly its periodicity), they do benefit from periodic designs (Schoukens et al., 1994; Nowak and Van Veen, 1994). It is shown by Godfrey, McCormack and Flower (1995) that the accuracy achieved when discrete-time estimators are used with a periodic excitation is similar to that obtained when estimators specifically designed to utilise the periodic nature of the data are used.

1.7 Conclusions

Multilevel signals yielding a specified set of spectral amplitudes have been discussed. Their properties have been shown to be useful for realising arbitrary Fourier specifications to a high degree of approximation. They are suitable for the realisation

both of optimal spectra, based on prior estimates of the system parameters, and of spectra selected on a more heuristic basis. Taking this into account with the possibilities offered by harmonic suppression, gives them an intermediary position between the properties of binary and multisine signals.

Chapter 2

The Application of Software and Hardware Tools in System Identification

2.1 Abstract

System identification methods are now incorporated into several commercially available products, including two MATLAB Toolboxes and a number of frequency response analyzers. This chapter assesses the accuracy obtained with the two MATLAB Toolboxes, when used with idealised noisy data, and also their performance when used with real measurement data. The performance of the curve-fitting algorithms on two commercial FFT analyzers is also compared using two applications for illustration.

2.2 Introduction

The availability of commercially available identification tools means that, today, it is relatively straightforward for the practising engineer to use the vast amount of knowledge that has been gathered from several decades of research into system identification and parameter estimation. Several textbooks are available, of which Eykhoff (1974), Ljung (1987), Norton (1986), Schoukens and Pintelon (1991) and Söderström and Stoica (1989) are examples. Interactive analysis tools are available for estimation procedures developed under the assumption of piecewise constant excitation (Ljung, 1986) and using frequency-domain methods (Kollár, 1994). There is also a noticeable trend towards the design of spectral analysis equipment which allow parametric models to be obtained, along with the more traditional non-parametric techniques. Examples are Hewlett Packard Co. (1989) and Advantest Corporation (1989). The applicability of each method depends greatly on the experimental setup, the intended use of the model and also on the required accuracy, but it is fair to say that for most control and measurement requirements, a linear system model can now be obtained comparatively easily.

The objectives here are to list some of the approaches which have made it to commercially available products, to describe some practical applications of these identification methods, and also to emphasise the often-neglected role of perturbation signal design in linear system identification (Godfrey, 1993c). Several notable surveys have appeared in the literature on system parameter estimation, both in discrete-time (Åström, 1980; Ljung, 1985) and in continuous-time (Young, 1981; Sinha and Rao, 1991). Transfer function estimation from frequency response data has been reviewed by Whitfield (1986).

It is assumed here that the identification is carried out off-line, and that a periodic excitation can be applied to the system under investigation. The discrete-time estimators have the advantage that they can be formulated for on-line implementation (Ljung and Söderström, 1983) and do not require the excitation to be periodic. However, significant advantages can be obtained from the use of periodic excitations (Schoukens et al., 1994), while the formulation of the frequency-domain estimators becomes tractable due to the algebraic relationship between the input and output Fourier coefficients (Van den Bos, 1970). The problem of obtaining a Laplace transfer function from an empirical frequency response function is treated separately, since it is not based on a statistical framework.

The chapter is split into two parts. In the first, the two MATLAB identification Toolboxes mentioned above are used to identify two systems, one of them simulated and the other a bench-scale heating process. The first Toolbox (Ljung, 1986) assumes a zero-order-hold (ZOH) excitation, and is aimed at estimating the coefficients of a discrete-time model of the process and, if required, a model of the disturbing noise. The second Toolbox (Kollár, 1994) requires a bandwidth-limited (BL) excitation, and is aimed at estimating the coefficients of either a discrete- or continuous-time model. In the applications described here, the two Toolboxes are used *correctly*. Substantial systematic errors can arise from application of ei-

ther Toolbox with data that violate the basic assumptions (Pintelon, Schoukens and Chen, 1994b; Schoukens et al., 1994). The objective therefore is to compare models obtained through the correct use of the two Toolboxes. Undoubtedly the most common source of systematic error in system identification is the presence of nonlinearities. From the point of view of linear model estimation, in the presence of system nonlinearities, there is no explicit advantage in using the identification algorithms of either Toolbox in favour of the other. However the periodic excitation design algorithms of the frequency-domain Toolbox can prove useful, e.g. the user is given the opportunity to design a signal with an arbitrary power spectrum and low peak-to-peak value.

In the second part of the chapter, two applications of estimating a Laplace transfer function of a process by curve-fitting to an empirical frequency response function (FRF), measured using a spectrum analyzer, are described. Results have been obtained from two different analyzers.

2.3 Application of MATLAB Toolboxes

2.3.1 Application 1: Simulated Third-Order Process

For this application, the process is a (software) simulated third-order process with a transfer function consisting of three equal time-constants, each of 10ms. The Laplace transfer function is thus given by

$$G(s) = \frac{1}{10^{-6}s^3 + 3.10^{-4}s^2 + 0.03s + 1}. \tag{2.1}$$

Two different perturbation signals are used. The first consists of 50 consecutive-harmonics (1 to 50), while the second consists of 400 consecutive-harmonics (1 to 400). Both signals have uniform power spectra. Other broadband excitation sign-

als, e.g. pseudorandom signals, may be employed with both Toolboxes, but here the multisine signals are favoured due to their ability to realise an arbitrary power spectrum. The peak-to-peak value of both signals is minimised using a time-frequency domain swapping algorithm (Van der Ouderaa et al., 1988a). The frequency ranges of the signals are 1 Hz to 50 Hz for the 50-harmonic signal and 0.125 Hz to 50 Hz for the 400-harmonic signal. The total power in each signal is set to 0.5, and Gaussian noise with variance equal to 0.04 is added to the output; the noise is band-limited, with constant power spectral density up to 50 Hz and zero power above 50 Hz.

The empirical FRF estimated from one period of the response, to each of the two perturbation signals (BL data), is shown in Figure 2.1, together with the theoretical FRF. From Figure 2.1 it can be seen that the increased signal-to-noise ratio afforded by the 50 harmonic signal gives a more accurate estimate of the frequency characteristic. This proves very useful in assessing model errors.

For the identification of a parametric model, eight periods of the response to each excitation signal (50-harmonic and 400-harmonic) are used to obtain each model estimate. For the time-domain identification (using the System Identification Toolbox) a ZOH excitation is used. Time-domain averaging is performed on the eight periods, with the sampling periods used for the 50- and 400-harmonic excitations made the same by specifying 1024 and 8192 samples per period, respectively. An output error model structure is used, which is of the form

$$y_r = \frac{B(q)}{F(q)} u_{r-nk} + e_r \tag{2.2}$$

where

$$B(q) = b_1 + b_2 q^{-1} + b_3 q^{-2}, \quad F(q) = 1 + f_1 q^{-1} + f_2 q^{-2} + f_3 q^{-3}, \text{ and } nk = 1.$$

This is the simplest model capable of fitting the measurements in the frequency-

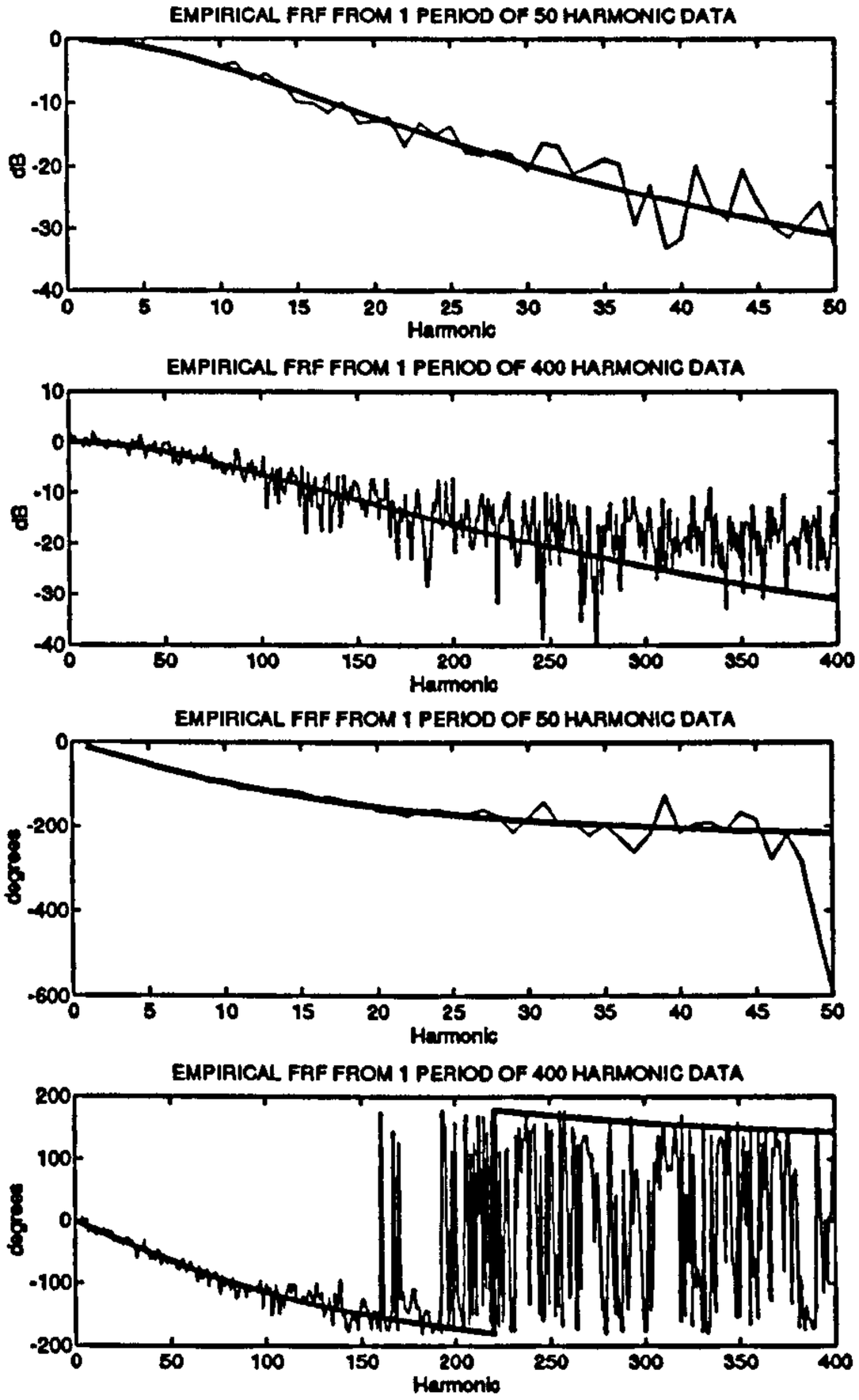

Figure 2.1: Empirical frequency response functions for Application 1, each one estimated from one period of data.

domain.

For the frequency-domain identification (using the Frequency Domain System Identification Toolbox), the output Fourier coefficients are averaged over the eight periods, with the variances of the real parts (equal to the variances of the imaginary parts) calculated during averaging. The model structure used is

$$Y_m(\jmath\omega_k) = G(\jmath\omega_k, P)X(\jmath\omega_k) + N_y(\jmath\omega_k), \qquad k = 1, 2 \cdots F \tag{2.3}$$

with $X(\jmath\omega_k)$, $Y_m(\jmath\omega_k)$ and $N_y(\jmath\omega_k)$, representing the Fourier coefficients of the true input, measured output, and output noise, respectively, and with F the number of harmonics. The vector P contains the unknown parameters of the transfer function, $G(s)$, $(s = \jmath\omega)$ of eqn. (2.1).

Note that, for both model structures, input noise is not added or considered. Since the frequency-domain algorithms are based on an error-in-variables model, the input noise variance is set equal to zero for the calculations. The time-domain and frequency-domain parameter estimates are used to compute the estimated frequency response over the range 0.125 Hz to 50 Hz (400 points), with the estimation repeated with 100 realisations of the output noise. The estimated mean and standard deviation of the error, in the estimated frequency characteristics, are shown in Figures 2.2 and 2.3, respectively. From Figures 2.2 and 2.3 it can be seen that all of the errors are small, with low uncertainties, showing that in this idealistic case, with the algorithms used under the correct assumptions, there is little to choose between the two MATLAB Toolboxes.

2.3.2 Application 2: Hot-Air Flow Device

The second application involves the identification of a bench-scale hot-air flow device (the PT326 Process Trainer). This is quite a realistic laboratory scale process, exhibiting many of the problems of a typical industrial process, including output drift (trend), inherent process noise and a limited linear range. The 50-harmonic signal is used to excite the heater within the linear range, with the period of the signal equal to 40 seconds. Averaging of the output data is over two periods of the signal.

A second-order plus time-delay model (as suggested by the manufacturers) is fitted using both the MATLAB Toolboxes, and the frequency responses of the fitted models are shown in Figure 2.4, together with the empirical FRF. It can be seen

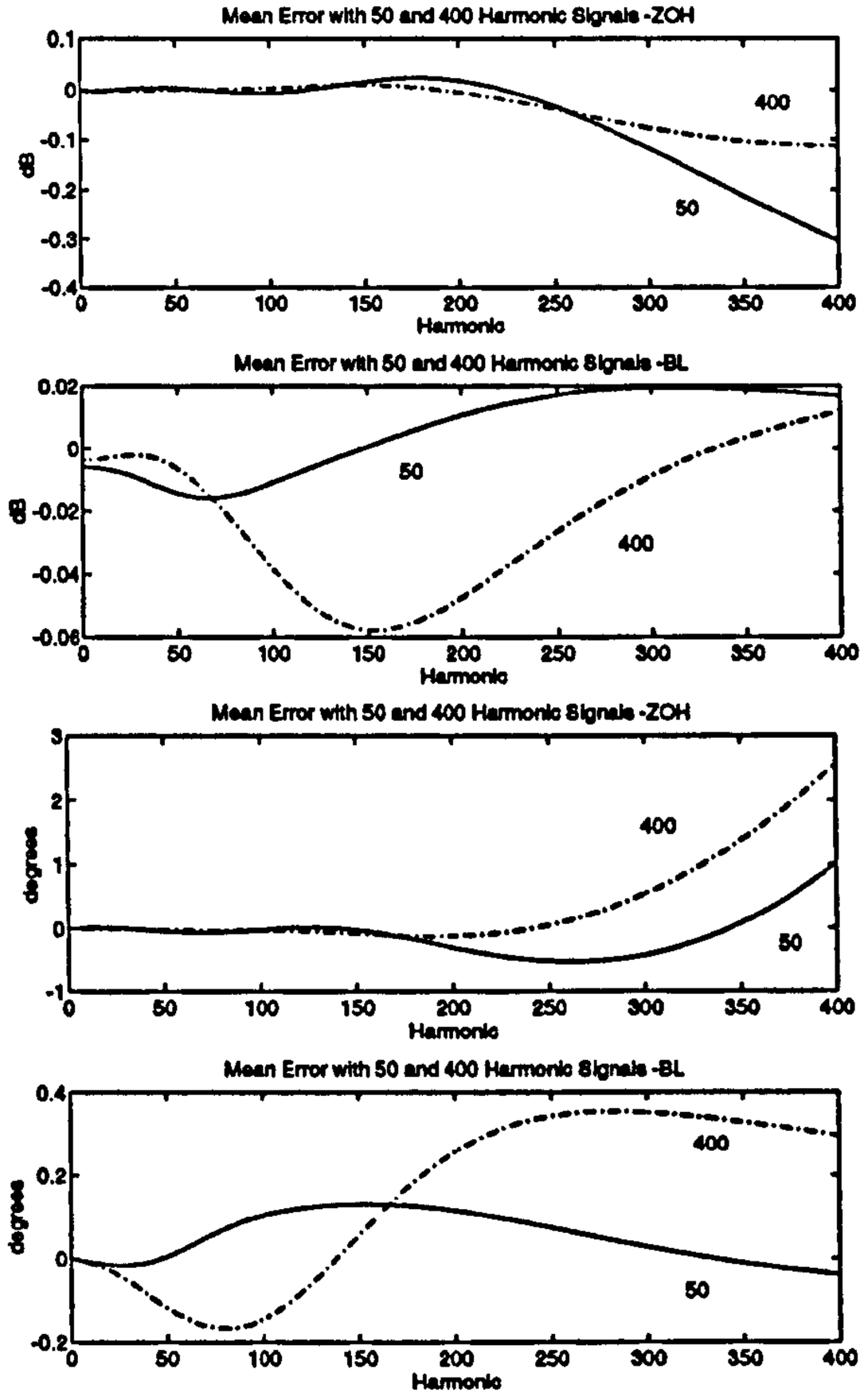

Figure 2.2: Mean error of estimated frequency response functions calculated from the parametric models

that the results obtained from this experiment again highlight the fact that both identification methods give very good results, if used correctly.

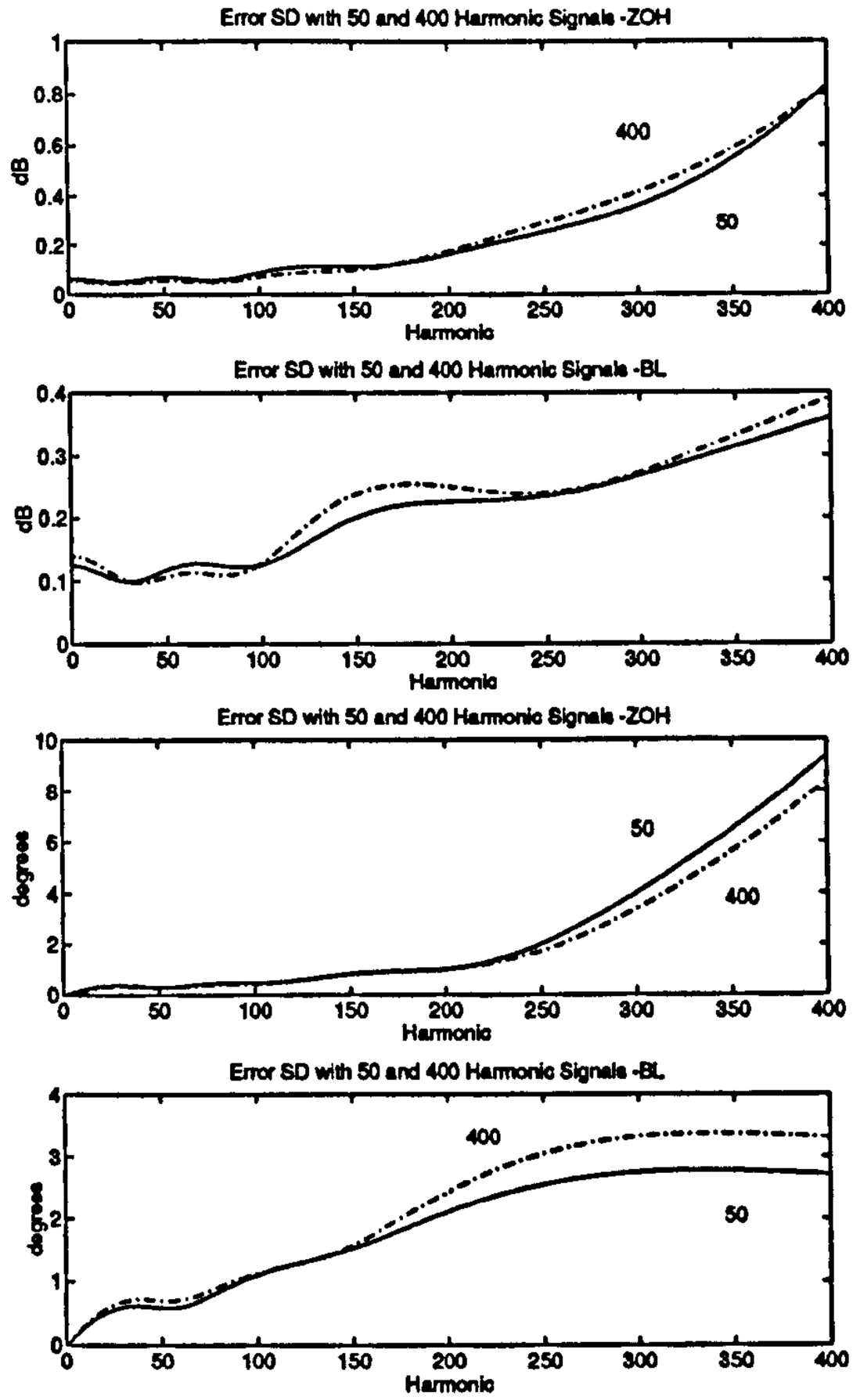

Figure 2.3: Standard deviations of frequency response function errors calculated from the parametric models

2.4 Curve-fitting to Frequency Response Functions

As noted in the introduction, there is an increasing trend in spectral analysis equipment towards incorporating facilities for obtaining parametric models by curve fitting to empirical frequency response functions. Some examples are the Hewlett-Packard 3562A Dynamic Signal Analyzer and 3563A Control Systems Analyzer,

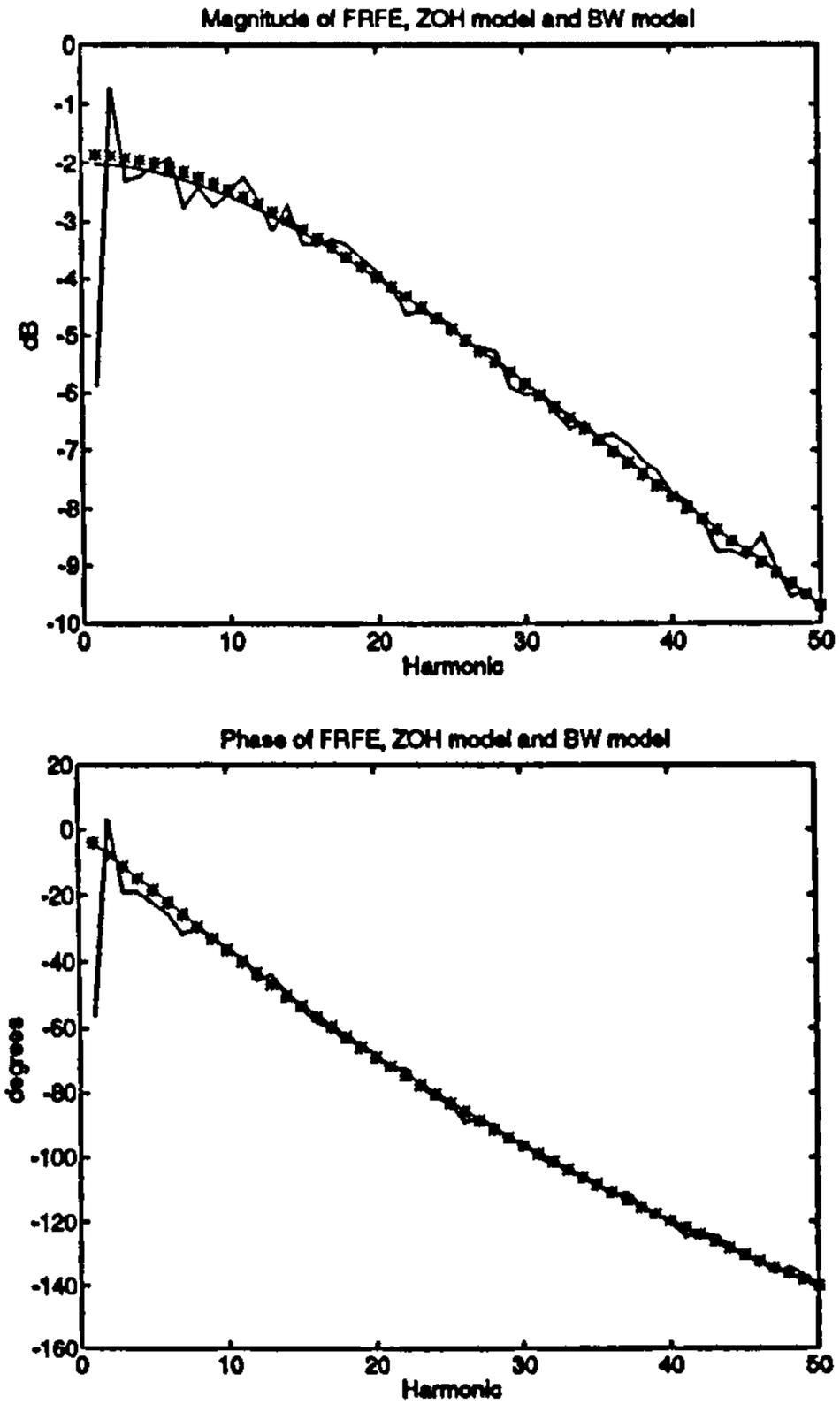

Figure 2.4: Frequency response function estimates for hot-air flow device. Empirical (-), ZOH model (smooth -), and BL model (*).

both of which can estimate a Laplace transfer function with up to 40 poles and 40 zeros. Another example is the Advantest R9211C FFT Servo Analyzer which has a limit of 20 sets of poles and zeros. The HP 3563A can also fit up to 40 poles and 40 zeros in a z-domain transfer function fit. To illustrate some of the problems encountered when using such techniques for system identification, two further applications are given below, both using the HP 3562A and the R9211C.

2.4.1 Application 3: Fourth-Order Butterworth Low-Pass Filter

The process in this case is a fourth-order Butterworth low-pass filter with cut-off frequency equal to 500 Hz. The *nominal* Laplace transfer function is given by

$$G(s) = \frac{9.741.10^{13}}{(s + 2903 \pm \jmath 1203)(s + 1203 \pm \jmath 2903)} \tag{2.4}$$

The word "nominal" is important here, because this is not a software simulated process, but a device constructed from the nearest preferred values of resistors and capacitors available. Therefore one might expect reasonable, but not exact, agreement between the theoretical and measured transfer functions. For both spectrum analyzers, the frequency response is measured with a maximum frequency of 2kHz (12566 rad s^{-1}). No noise is added at the output. Perturbation signals are obtained from the signal generators that form an integral part of the overall analyzer system and have approximately flat amplitude spectra over the frequency range up to 2 kHz.

For the HP 3562A results, the signal used is a periodic chirp (i.e. a swept sine wave). Fitting is carried out using the automatic fit facility, where the user does not specify the model structure. The estimated transfer function is

$$\hat{G}(s) = \frac{129158.10^{3}(s + 747253)}{(s + 2942 \pm \jmath 1078)(s + 1208 \pm \jmath 2879)}. \tag{2.5}$$

The denominator is in reasonable agreement with that of $G(s)$ and so is the numerator once it is realised that within the frequency range up to 2 kHz (12566 rad s^{-1}), the zero can effectively be regarded as a constant and combined with the multiplier of s in the numerator of $\hat{G}(s)$ to give a numerator constant of $9.651.10^{13}$. The agreement between $\hat{G}(s)$ and $G(s)$ is certainly as close as might be expected

given that only the nearest preferred values of resistors and capacitors were used in constructing the filter.

For the 9211C results, the perturbation signal is a 400 consecutive-harmonic multisine, with equal power in each harmonic. The signal is very similar to a Schroeder-phased multisine. The estimated transfer function, $\hat{G}(s)$, is

$$\hat{G}(s) = \frac{-159.4.10^{-6}(s+20306)(s-32112)(s-5221 \pm j30347)}{(s+2896 \pm j1078)(s+1202 \pm j2875)}. \qquad (2.6)$$

It is clear that the denominators of $G(s)$ and $\hat{G}(s)$ are in reasonable agreement, but the numerator of $\hat{G}(s)$ appears, at first sight, to bear no resemblance to that of $G(s)$.

At this point, it is necessary to consult the "know-how of use" in the Instruction Manual of the R9211C, where one reads that "large enough" zeros and poles should be removed by regarding them as constants (taking the modulus in the case of complex zeros and poles). All the zeros are larger in magnitude than the maximum radian frequency of 12566 rad s^{-1} and so replacing them by constants, the numerator becomes $9.392.10^{13}$, which is close to that of $G(s)$. The main problem associated with this procedure is precisely what is meant by "large enough". In this example, the magnitude of the smallest zero is 20306, compared with a maximum radian frequency of 12566, the ratio between them of 1.616 not appearing particularly large. The frequency response calculated from $\hat{G}(s)$ is indistinguishable from the measured frequency response and is therefore not reproduced here.

2.4.2 Application 4: First-Order System with Noise

In the final application, the process is a first-order lag, set up on a Feedback PCS327 Process Simulator, with a nominal transfer function consisting of a time-constant of 0.01s (denoted a) and a numerator of 1 (denoted b). As in the preceding example,

the word nominal is important because the process is part of a hardware simulator. For both spectrum analyzers, the frequency response is measured with a maximum frequency of 100 Hz.

For the HP 3562A results, the perturbation signal used is a periodic chirp, and fitting is carried out using the automatic fit facility. Averaging over 16 periods of the input signal, with no noise added, the fitted transfer function, $\hat{G}_{ref}(s)$, has parameters of $\hat{a} = 0.00982$ and $\hat{b} = 0.95998$, which is reasonably close to the nominal $G(s)$. The experiment was then repeated with wide-band Gaussian noise added at the output. The noise level selected meant that the (squared) coherence function drops to 0.9 at the top end of the frequency scale. The first transfer function fit, $\hat{G}_1(s)$, has values $\hat{a} = 0.010067$ and $\hat{b} = 0.98214$, again reasonably close to the nominal $G(s)$. Very similar results are obtained when the experiment is repeated several times.

For the R9211C, as in Section 2.4.1, the perturbation signal is a 400 consecutive-harmonic multisine. Averaging over 16 periods of the input signal, and with no noise added at the output, the estimated transfer function, $\hat{G}_{ref}(s)$, has parameters $\hat{a} = 0.010027$ and $\hat{b} = 1.00484$, again close to the nominal $G(s)$. The experiment was then repeated with wide-band Gaussian noise added at the output, of the same amplitude and bandwidth as that used in the HP 3562A experiment.

According to the R9211C manual, "Curve fitting is inhibited in the case of an averaged frequency response function whose corresponding coherence function is less than 0.8". A major problem in using this type of analysis for system identification is illustrated in two consecutive runs under identical conditions, except for the specific realisation of the added random noise. For the first run, the empirical FRF and the FRF of the estimated model are shown in Figure 2.5. The measured coherence function is above 0.9 throughout the frequency range. For this run, the fitted transfer function, $\hat{G}_1(s)$, has parameters $\hat{a} = 0.01009$ and $\hat{b} = 1.00346$, which is close to the

nominal $G(s)$ and to the estimated $\hat{G}_{ref}(s)$, in the noise-free case. In the second run, results for which are shown in Figure 2.6, the empirical FRF is very similar to that of Figure 2.5, and the coherence is again above 0.9 throughout the frequency range. However, the fitted transfer function is completely different. There are two zeros, $s = -278.3$ and -1718.2, and fourteen poles, $s = -126.1 \pm \jmath 12.6$; $+217.8 \pm \jmath 178.9$; $-193.1 \pm \jmath 227.9$; $+148.2 \pm \jmath 470.6$; $-138.1 \pm \jmath 478.8$; $-50.9 \pm \jmath 650.1$; $+46.0 \pm \jmath 653.1$. Only one of the zeros could be replaced with a constant, while six of the poles are in the right half of the s-plane.

From inspection of Figure 2.6, it may be seen that the frequency response of the model is in poor agreement with the empirical FRF at higher values of frequency. This second fit is wholly unacceptable. Further runs with the same noise level produces either a first-order fit or multiple-zero/multiple-pole fits of varying order. The conclusion, therefore, is that the curve fitting route to obtaining a Laplace transfer function *on this particular device* is liable to produce a very poor fit when even a quite modest amount of noise is present in the system.

In all areas of system identification (and especially in the control area where processes are of comparatively low-order), it is necessary that the user is able to specify the order of both the denominator and the numerator, as in the MATLAB Toolboxes discussed earlier. With the HP 3562A, the user *does* have the option of specifying the number of poles and number of zeros required in any fit. In Application 3, when a no-zero, four-pole fit is specified, the estimated Laplace transfer function is

$$\hat{G}(s) = \frac{9.760.10^{13}}{(s + 2978 \pm \jmath 1044)(s + 1207 \pm \jmath 2875)} \tag{2.7}$$

The denominator is very close to that obtained when using the automatic fit facility, and the numerator constant is also similar to the $9.651.10^{13}$ obtained from the automatic fit once the constant term and the zero have been combined. In Application

4, the HP 3562A analyzer produces the correct structure of model, even when noise is added. In contrast, the R9211C does not allow the user to specify the order of the numerator and denominator for the transfer function. To compound the problem, the model-order estimation algorithm produces unusable results.

2.5 Conclusions

The first two application examples, the first on a simulated system and the second involving a bench-scale heating system, have shown that, in these examples, both MATLAB identification Toolboxes examined here produce good quality results if used correctly. The 50-harmonic and 400-harmonic perturbation signals used in Section 2.3.1 show that the parameter estimation methods used in both Toolboxes are not highly dependent upon the harmonic content of the excitation. However, for practical identification, the increased SNR at each harmonic frequency, obtained with the 50-Harmonic signal, makes it easier to inspect possible model errors, since the uncertainty on the empirical FRF is reduced.

The second two application examples, both on hardware systems, have shown that, for one of the analyzers (an Advantest R9211C FFT Servo Analyzer), considerable skill may be needed in interpreting a Laplace transfer function obtained by curve-fitting to a noise-free measured frequency response. When a comparatively small amount of noise is added to the system output, the results can be nonsensical. From the viewpoint of most system identification applications in the control engineering area, the second of these problems makes this a dangerous route to use. The problem is seen to be overcome if the user is able to specify the order of the numerator and denominator of the transfer function to be fitted, as in the MATLAB Toolboxes. The user is also able to specify the model structure in the other analyzer used (a Hewlett-Packard 3562A Digital Signal Analyzer), although very good results

are obtained from this even when using the automatic fit facility, in which the model structure is not specified. The advantages of good perturbation signal design using this means of identification are clear, since high accuracy of the empirical FRF is necessary.

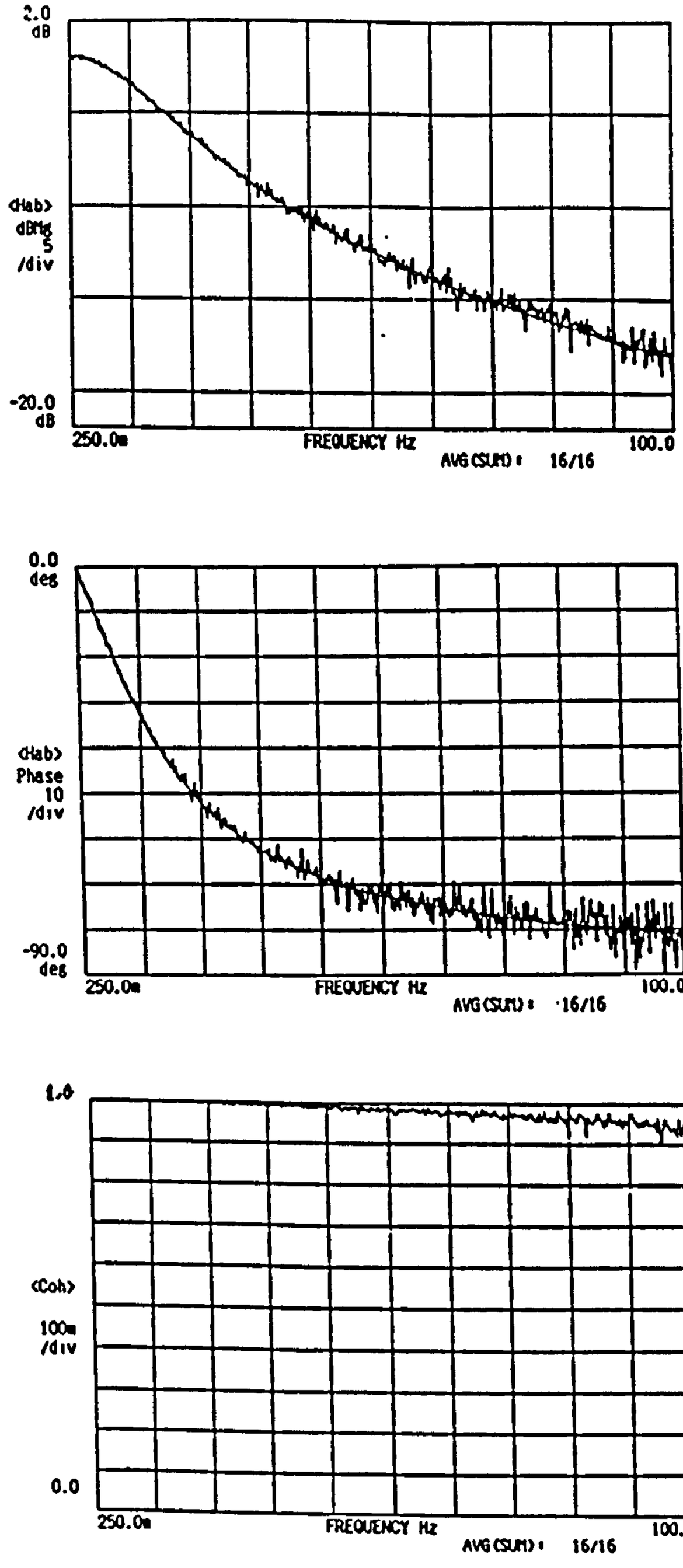

Figure 2.5: Empirical frequency response function for first analyzer run, together with frequency response function of the fitted model.

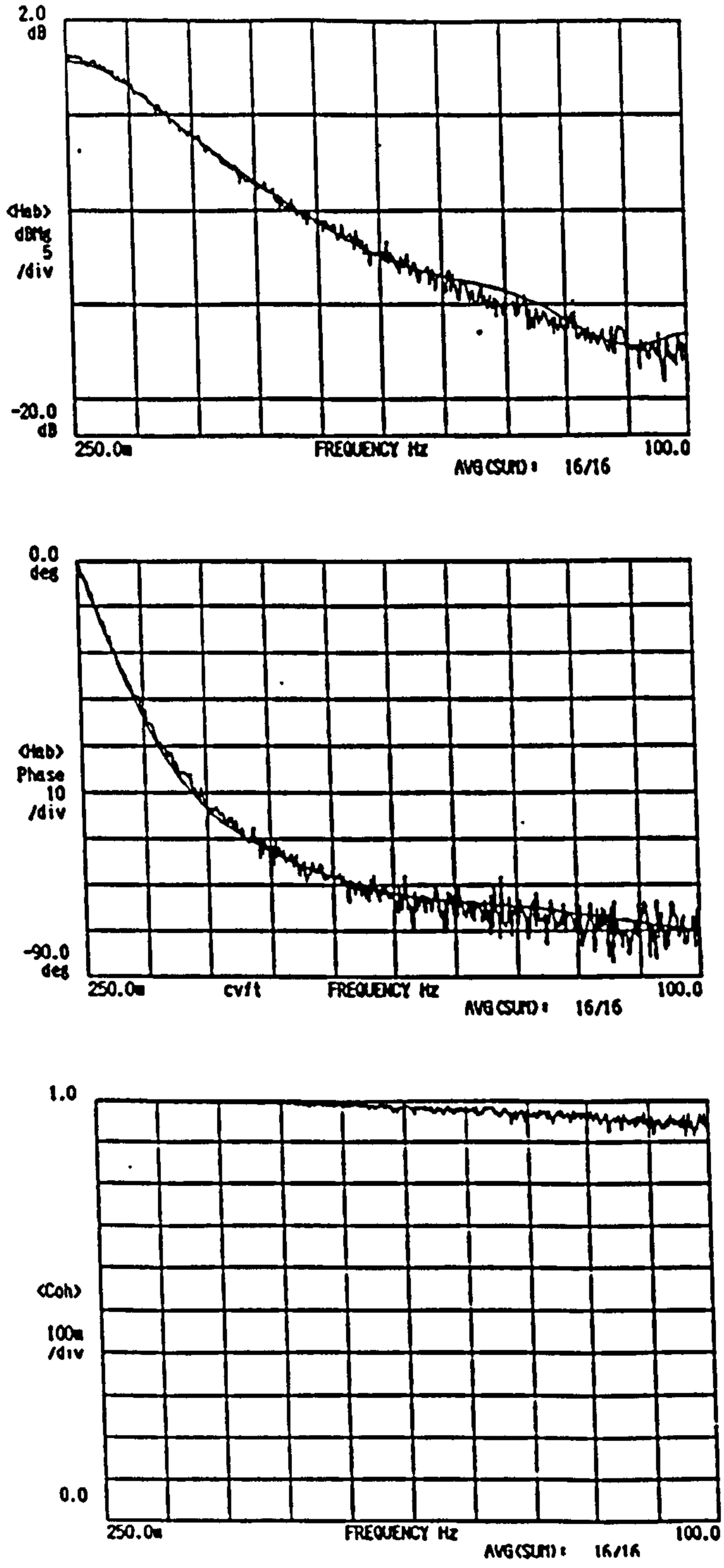

Figure 2.6: Empirical frequency response function for second analyzer run, together with frequency response function of the fitted model.

Chapter 3

The Suppression of Drift Effects for System Identification

3.1 Introduction

The treatment of periodic data records which are contaminated by drift disturbances (also termed trends or aperiodic noise) is considered in this chapter. Pre-processing methods are discussed, and so the system parameters may be estimated using any of the available identification techniques. However, the estimation techniques employed to approximate the drift disturbance may only be used when a periodic signal is used to excite the system under investigation. Due to the advantages to be gained from using periodic excitations in system identification applications (Schoukens et al., 1994), this is not seen as a practical limitation. Indeed the assumption of a periodic signal has been used by many authors to tackle the drift problem (Barker, 1967; Barker, 1969; Ream, 1968; Brown, 1968), where pseudo-random binary signals are the most frequently encountered excitation. Due to the increasing use of perturbation signals which are designed to meet frequency-domain constraints (Godfrey, 1993c), a re-examination of the drift problem, taking into account the properties of these excitation signals, seems timely.

Two frequency-domain approaches are considered here. The first, the decoupled approach, offers a simple method for establishing the drift characteristics by making use of the non-excited frequency lines in the measured response signal. This method can be employed in conjunction with every class of periodic excitation, but it is more suitable where considerable flexibility is allowed in the control of the power spectrum, i.e. multisine and multilevel signals. The method can be employed in characterising highly nonlinear forms of drift disturbance, with increasing constraints on the perturbation; thus it is most attractive when dealing with the frequently encountered lower-order forms of disturbance. The second method, the coupled approach, needs no assumptions to be made about the harmonic content of the excitation, and so is appropriate for all experiments where a periodic sig-

nal is used. This advantage however needs to be weighed against the considerably increased computational burden.

The practical identification of type-2 systems is dealt with in Section 3.6. This material is included here, since the application of a periodic signal to a type-2 system results in the response signal drifting away from the desired operating region. It is shown that the reference phase proposed by Flower, Forge, Radcliffe and Roust (1978), to eliminate the change in the mean level of the output signal encountered when testing type-1 systems, can be used unchanged in this application.

The treatment of drift effects by other workers is reviewed in Section 3.2. The decoupled and coupled approaches to the drift problem are presented in Section 3.3, together with a discussion of the required perturbation signal properties. The resulting accuracy of the drift parameters and the Fourier coefficients is examined in Section 3.4, while Section 3.5 examines the accuracy of the complete experiment when a perturbation signal, suitable for use with the decoupled estimator, is used as an excitation. The treatment of drift effects for multivariable system identification, using the approaches developed here, is examined in Section 3.5.1. The reference phase method is discussed in Section 3.6, followed by some conclusions in Section 3.7

3.2 Time-Domain Drift Compensation Schemes

The established methods for the treatment of drift in the time-domain can be split into reference phase methods and weighting methods. Both of these approaches share the common property that the characteristic of the drift is not directly estimated, and so they are expected to be immune to additional stochastic disturbances present in the measurement system. However these methods generally depend on the perturbation signal used, and also on the *a priori* knowledge of the likely drift

characteristic.

The suppression of drift effects when pseudo-random signals derived from m-sequences are used as the perturbation signal, has been examined in Barker (1967) and Barker (1969). It was shown that it is possible to generate several reference phases, for the pseudo-random signals, which have the property that the correlation with constant and linear signals is zero. It was further shown that out of these reference phases it is possible to select one which minimises the correlation with quadratic signals. For drift signals containing cubic contributions, it is necessary to use a weighting function, with the correlation being taken over two periods of the measured response to the perturbation. This concept is similar to the approach adopted by Brown (1968) where an inverse-repeat perturbation signal is studied and also to that of Davies and Douce (1967). In the former, the system estimate is shown to be immune to drift if the correlation is taken over a number of half-periods corresponding to the nature of the drift influencing the response. The latter authors generalise the weighting concept so that it can be used with an arbitrary excitation. A set number of periods still being required.

3.3 Frequency-Domain Characterisation of Drift

In contrast to the time-domain methods, the drift-suppression techniques detailed here are based on frequency-domain results, and rely on direct estimation of the drift characteristics, prior to system parameter estimation. They have the advantage that a fixed number of observation periods is not required, and also that they can be used with most classes of excitation signal. The first method to be presented is the decoupled approach.

3.3.1 Decoupled Estimation of Output Signal Parameters

The measurement problem is illustrated in Figure 3.1. The un-measurable distur-

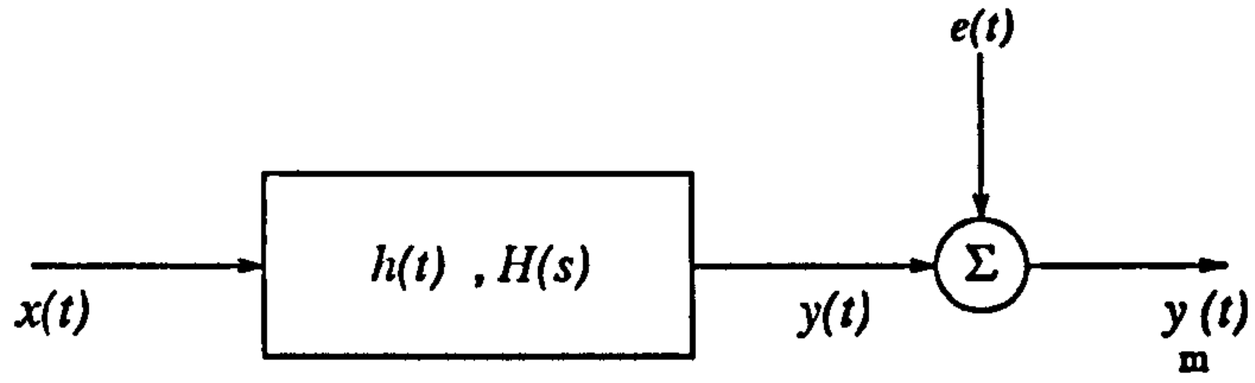

Figure 3.1: System response disturbed by un-measurable $e(t)$.

bance $e(t)$ is assumed to be given by

$$e(t) = n(t) + d(t) \tag{3.1}$$

where $n(t)$ is assumed to represent a stochastic disturbance, and $d(t)$ is a low frequency disturbance which may be represented by

$$d(t) = \sum_{m=1}^{M} a_m.t^m. \tag{3.2}$$

The exclusion of the constant term, $(m = 0)$, in $d(t)$, is due to the frequency-domain analysis which follows. The input signal in Figure 3.1 is assumed to be periodic, and the measurement record, $y_m(t)$, is assumed to contain no transient contribution, i.e the steady-state has been reached (see Chapter 4). In this case, representing the input signal by its Fourier series

$$x(t) = \sum_{k}[\gamma_k \cos(k\omega_0 t) + \delta_k \sin(k\omega_0 t)] \tag{3.3}$$

where ω_0 and $k \in K$ are respectively the fundamental frequency and harmonics of $x(t)$, the p^{th} period of the measured response, $y_m(t)$, is given by

$$y_{m_p}[t+(p-1)T] = \sum_k [\zeta_k \cos(k\omega_0 t) + \eta_k \sin(k\omega_0 t)] + \sum_{m=1}^{M} a_m (t+(p-1)T)^m + n(t). \quad (3.4)$$

If measurements are made over one period of $x(t)$, $(P = 1)$, Fourier transformation of eqn. (3.4) results in,

$$Y_m(\jmath k\omega) = Y(\jmath k\omega) + \sum_{m=1}^{M} a_m C_m(k) \quad (3.5)$$

where,

$$Y(\jmath k\omega) = \frac{1}{2}(\zeta_k - \jmath \eta_k) \quad (3.6)$$

and

$$\begin{aligned} C_m(k) &= \int_0^T t^m \exp(-\jmath k\omega_0 t) dt \\ &= \frac{\jmath T^{m+1}}{2\pi k} - \frac{\jmath m T}{2\pi k} C_{m-1} \end{aligned} \quad (3.7)$$

with

$$C_1 = \frac{\jmath T^2}{2\pi k}. \quad (3.8)$$

Although the preceding analysis considers continuous-time signals and in reality, the measurements will consist of a discrete sequence for $y_m(t)$, the following points can be made:

- If $X(\jmath k\omega)$, the complex amplitudes of $x(t)$, equal zero, the solution for a_m is a linear least-squares problem.

- For identifiability, at least M harmonic amplitudes of $X(\jmath k\omega)$ must equal zero. This is due to the linear dependency exhibited by the values for $C_m(k)$.

Recognising that the measured response signal will represent a discrete set of values, and due to fact that $d(t)$ is assumed to be time limited (and hence not band limited), the values for $C_m(k)$ required for estimation purposes are calculated using

$$C_m(k) = \frac{2}{P.N} \sum_{n=0}^{N-1} \sum_{p=1}^{P} (n + (p-1)N)^m \exp(\frac{-\jmath 2\pi kn}{N}) \tag{3.9}$$

with $P.N$ the total number of observations of $y_m(.)$. From eqn. (3.5) with $C_m(k)$ calculated using eqn. (3.9), and assuming a linear relationship between $x(t)$ and $y(t)$, it can be seen that by setting $X(\jmath k\omega)$ to zero for $D \geq M$ values of k $(k_1, k_2 \cdots k_D)$, the parameters $a_1 \cdots a_M$ may be estimated from $Y_m(k)$ using

$$\hat{\Psi}_{dc} = (\Omega_{dc}^T \Omega_{dc})^{-1} \Omega_{dc}^T \Upsilon_{dc} \tag{3.10}$$

where,

$$\Omega_{dc} = \begin{pmatrix} \alpha_1(k_1) & \alpha_2(k_1) & \cdots & \alpha_M(k_1) \\ \beta_1(k_1) & \beta_2(k_1) & \cdots & \beta_M(k_1) \\ \alpha_1(k_2) & \alpha_2(k_2) & \cdots & \alpha_M(k_2) \\ \beta_1(k_2) & \beta_2(k_2) & \cdots & \beta_M(k_2) \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1(k_D) & \alpha_2(k_D) & \cdots & \alpha_M(k_D) \\ \beta_1(k_D) & \beta_2(k_D) & \cdots & \beta_M(k_D) \end{pmatrix} \Upsilon_{dc} = \begin{pmatrix} \sigma(k_1) \\ \tau(k_1) \\ \sigma(k_2) \\ \tau(k_2) \\ \vdots \\ \sigma(k_D) \\ \tau(k_D) \end{pmatrix} \hat{\Psi}_{dc} = \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_M \end{pmatrix}$$

with,

$$C_m(k) = \alpha_m(k) + \jmath\beta_m(k)$$

$$Z(k) = \sigma(k) + \jmath\tau(k).$$

The estimated drift signal, $\hat{d}(t)$, can then be removed from the measured response. The accuracy of this method for drift estimation is examined in Section 3.4

3.3.2 Signal Design for Decoupled Drift Estimation

The estimator $\hat{\Psi}_{dc}$ relies on the fact that $y_m(n)$ contains no information about the system $H(s)$ at certain Fourier coefficients. The class of perturbation signals which allow the experimenter to create this situation needs to be discussed. The most suitable choice for the problem is the multisine signal, since it affords complete control over the power spectrum. In this case, $D \geq M$ harmonic amplitudes can be set to zero, or if experimental restrictions allow, $P > M$ periods of the response signal may be measured and the Fourier analysis taken over the $P.N$ measurements. In situations where the stochastic disturbance, $n(t)$, is substantial, this may be necessary to get an accurate estimate of the drift if it has to be described by a high-order polynomial ($M \geq 3$).

For binary signals, both discrete-interval binary signals (Van den Bos and Krol, 1979) and pseudo-random signals (Godfrey, 1993b), the inability to suppress harmonics from the power spectrum (other than suppressing all odd- or all even-harmonics of the fundamental), means that it will generally be necessary to record multiple periods of the response. This is equivalent to using an excitation designed by concatenating at least $M + 1$ periods of a consecutive harmonic signal together. Multilevel signals, both m-level (see McCormack, Godfrey and Flower (1995) and Chapter 1) and pseudo-random signals (Barker, 1993), do allow, however, greater control of the power spectrum. If these are used, it is possible to have an input spectrum with the Fourier coefficients equal to zero at all harmonics which are multiples of a set of prime numbers. For example, the complex amplitudes at all multiples of the second- and third-harmonics may be set to zero with an m-level signal (three amplitude levels) or a pseudo-random signal (seven amplitude levels). Hence it

will be possible to characterise a drift signal containing linear, quadratic and cubic contributions.

In Section 3.5 the effect on the accuracy of system frequency response (estimated via a parametric model) is examined when a multisine signal, with gaps left in the low frequency end of the spectrum, is used to excite the system. The ability to compress this type of signal in the time-domain (effect on signal-to-noise ratio), compared to a signal containing consecutive-harmonics is also examined.

3.3.3 Coupled Estimation of Drift and Signal Parameters

The estimation of the drift parameters prior to calculation of the Fourier coefficients is attractive from a computational point of view, but for situations where time is not a constraint on the off-line calculations, an alternative estimator is considered here. Examination of eqn. (3.4) shows that the output $y(t)$ is linear in the unknown Fourier coefficients, η_k and ζ_k, as well as the drift parameters, a_m. These can therefore be estimated in a straightforward manner, with the n^{th} row of the regression matrix, Ω_c, now given by

$$\left(\cos(2\pi kn) \quad \sin(2\pi kn) \quad \cdots \quad \cos(2\pi Kn) \quad \sin(2\pi Kn) \quad n \quad \cdots \quad n^m \right) \tag{3.11}$$

with $n = 0 \cdots P.N-1$ and Υ_c a vector containing the time-domain samples. Clearly the size of Ω_c, $(P.N \text{ x } (2K + M))$, is quite unwieldy for large K and PN, but estimation of the parameters via this route will be more accurate, due to the use of all the time-domain samples. The accuracy loss, and its significance, using the decoupled approach is examined in Section 3.4.

3.4 Accuracy of the Drift Estimators

In this section the accuracy of the decoupled drift estimator is compared to that obtained using the coupled approach. To begin with, the loss in accuracy in the parameter estimates, resulting from processing only the lower end of a polynomial spectrum, is examined. Following this, a comparison is given of the results achieved with the two estimators, in the presence of a multiharmonic signal. Two conceptual situations for the decoupled approach are treated. The first corresponds to the case where harmonics are left free in the perturbing spectrum, and the second when at least $M + 1$ periods of the response are measured, allowing a minimum of M harmonics to be used for drift estimation. The former case will assume that the fundamental frequency is not present, since, if this is not true, the two cases are identical. For both situations the accuracy of the Fourier coefficients calculated from the response signal is examined, and so the accuracy of the DFT is included for comparison purposes.

For both estimators, the covariance matrix of the parameter estimates, P_z, is given by

$$P_z = (\Omega_z^T \Omega_z)^{-1} \Omega_z^T R \Omega_z (\Omega_z^T \Omega_z)^{-1} \tag{3.12}$$

where P_{dc} and P_c correspond to the covariance matrices for the decoupled (regression matrix Ω_{dc}) and the coupled (regression matrix Ω_c) methods respectively. If the stochastic disturbance on the Fourier coefficients is assumed to have covariance matrix

$$R = \sigma_y^2(k).I, \tag{3.13}$$

with $\sigma_y^2(k)$ representing the variance of the real or imaginary part of the Fourier coefficients at harmonic k, the expression for the covariance matrix reduces to

$$P_z = \sigma_y^2(k)(\Omega_z^T \Omega_z)^{-1}. \tag{3.14}$$

If the noise spectrum, $e(\jmath\omega)$, can be assumed to be flat in the region of interest, then the noise variances, $\sigma_y^2(k)$ can be calculated from the time domain variance using

$$\sigma_y^2(k) = \frac{2.\sigma_t^2}{N} \quad \forall k \in K \tag{3.15}$$

where σ_t is the standard deviation of $e(t)$.

As an illustration of the decrease in the parameter accuracy from, firstly increasing the number of harmonics and, secondly using harmonics $1 \cdots D$, to estimate the polynomial coefficients, Figure 3.2 shows the normalised variances obtained for the parameters of a second-order function. The first and second values on each

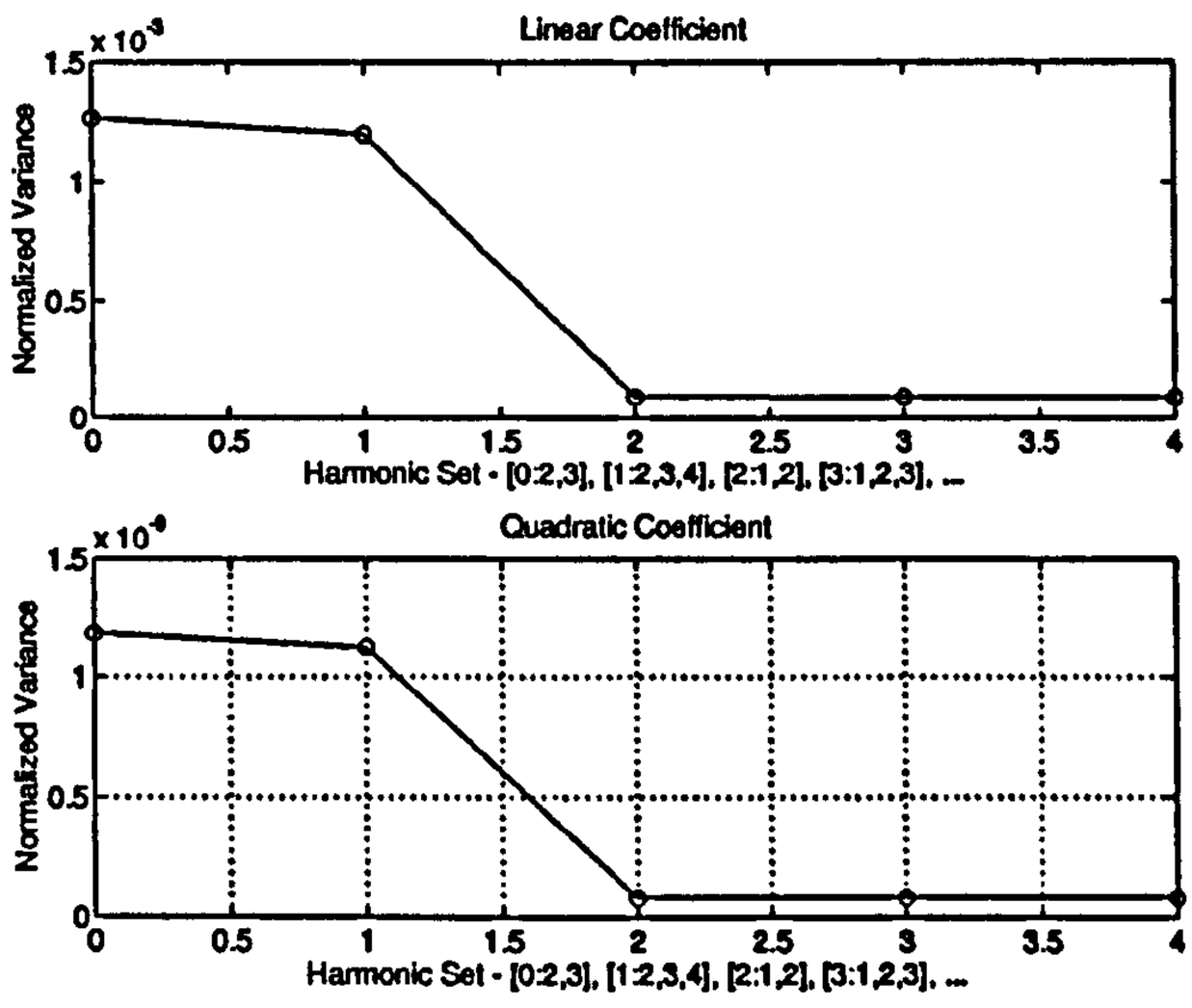

Figure 3.2: Normalised variances for linear and quadratic parameters of drift polynomial

graph are obtained from processing harmonics $[2,3]$ and $[2,3,4]$ respectively, while successive points correspond to harmonics $[1,2], \cdots, [1,2,3,4]$. In the limit, using harmonics $[1 \cdots P.N/2]$ corresponds to the coupled method. The results of Figure 3.2 show that there is a large drop in the variance of the estimates when the first harmonic can be used for drift estimation. However it also shows that there is little

to be gained from using the computationally intensive coupled estimator if multiple periods of the response can be measured.

The accuracy of the Fourier coefficients is the main consideration, and so this will be considered now. The DFT will be taken as the benchmark, which under the given noise assumptions, yields parameter variances given by eqn. (3.15). The variance of the Fourier coefficients obtained via the coupled estimator, which will intuitively be very close to that obtained to the DFT, can be obtained from eqn. (3.14). Since the decoupled approach is a two-step procedure, the variance of the Fourier coefficient estimates will be obtained from an analysis of 100 sets of data. For comparison, this procedure will also be carried out for all cases where the theoretical variance values are available.

The results given in Table 3.1 are the variances obtained using each estimator with a data set consisting of a multisine signal (N=1024) superimposed upon a second-order polynomial. The perturbation signal is designed to have equal power in

Estimator	*Estimate* $\|Y\,j\omega\|$	
	Theoretical	Measured
DFT 1 Period	8.84e-03	9e-03
Coupled 1 Period	9.5e-03	9.5e-03
Coupled 2 Periods	6.3e-03	6.2e-03
Decoupled [2,3,4]	-	10.2e-03
Decoupled [1,3,···9]	-	6.2e-03

Table 3.1: Statistical results for Fourier coefficients obtained via the DFT, the coupled and decoupled drift estimators

harmonics 1,5,6···100. The performance of the drift estimators is given for the cases where one and two periods of the response is measured. In the former, harmonics 2,3

and 4 are available, while in the latter, harmonics 1,3,4$\cdots$9 can be used to estimate the drift parameters. The term $|Y(j\omega)|$ has been chosen for examination since its standard deviation is approximately the same as the real and imaginary parts of $Y(j\omega)$. The first 20 harmonics of the signal have been used for the calculation of the mean standard deviation.

The results of Table 3.1 indicate superficially that the decoupled drift estimator gives good performance (equal to the coupled method when two periods are measured), however, the harmonic standard deviations given in Figure 3.3 show the true story. It can be seen that the decoupled estimator gives a significant increase

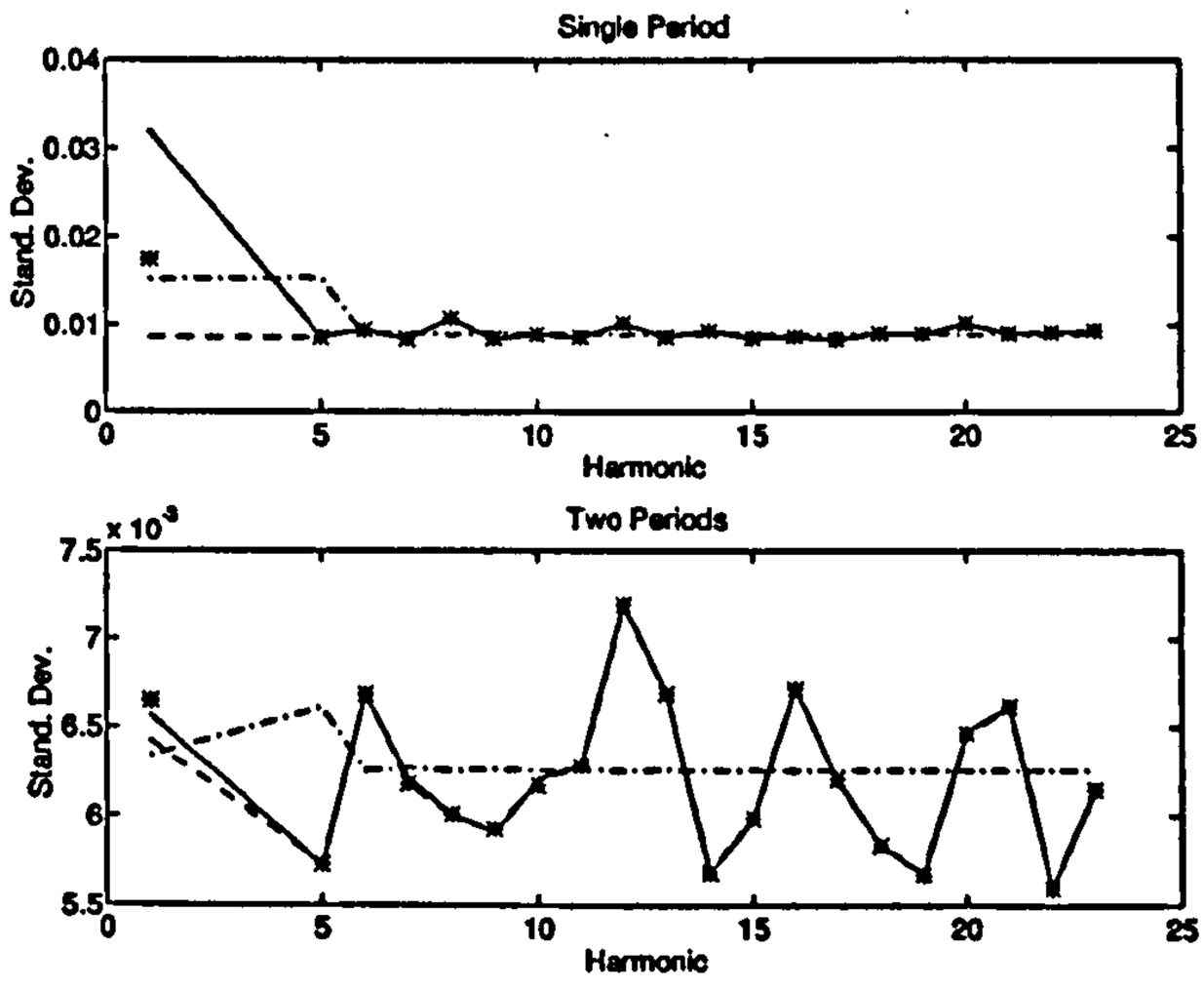

Figure 3.3: Standard deviations obtained with DFT (--), decoupled drift (–) and coupled drift (*) estimators, with theoretical value for coupled drift (-.).

in the standard deviation of the spectral estimate when one period of the response is measured. However this decreases to the same level as that obtained with the coupled and DFT estimators when the first harmonic can be used, which in this case was possible when two periods of the response were processed.

3.5 Input Optimisation with Decoupled Signal

The effect of using an excitation signal with missing harmonics, on the final parameter estimates will now be examined. The non-parametric estimate is not considered since it is absent at the harmonics used for drift estimation. The treatment will assume that after drift estimation, the drift contribution is zero, and hence the properties of the input signal alone are important. The effect on the uncertainty is considered followed by the effect on the ability to compress a signal with the designated spectrum in the time-domain.

The decoupled drift estimator requires that at least M harmonic frequencies, in the excitation signal, are not used to perturb the system. Since these harmonics need to be as low as possible, to accurately characterise the drift, there is the potential to leave out the frequencies at which the system parameters display maximum sensitivity to the input signal. Since the majority of systems encountered in the process industries display low-pass characteristics, it will be the energy at low frequencies (around the bandwidth of the system) which contribute the most to the accuracy of the estimated model. Although optimal power spectra are rarely calculated in practice, the algorithm given by Van den Eijnde and Schoukens (1991) and implemented in Kollár (1994) is useful for giving an indication of the frequencies at which the system should be excited. Here, for illustration purposes and simplicity, the system to be identified is assumed to be given by a first-order system with a d.c. gain of 1 and the pole at -1. The set of harmonics (1,5,$\cdots$ 50) is used to construct the input. This will be compared to a signal with consecutive harmonics (1,2 $\cdots$ 50), with both signals designed, initially, to contain equal powers.

The fundamental frequency for both signals is set equal to a quarter of the cut-off

frequency for the system, so that the harmonic frequencies are

$$F = 1/8\pi, 2/8\pi, 3/8\pi, 4/8\pi, 5/8\pi, \cdots 50/8\pi. \tag{3.16}$$

The frequencies in **bold** are used in the first signal to estimate the drift, and so will contribute no information to the system parameter estimates. It should be noticed that the fourth harmonic is exciting the system at the -3dB point. The dispersion function (Goodwin and Payne, 1977; Schoukens and Pintelon, 1991) is used to examine the effect on the model variance when the above frequency grids are used for the experiment. The input amplitudes are initially set to a uniform distribution, which gives the two dispersion functions shown in Figure 3.4. Note

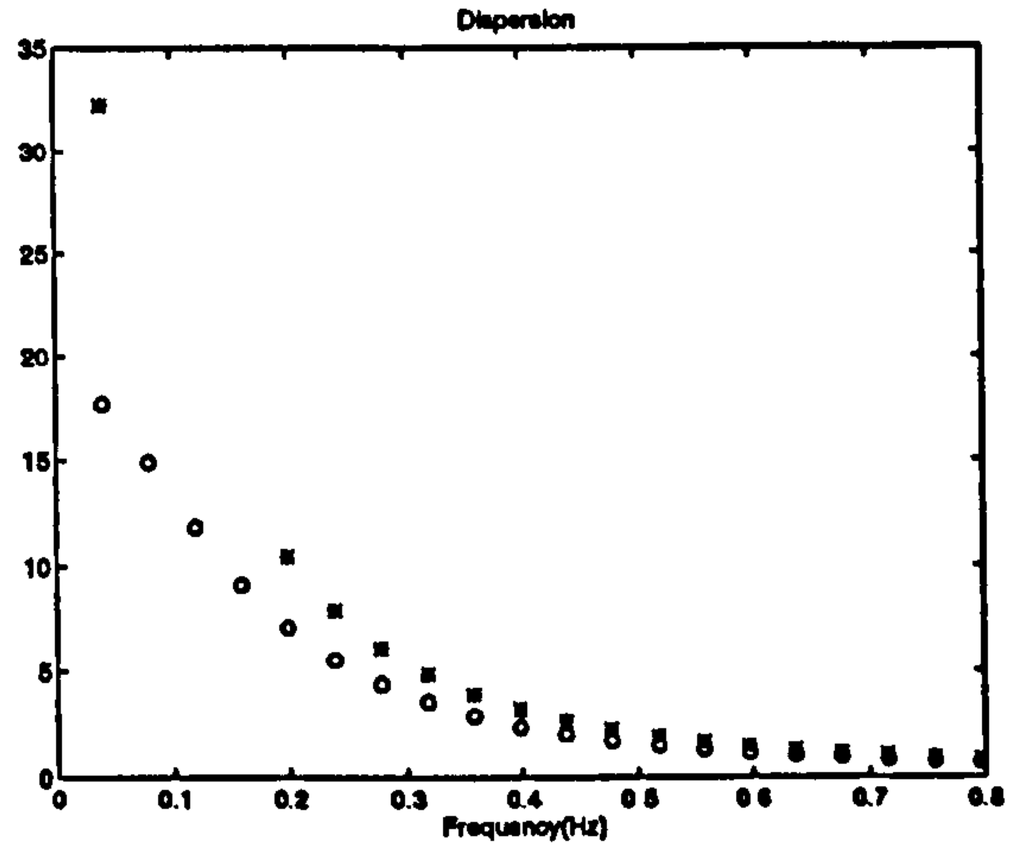

Figure 3.4: Dispersion functions obtained with uniform amplitude spectra. Consecutive (o) and non-consecutive (*) harmonic signal.

that the dispersion function for the non-consecutive harmonic signal has a gap in the low-frequency harmonics. It is expected that the recorded uncertainty at these frequencies will be greater than at the excited frequencies, due to the necessary interpolation. When the model structure is correctly specified, or in practical terms, the system is not over-modelled, this increased uncertainty will not be significant.

In Figure 3.4, the maximum frequency shown is $20/8\pi$, since the uncertainty drops off very rapidly. In this case, the maximum value of the dispersion function, calculated with the non-consecutive harmonic signal, is approximately 1.8 times that obtained with the signal containing consecutive harmonics. Although a minimum dispersion function is seldom sought, due to the general lack of precise prior knowledge of the system, it may be useful to carry out a small number of iterations of the optimisation algorithm. Here, the results after just one iteration will be examined, which corresponds to placing the majority of the signal power in the pass-band of the system. The new amplitude spectra for the experiment are shown in Figure 3.5. These new amplitudes can be seen to be placed where they are likely to contribute

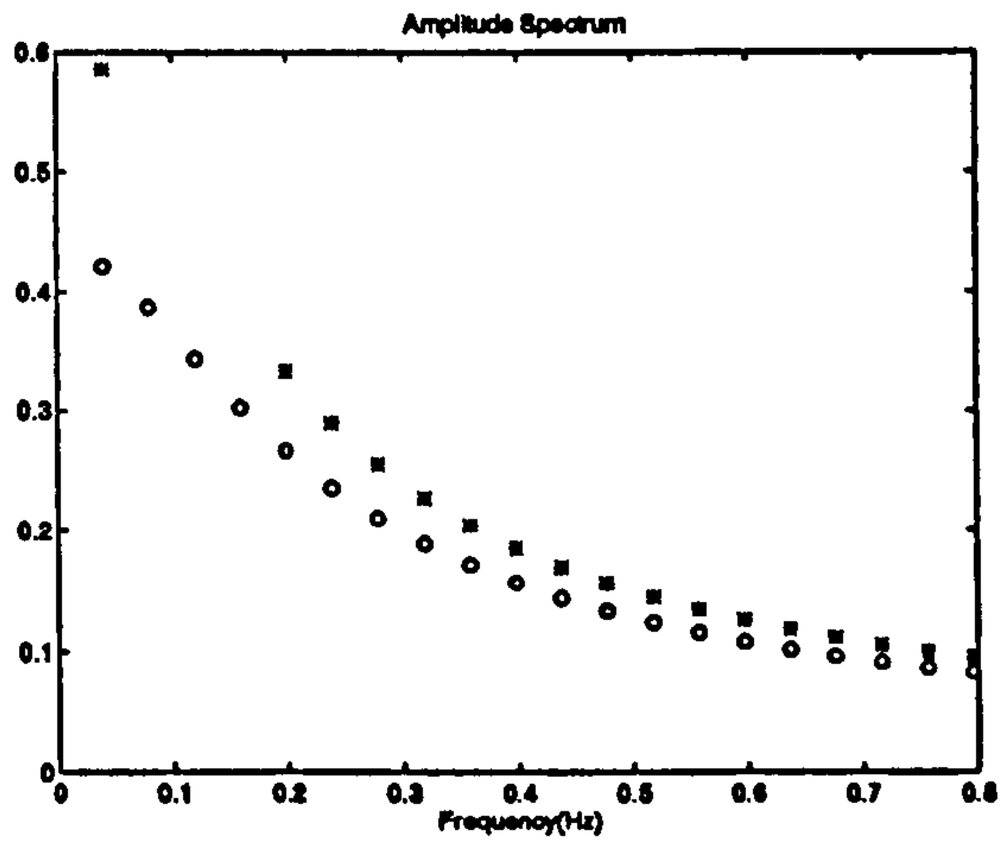

Figure 3.5: Amplitude spectra obtained after one iteration of the optimisation algorithm. Consecutive (o) and non-consecutive (*) harmonic signal.

the most information about the system. The corresponding dispersion functions are shown in Figure 3.6. In this case the maximum value of the dispersion function obtained with the non-consecutive harmonic signal is 1.03 times that obtained with the consecutive harmonic signal. Hence with proper selection of the signal bandwidth, no significant loss in accuracy is expected to result from using the non-consecutive harmonic signal.

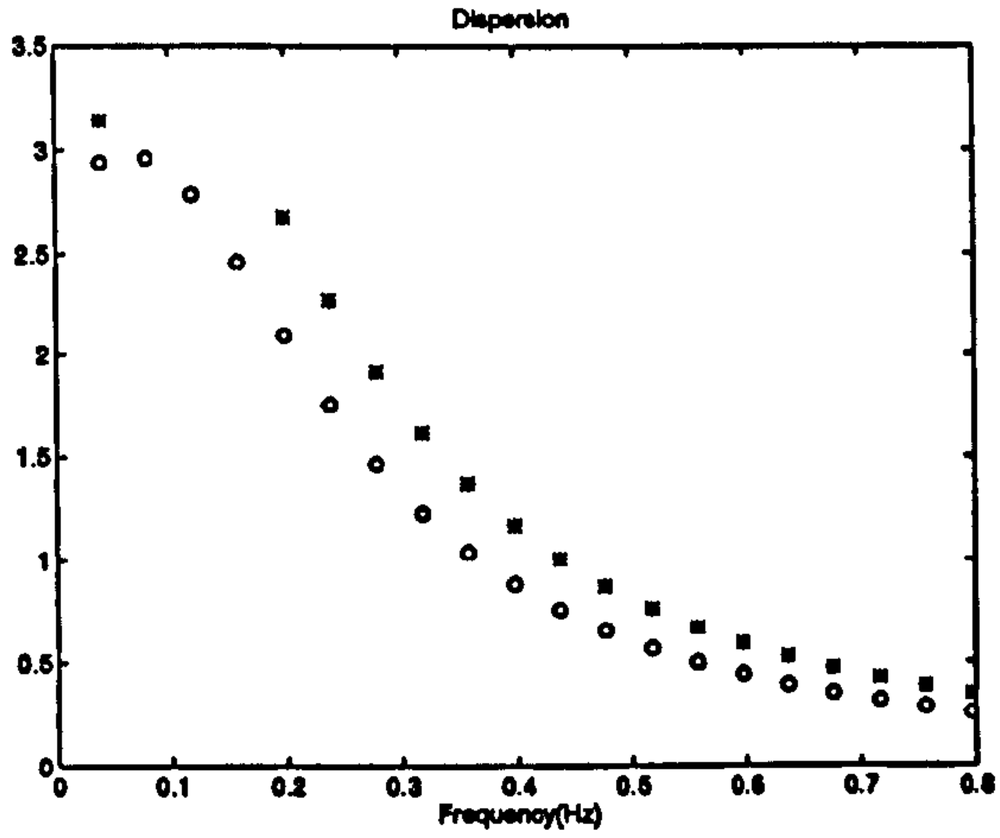

Figure 3.6: Dispersion functions obtained with amplitude spectra resulting from one iteration of optimisation algorithm. Consecutive (o) and non-consecutive (*) harmonic signal.

It is demonstrated above that the expected loss in accuracy in the model estimate is low when a signal designed for use with the decoupled drift estimator is used for the experiment. The two signals used in the comparison were assumed to have equal power, and so no consideration was given to the ability of the time-domain optimisation algorithms to realize the desired power spectra. The desired spectra are assumed to be realized with multisine signals, since multilevel signals are studied in detail in Chapter 1. If the clipping algorithm (Van der Ouderaa et al., 1988a) is used to optimise the time-domain signals, with 200 clipping iterations, the crest factors are 1.49 and 1.58 respectively for the uniform amplitude spectra, and 1.33 and 1.38 respectively for the "optimised" spectra. Since the dispersion function is normalised by the power in the signal (Schoukens and Pintelon, 1991), the relative increase from the equal power case can be calculated from the ratio of the power in each signal after they have been scaled to have equal peak values. In this case, the maximum expected ratio of the dispersion functions increases by a factor of 1.12 for the uniform amplitude spectrum and 1.08 for the amplitude spectra after one

iteration of the optimisation procedure. Hence if correct signal design procedures are followed (high signal power in useful part of system bandwidth), no significant drop in accuracy is to be expected if the signal designed to be used with the decoupled estimator is used for the experiment.

3.5.1 Drift Suppression for Multivariable System Identification

Frequency-domain identification of MIMO systems is conceptually identical to the scalar case if uncorrelated input signals are applied to each input (although global approaches are developed by Guillaume, Pintelon and Schoukens (1992) and Bayard (1994)). Since with the decoupled drift estimator, the drift contribution is modelled as another uncorrelated signal (at certain frequencies), the number of harmonics that can be used for the system identification, reduces in the lower frequency band, if the system has several inputs. In these cases, the decoupled estimator will give accurate results if multiple periods of the responses to the excitations can be measured. The coupled estimator is of course independent of the number of inputs present in the system. The same holds for the decoupled estimator when the system is single-input multiple-output.

3.5.2 Application Example

The drift suppression technique will be illustrated on the Feedback PT326 process trainer. An identification experiment was carried out using an odd-harmonic signal (harmonics $[1, 3, 5 \cdots 49]$) with a fundamental frequency of 128.02/1024 Hz. One period of steady-state data was recorded. The measured output signal is shown in Figure 3.7, together with the result of removing a linear drift term. The harmonics used for drift estimation were two and four, and as can be seen, are adequate for

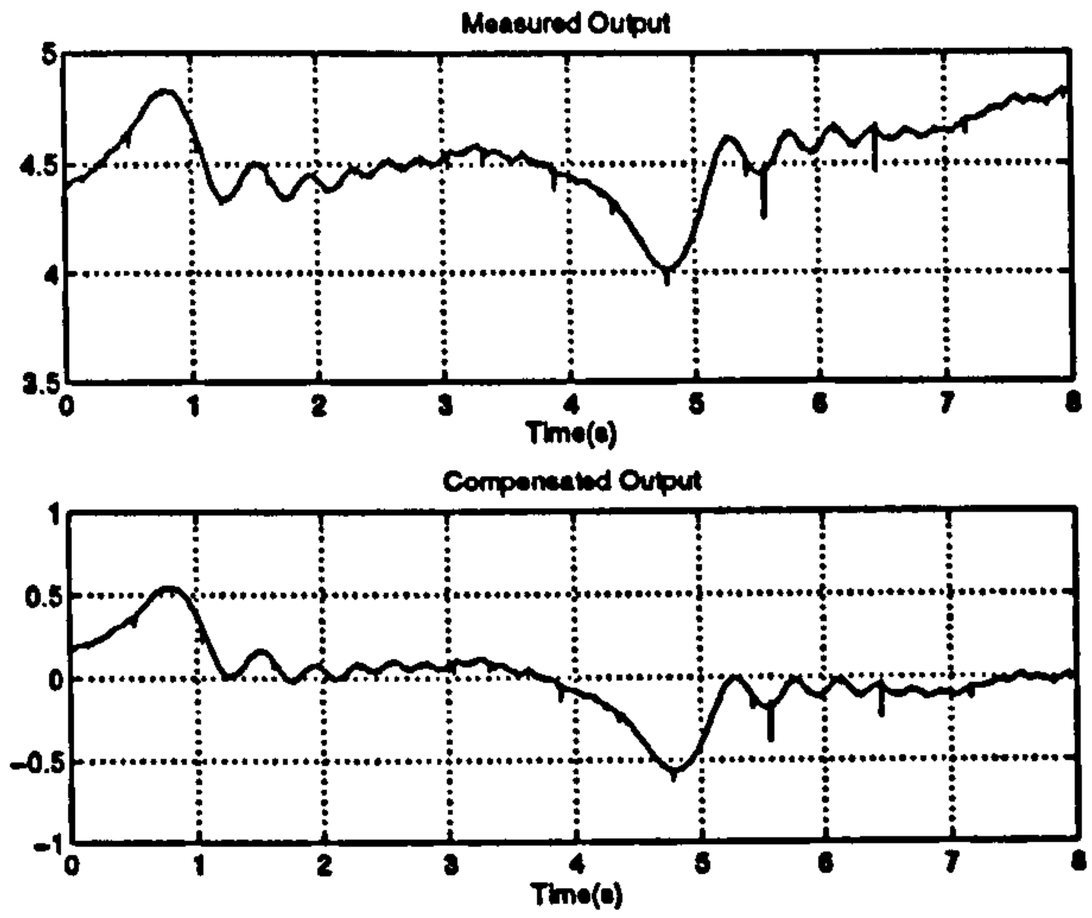

Figure 3.7: Upper: measured output data from the Feedback PT326 process trainer. Lower: drift corrected output data

accurate characterisation of the drift disturbance. The empirical frequency response functions calculated from the measured data and drift corrected data are shown in Figure 3.8. These show a significant improvement in accuracy after the drift has been removed. A similar more detailed identification example, also using the Feedback PT326 process trainer, is given in Section 4.3.3 of Chapter 4.

3.6 Excitation Induced Drift

This section illustrates a use of the reference phase (see Flower et al. (1978) and Chapter 4) of multiharmonic signals. In the appendix of Flower et al. (1978) it is shown how to specify the reference phase so as to eliminate the standing offset produced in the output of the process during frequency response testing of type-1 systems. This is accomplished by setting the coefficient of the $1/s$ term to zero in the partial fraction expansion of $G(s)X(s)$, where $G(s)$ is type-1, thus eliminating the offset term. Here it is shown how to specify the reference phase to eliminate the

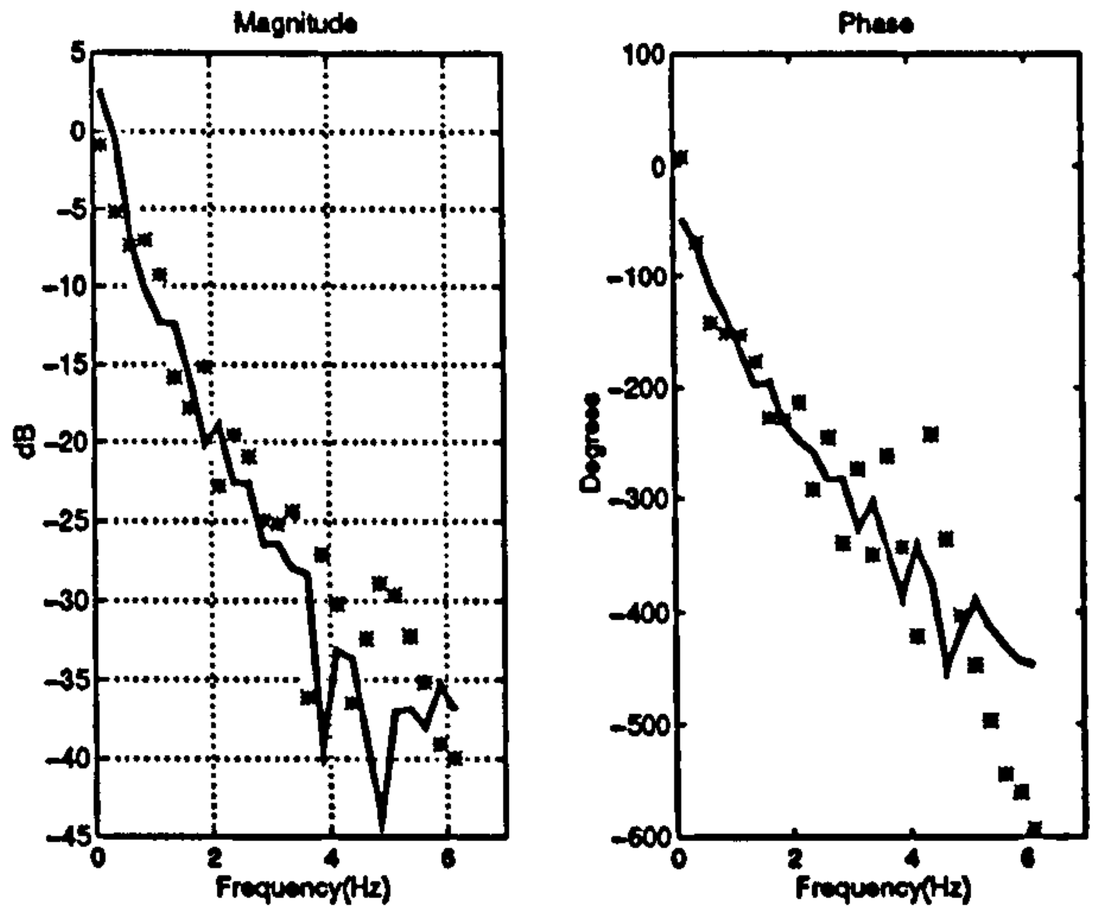

Figure 3.8: Frequency response functions for the Feedback PT326 process trainer. (*): from un-compensated time domain measurements, (-): from drift corrected data

drift due to the input signal when the system is type-2.

The multisine signal, $x(t)$, has a Laplace transform, $X(s)$, given by

$$X(s) = \mathrm{Re} \sum_k \frac{A_k \exp(\jmath\phi_k)}{s - \jmath k\omega_0} \tag{3.17}$$

where the phase set, $\{\phi_k\}$, which is usually used to minimise the crest factor of $x(t)$, can be written as

$$\phi_k = \phi'_k + \phi_0 \tag{3.18}$$

with ϕ_0 representing the reference phase of $x(t)$ (Flower et al., 1978). This reference phase is open to choice and is usually set equal to zero. The system, $G(s)$, is assumed to be given by

$$G(s) = \frac{G_0(s)}{s^2} \tag{3.19}$$

where $G_0(s)$ is type-0. In this case $Y(s)$ is given by

$$Y(s) = \frac{K_2}{s^2} + \frac{K_1}{s} + T(\alpha_m) + P(\jmath k\omega_0) \tag{3.20}$$

where the coefficients, K_1 and K_2, are calculated from

$$K_{M_r-n} = \frac{1}{n!}\frac{d^n}{ds^n}[(s+P_r)^{M_r}G(s)X(s)]_{s=-P_r} \tag{3.21}$$

where $n = 0, 1$ and M_r is the order of the pole $-P_r$, which in this case is two (Seifer and Steeg, 1960). Therefore

$$Y(s) = \frac{G_0(0)}{s^2}\mathrm{Re}\sum_k \frac{A_k\exp(\jmath\phi_k)}{-\jmath k\omega_0} + \frac{K_1}{s} + T(\alpha_m) + P(\jmath k\omega_0) \tag{3.22}$$

where K_1 is calculated from

$$K_1 = [\frac{d}{ds}(G_0(s)X(s)]_{s=0}. \tag{3.23}$$

Therefore

$$\frac{K_1}{s} = \frac{1}{s}[G_0'(0)\mathrm{Re}\sum_k \frac{A_k\exp(\jmath\phi_k)}{-\jmath k\omega_0} + G_0(0)\mathrm{Re}\sum_k \frac{A_k\exp(\jmath\phi_k)}{(k\omega_0)^2}] \tag{3.24}$$

where the prime, ($'$), denotes the first derivative. Now the term

$$\frac{G_0(0)}{s^2}\mathrm{Re}\sum_k \frac{A_k\exp(\jmath\phi_k)}{-\jmath k\omega_0} \tag{3.25}$$

produces a ramp term (drift) in $y(t)$. The presence of this drift term in $y(t)$, as well as causing practical problems, has the effect of producing a bias on the frequency response function estimate, $\hat{G}(\jmath k\omega_0)$. The drift term in this case is given by a linear function of time, which contributes an additive term to the Fourier coefficients of

$y(t)$, given by $\mathrm{F}(at)$, where F(.) denotes Fourier transformation and a equals the coefficient of $1/s^2$ in eqn. (3.25). If the Fourier analysis is taken over P periods of $y(t)$, the contribution of $\mathrm{F}(at)$ to harmonic k of $Y(\jmath k\omega_0)$, is calculated from eqn. (3.8) as

$$\mathrm{F}(at)\mid_k = \frac{\jmath aT^2}{2\pi k} \tag{3.26}$$

where T denotes the signal period. Therefore the elimination of the drift is necessary from an experimental and a measurement viewpoint. This is accomplished by equating the summation in eqn. (3.25) to zero and noting eqn. (3.18), which yields the following value for ϕ_0

$$\phi_0 = -\tan^{-1}\frac{\sum_k \frac{A_k}{k}\sin\phi_k'}{\sum_k \frac{A_k}{k}\cos\phi_k'} \tag{3.27}$$

This is the same reference phase that eliminates the offset produced in the output signal when testing type-1 systems (Flower et al., 1978). Note that the term

$$G_0(0)\sum_k \frac{A_k\cos\phi_k}{(k\omega_0)^2}, \tag{3.28}$$

representing a constant offset, is still present in the output. It is not possible, generally, to eliminate the ramp and offset terms simultaneously by use of the reference phase.

3.7 Conclusions

This chapter discussed methods for tackling the practical problem caused by drift disturbances during identification experiments. It was shown that frequency-domain measurement techniques can be used to suppress the drift effect from the measurement results. This is applicable to drift disturbances that can be modelled using

low-order polynomials, which are the most common type found in practice. The accuracy of the techniques was demonstrated through simulation results and a realistic experimental example. The techniques, as reported in this chapter, are restricted to situations where periodic excitations are used, and also when steady-state records are available. The latter restriction is dropped in the next chapter, where transient corrupted data records are considered. Included in this chapter was an application of the reference phase which can eliminate the linear drift which is inherent when using periodic signals to identify type-2 systems.

Chapter 4

The Suppression of Transient Effects for System Identification

4.1 Introduction

Experimental design for dynamic system identification considers many different constraints, one of which being the duration of the experiment. In this thesis and in many recent publications (Godfrey, 1993c; McCormack et al., 1994a; Schoukens et al., 1994), the properties of periodic excitations have been examined and many advantages have been highlighted. The price to be paid for these useful properties in frequency-domain identification is the necessary use of the DFT without systematic errors, which means that the data records, used for processing, must contain no transient information from the system. This can sometimes mean a large amount of time is wasted at the start of the experiment, which can be a costly exercise for systems with large settling times.

In attempting to alleviate this problem, some authors (Kvashnin, 1969; Darnell and Kemp, 1988; Darnell, 1993), have examined the properties of aperiodic signals and correlation schemes for both single- and multi-input system identification. The general methodology employed in the single-input case is to trail the signal with a suitable number of zeros and record the response of the system to the composite signal (aperiodic signal + zeros). An aperiodic correlation scheme is then used to obtain the system impulse response. The method however has the unfortunate disadvantage that the period of the excitation must be very much greater than the settling time of the system for accurate results to be obtained (Chen, 1993), hence defeating the original goal of a quick experiment. Also, for multi-input systems, the method seems to give very large systematic errors due to the non-zero cross-correlation between the aperiodic excitation signals available at the present time.

In this chapter, two simple methods for suppressing transient effects are given which may be used in conjunction with standard periodic excitations. The aim is to suppress the effect of the transient response on the Fourier coefficients, and so

the means of system parameter estimation does not need to be specified. Firstly in Section 4.2, a reference phase method is given for use with multisine signals, although its extension to general periodic excitations is straightforward. This method is applicable to systems where a single long time-constant dominates the transient response, and *a priori* knowledge of its value exists. It is shown that approximate knowledge of the value of this time-constant is sufficient to obtain accurate results.

To treat the more general case, an interpolated fast Fourier transform (IFFT) algorithm (Renders, Schoukens and Vilain, 1984) is examined in Section 4.3. This can be used to calculate the Fourier coefficients from data which does not contain an integer number of periods of each harmonic frequency. The philosophy employed is to truncate the transient corrupted section of the first period of the response and to process the remainder using the IFFT instead of the DFT (FFT). Hence by using the IFFT, the problem of dealing with transients is transferred to one of dealing with spectral leakage, which has been treated much more extensively in the literature (Jain, Collins and Davis, 1979; Grandke, 1983; Andria, Savino and Trotto, 1989; Offelli and Petri, 1990; Schoukens, Pintelon and Van hamme, 1992). The algorithm and its derivation is given in Section 4.3, with Section 4.3.1 comparing its statistical properties to those of the DFT. The extension of the IFFT algorithm to estimate a polynomial contribution to the signal is given in Section 4.3.2, which can be used where the response is corrupted by a drift term. In Section 4.3.3, an application example is given.

4.2 Transient Suppression via Reference Phase

The application of $X(s)$, given in eqn. (3.17), to a system $G(s)$ with poles α_m results in the following dynamic response:

$$Y(s) = \sum_m [\frac{K_m}{s+\alpha_m} \mathrm{Re} \sum_k \frac{A_k \exp(\jmath\phi_k)}{\alpha_m + \jmath k\omega_0}] + P(\jmath k\omega_0) \tag{4.1}$$

with K_m constants associated with $G(s)$, and $P(\jmath k\omega_0)$ the desired periodic mode of response. If $G(s)$ can be written as

$$G(s) = \frac{G_0(s)}{s+\alpha_l} \tag{4.2}$$

where $|\alpha_l|$ is small compared to $|\alpha_{m\neq l}|$, then $Y(s)$ can be written as

$$Y(s) = \frac{-G_0(-\alpha_l)}{s+\alpha_l} \mathrm{Re} \sum_k \frac{A_k \exp(\jmath\phi_k)}{\alpha_l + \jmath k\omega_0} + T(\alpha_{m\neq l}) + P(\jmath k\omega_0) \tag{4.3}$$

with $T(\alpha_{m\neq l})$ containing the transient terms due to $G_0(s)$. It is assumed that the transient terms in $T(\alpha_{m\neq l})$ die out rapidly compared with that produced by α_l. Writing

$$\mathrm{Re} \sum_k \frac{A_k \exp(\jmath\phi_k)}{\alpha_l + \jmath k\omega_0} = \sum_k \frac{A_k}{R_k} \cos(\phi_k - \theta_k) \tag{4.4}$$

with

$$R_k = \sqrt{\alpha_l^2 + (k\omega_0)^2} \text{ and } \theta_k = \tan^{-1} \frac{k\omega_0}{\alpha_l}$$

and equating eqn. (4.4) to zero, while noting (3.18) gives

$$\phi_0 = \tan^{-1}(\frac{\sum_k \frac{A_k}{R_k} \cos(\phi'_k - \theta_k)}{\sum_k \frac{A_k}{R_k} \sin(\phi'_k - \theta_k)}) \tag{4.5}$$

Therefore with knowledge of α_l, its transient response may be set to zero if ϕ_0 is calculated according to eqn. (4.5).

Since it is unlikely that exact knowledge of α_l will be available prior to the experiment, the sensitivity of the resulting periodic estimate, $P(k\jmath\omega_0)$, to inexact knowledge of α_l is important. To assess the efficiency of the method, a simulation was carried out with an input signal consisting of the first 300 harmonics of a fundamental, designed using the method of Van der Ouderaa et al. (1988a). The following error function is used to evaluate the results.

$$Err(\hat{\tau}, T) = \frac{1}{300} \sum_{k} (\mid G(\jmath k\omega_0) \mid - \mid \hat{G}(\jmath k\omega_0) \mid)^2 \tag{4.6}$$

where $G(\jmath k\omega_0)$ is the correct value of the frequency response function at harmonic k and $\hat{G}(\jmath k\omega_0)$ is calculated with $x(t)$ incorporating the reference phase. T and $\hat{\tau}$ denote the signal period and the estimate of the time constant respectively. Figure 4.1 shows the results of 90 simulations with $\hat{\tau}$ varying from being 10% in error to 150% in error ($\hat{\tau}$ being over-estimated each time), and T varying from $6.25/\alpha$ to $200/\alpha$. The errors involved for these practical levels of $\hat{\tau}$ and T are seen to be small thus giving high confidence in the use of the reference phase method.

4.3 Truncation and Interpolation

The fundamental frequency of the perturbation signal relative to the system bandwidth depends heavily on the system being identified, but generally the signal period is greater than the settling time of the system by a factor of at least three or four. Therefore, upon measurement of the first period of the response to the excitation, a large amount of the latter section of the data contains insignificant amounts of transient information. It therefore seems appropriate to explore methods of extract-

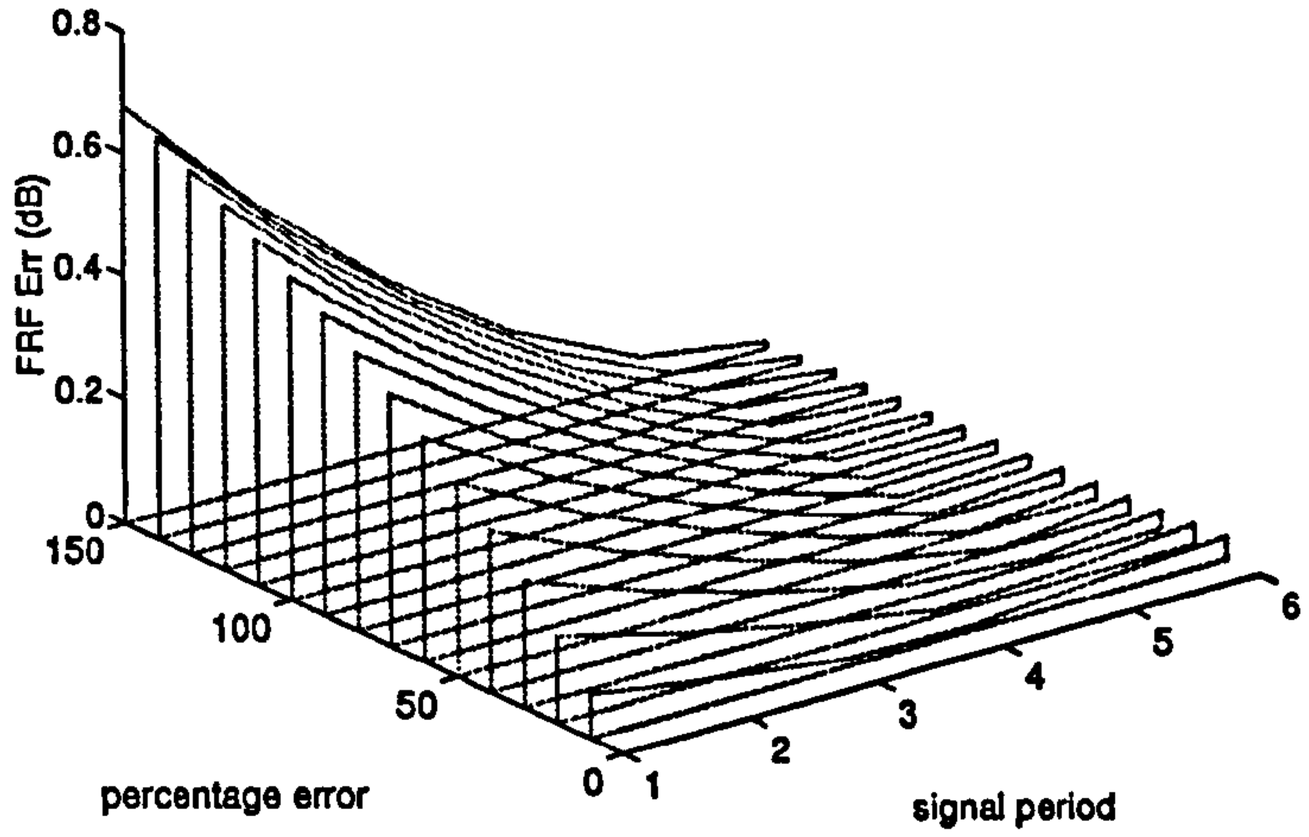

Figure 4.1: Average frequency response error obtained when incorrect values of the dominant pole are used for the calculation of the reference phase.

ing the Fourier coefficients from the latter section of the data record. Calculation directly with the DFT is not a possibility (even with windowing) due to the presence of spectral leakage. The interpolated FFT does offer a possibility, since its use is intended for data records containing non-integer cycles of the harmonic frequencies of the fundamental. The interpolated FFT used in this application (Renders et al., 1984) is a closed form solution due to the fact that the frequencies are known *a priori*.

Given a multifrequency signal $x(t)$ with a harmonic set $\{K\}$, sampled at discrete intervals Δt, with N representing the number of samples, a truncation of $x(\Delta t)$ by κ samples results in $x'(\Delta t)$ which can be written as

$$x'(n\Delta t) = \sum_{k'} X_{k'} \cos(2\pi f_{k'} n\Delta t) - Y_{k'} \sin(2\pi f_{k'} n\Delta t). \tag{4.7}$$

where

$$X_{k'} + \jmath Y_{k'} = A_{k'} \exp(\jmath \phi_{k'}), \tag{4.8}$$

$$f_{k'} = k \frac{N'}{N} f_0 = \lambda_{k'} f_0 \tag{4.9}$$

with $N' = N - \kappa$, and

$$\phi_{k'} = \phi_k + \frac{2\pi k \kappa}{N}. \tag{4.10}$$

The DFT of eqn. (4.7) at line i is given by

$$\begin{aligned} U_i + \jmath V_i = \sum_{k'} \sum_{n=0}^{N'-1} & X_{k'} \cos(2\pi\lambda n/N').\cos(2\pi i n/N') \\ & - Y_{k'} \sin(2\pi\lambda n/N').\cos(2\pi i n/N') \\ & - \jmath X_{k'} \cos(2\pi\lambda n/N').\sin(2\pi i n/N') \\ & + \jmath Y_{k'} \sin(2\pi\lambda n/N').\sin(2\pi i n/N') \end{aligned} \tag{4.11}$$

which gives

$$\begin{aligned} U_i + \jmath V_i = \sum_{k'} \sum_{n=0}^{N'-1} & \frac{X_{k'}}{2} \{\cos(\frac{2\pi(\lambda - i)n}{N'}) + \cos(\frac{2\pi(\lambda + i)n}{N'})\} \\ & - \frac{Y_{k'}}{2} \{\sin(\frac{2\pi(\lambda + i)n}{N'}) + \sin(\frac{2\pi(\lambda - i)n}{N'})\} \\ & + \frac{\jmath X_{k'}}{2} \{\sin(\frac{2\pi(\lambda + i)n}{N'}) - \sin(\frac{2\pi(\lambda - i)n}{N'})\} \\ & + \frac{\jmath Y_{k'}}{2} \{\cos(\frac{2\pi(\lambda - i)n}{N'}) - \cos(\frac{2\pi(\lambda + i)n}{N'})\}. \end{aligned} \tag{4.12}$$

Now using

$$C(z) = \sum_{n=0}^{N'-1} \cos(\frac{2\pi n}{N'} z) = \cos(\frac{\pi}{N'}(N' - 1)z).\frac{\sin(\pi z)}{\sin(\frac{\pi z}{N'})}$$

and

$$S(z) = \sum_{n=0}^{N'-1} \sin(\frac{2\pi n}{N'}z) = \sin(\frac{\pi}{N'}(N'-1)z).\frac{\sin(\pi z)}{\sin(\frac{\pi z}{N'})}$$

and equating the real and imaginary parts of eqn. (4.12) gives

$$\sum_{k'} X_{k'}(C(\lambda - i) + C(\lambda + i)) - Y_{r'}(S(\lambda - i) + S(\lambda + i)) = 2U_i \qquad (4.13)$$

$$\sum_{k'} X_{k'}(S(\lambda + i) - S(\lambda - i)) + Y_{k'}(C(\lambda - i) - C(\lambda + i)) = 2V_i. \qquad (4.14)$$

Written in matrix notation, this becomes $Ax = B$ where

$$A = \begin{pmatrix} \Gamma_{1i_1} & \Gamma_{2i_1} & \cdots & \Gamma_{K'i_1} \\ \Gamma_{1i_2} & \Gamma_{2i_2} & \cdots & \Gamma_{K'i_2} \\ \vdots & \vdots & \vdots & \vdots \\ \Gamma_{1i_F} & \Gamma_{2i_F} & \cdots & \Gamma_{K'i_F} \end{pmatrix}, B = \begin{pmatrix} \Phi_{i_1} \\ \Phi_{i_2} \\ \vdots \\ \Phi_{i_F} \end{pmatrix} \text{ and } x = \begin{pmatrix} \eta_{n'_1} \\ \eta_{n'_2} \\ \vdots \\ \eta_{n'_{K'}} \end{pmatrix}$$

with

$$\Gamma_{ef} = \begin{pmatrix} a_{ef} + b_{ef} & -d_{ef} - c_{ef} \\ c_{ef} - d_{ef} & a_{ef} - b_{ef} \end{pmatrix}, \quad \Phi_{if} = \begin{pmatrix} 2U_{if} \\ 2V_{if} \end{pmatrix} \quad , \eta_{n'_s} = \begin{pmatrix} X_{k'_s} \\ Y_{k'_s} \end{pmatrix}$$

and

$$a_{ef} = C(k'_e - f), \quad b_{ef} = C(k'_e + f), \quad c_{ef} = S(k'_e + f), \quad d_{ef} = S(k'_e - f).$$

This leads to a linear least squares estimate for X_k and Y_k. The regression matrix A must have rank $2K$, and its behaviour with different frequency sets K and truncation lengths κ needs careful attention, before any statistical analysis of the estimator is

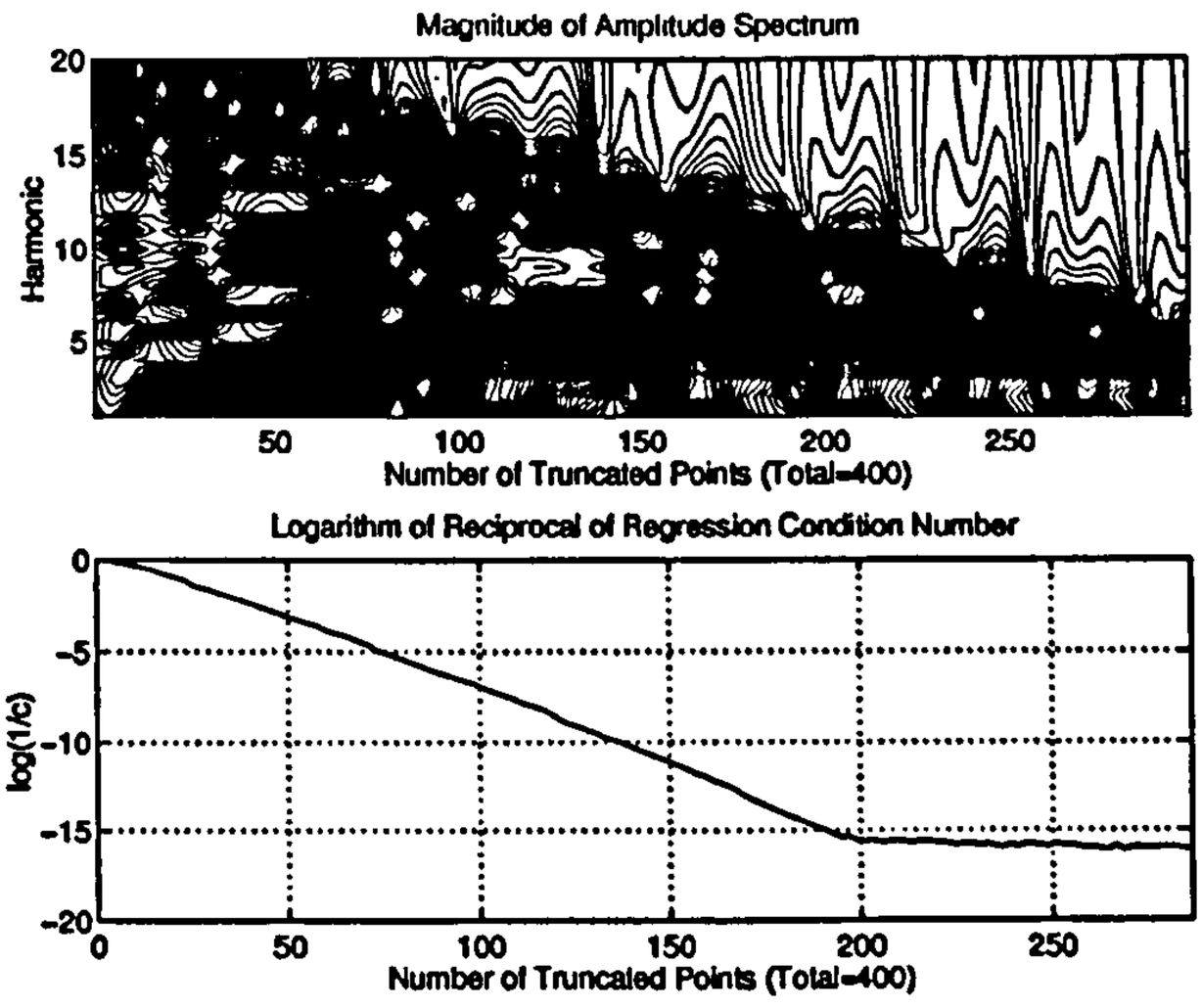

Figure 4.2: Upper: contour plot of amplitude spectrum of signal (N =512) containing uniform power in harmonics $1, 2 \cdots 20$ for truncated data points $1, 3 \cdots 299$, with dark corresponding to high energy. Lower: reciprocal of the condition number of the regressor for the given truncation lengths.

carried out.

The numerical ill-conditioning of A coincides with the DFT spectrum of the truncated signal having less than K dominant peaks. The maximum value of $\lambda_{r'}$ dictates this. To visualise the degeneration in rank, the absolute value of the DFT spectrum is shown in Figure 4.2, together with the reciprocal of the condition number of A (computed using singular value decomposition), for a 20 consecutive-harmonic signal ($k_{min} = 1$) being increasingly shortened. From this, the truncation method could only be used confidently with at least three-quarters of the signal. The loss in rank is complete at around $N/2$ ($\max(\lambda_{r'}) = 10.45$). Repeating the procedure with a signal containing the first 40 consecutive-harmonics of the fundamental, the regressor loses rank at about $N/3$.

This method of eliminating the transient information will seldom be used with very large values for κ compared to N, and so the results presented above do not

preclude its use. However the numerical ill-conditioning of the algorithm is potentially dangerous, and so frequency sets which alleviate the problem are useful. The immediate choice is an inverse-repeat signal, and so now the numerical conditioning of the algorithm with frequency sets containing only odd-harmonics is examined. The conditioning of the regressor as the number of truncated data points increases is shown in Figure 4.3, together with the corresponding DFT spectrum of a bandwidth-limited inverse-repeat signal, containing uniform power in harmonics $1, 3 \cdots 39$. It can be seen from Figure 4.3 that the regressor is well conditioned up to

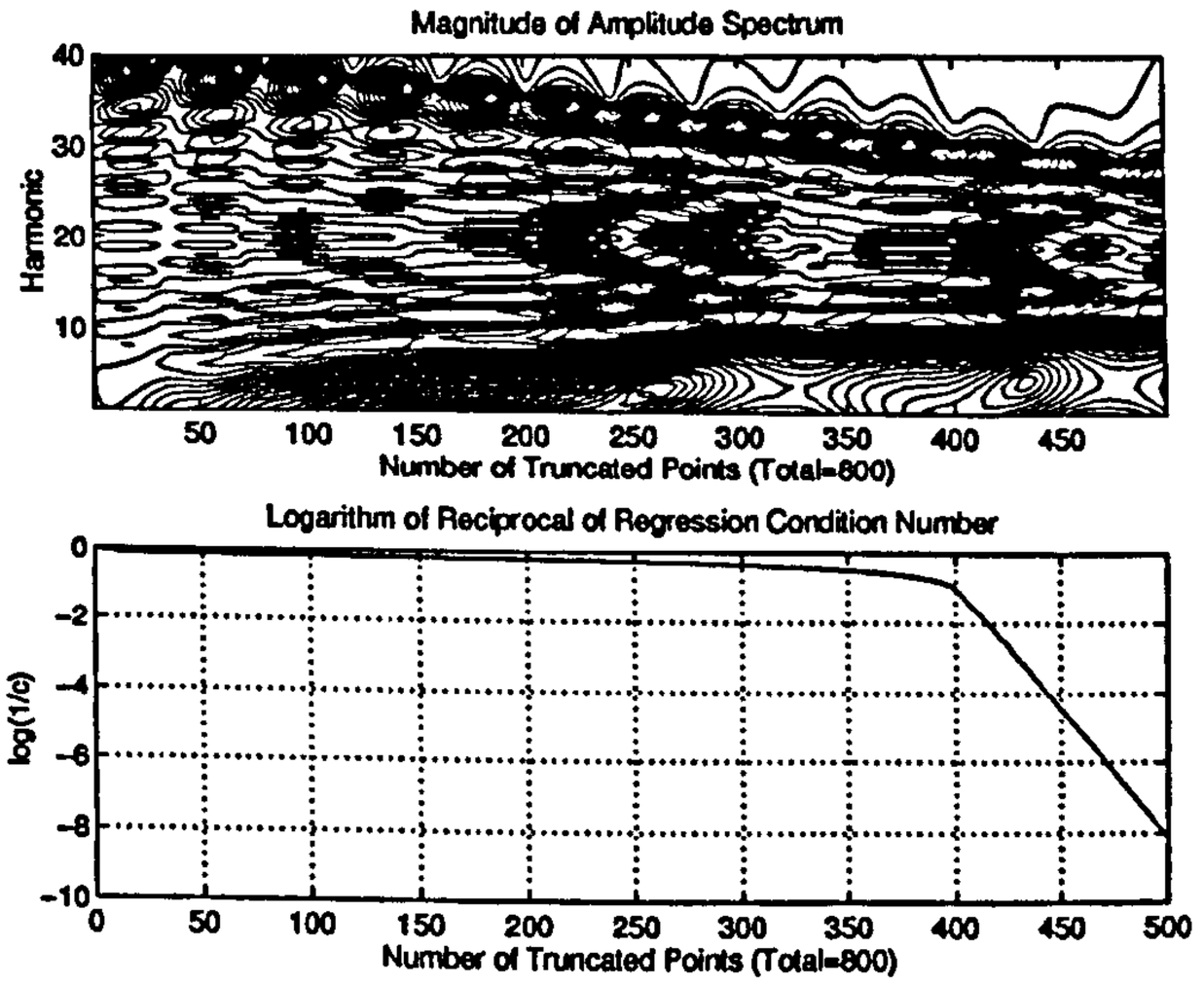

Figure 4.3: Upper: contour plot of amplitude spectrum of signal (N=512) containing uniform power in harmonics $1, 3 \cdots 39$ for truncated data points $1, 3 \cdots 299$, with dark corresponding to high energy. Lower: reciprocal of the condition number of the regressor for the given truncation lengths.

a truncation length of half the period length, at which point similar behaviour to the consecutive-harmonic case is observed. For this reason the accuracy of the Fourier coefficients estimated with the IFFT algorithm will be examined for band-limited inverse-repeat signals.

4.3.1 Accuracy of the Fourier Coefficients

Since the interpolated FFT algorithm used in this work processes quantities which result from a DFT, the corrupting noise can be assumed to have a Gaussian distribution and diagonal covariance matrix. This assumption is valid for data records of practical lengths ($N \geq 512$) (Schoukens and Renneboog, 1986). The multiharmonic signal is an odd-harmonic signal ($1, 3 \cdots 39$, $N = 1024$), and the maximum value of κ is 599. The theoretical and empirical standard deviations of the real and imaginary parts of the estimated Fourier coefficients are calculated for an increasing value of κ. The standard deviations are estimated from 100 independent records for each truncation length. The results are shown in Figures 4.4–4.9. From the results, it can be seen that the assumption of a diagonal covariance matrix for the noise gives excellent agreement between the theoretical and estimated standard deviations of the Fourier coefficients. A discrepancy is only noticeable when a very large value of κ is being used (short data record). However the main conclusion that can be drawn from the results is that the uncertainty of the Fourier coefficient estimates is no greater than double that of the DFT used with a transient free full data record, even when nearly half of the signal is being discarded.

4.3.2 The IFFT in the Presence of Drift

The expression (4.7) for the DFT of the truncated multifrequency signal can be modified to include the contribution to the Fourier coefficients of the polynomial

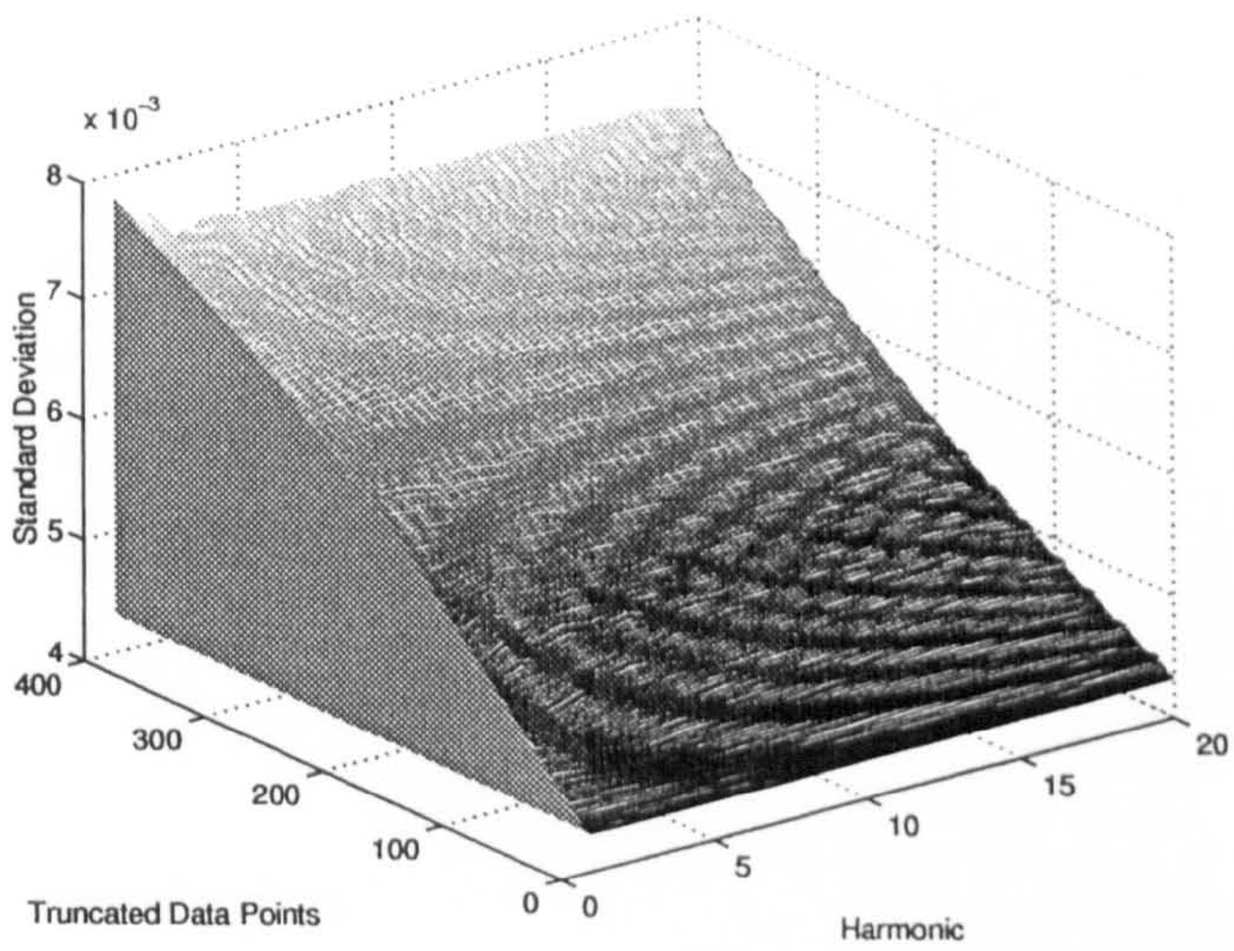

Figure 4.4: Theoretical standard deviation of the real part of the complex Fourier coefficients calculated via the IFFT (I).

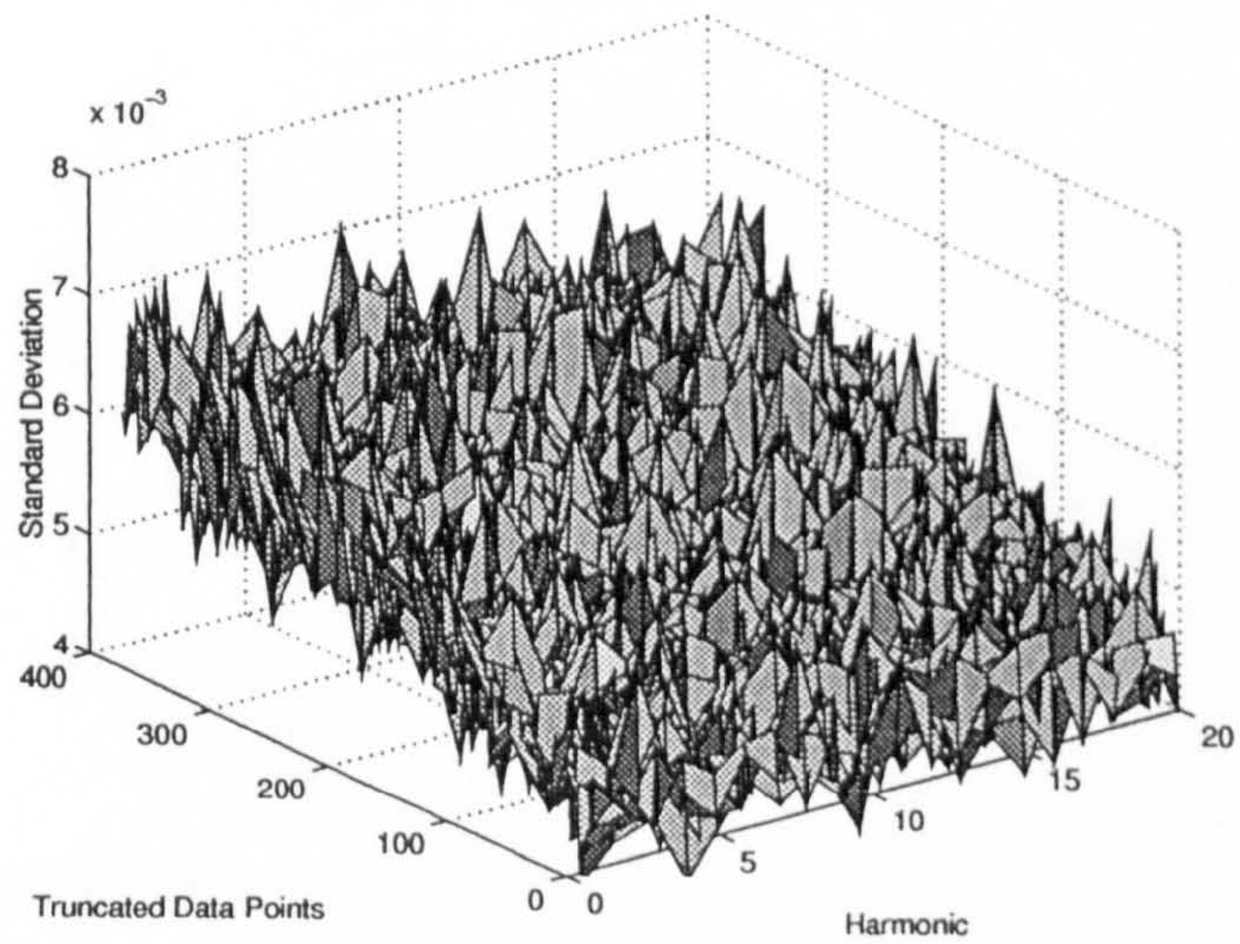

Figure 4.5: Estimated standard deviation of the real part of the complex Fourier coefficients calculated via the IFFT (I).

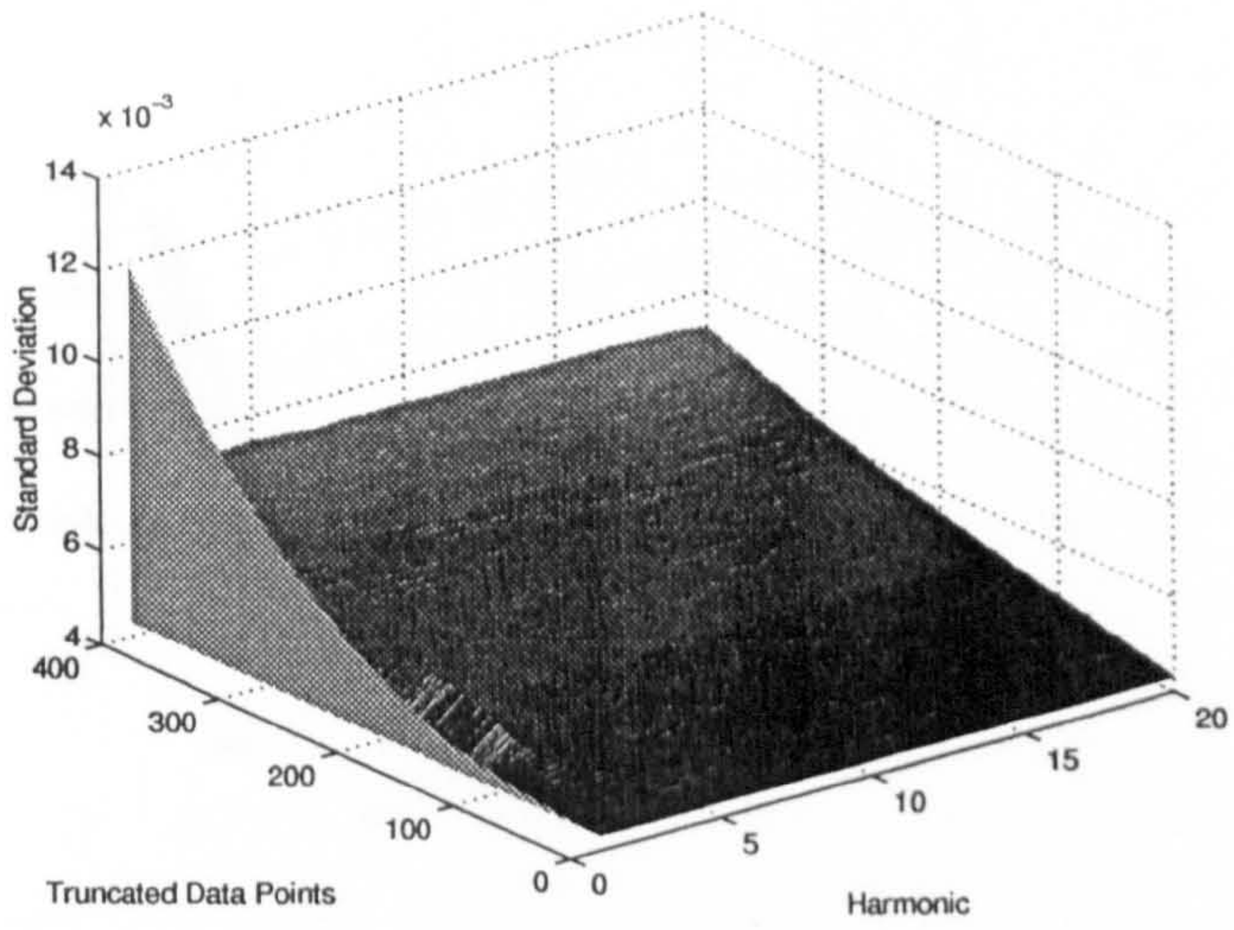

Figure 4.6: Theoretical standard deviation of the imaginary part of the complex Fourier coefficients calculated via the IFFT (I).

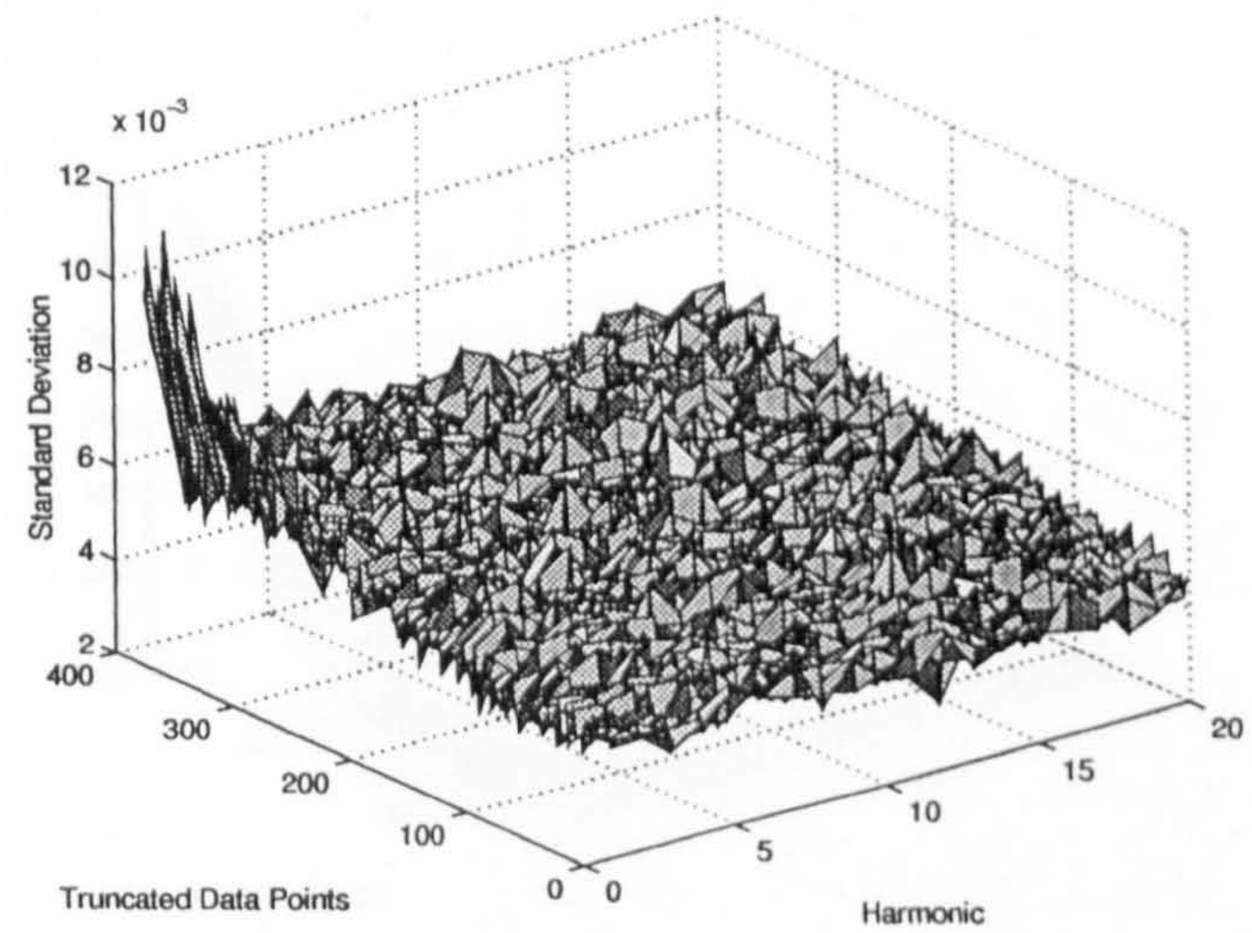

Figure 4.7: Estimated standard deviation of the imaginary part of the complex Fourier coefficients calculated via the IFFT (I).

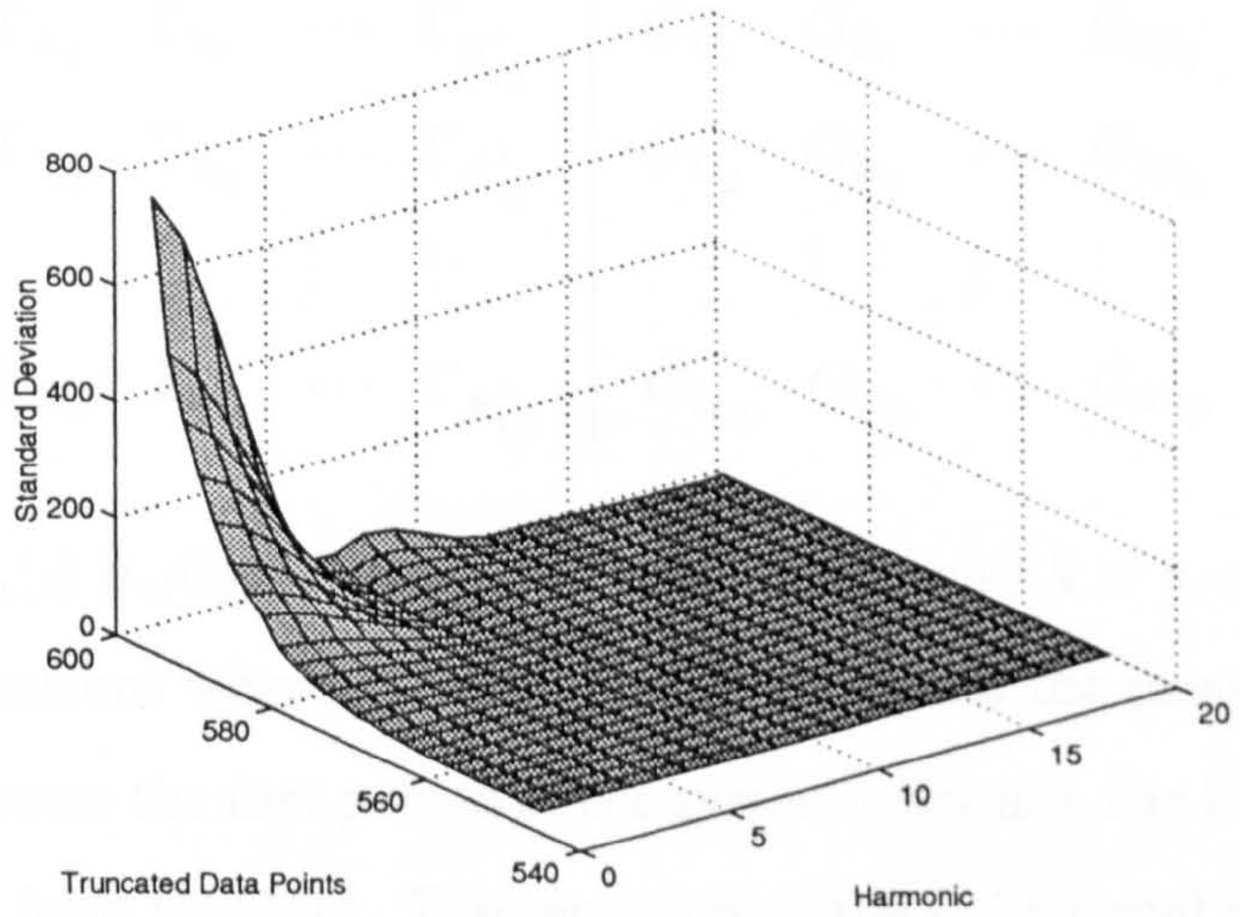

Figure 4.8: Theoretical standard deviation of the real part of the complex Fourier coefficients calculated via the IFFT (II).

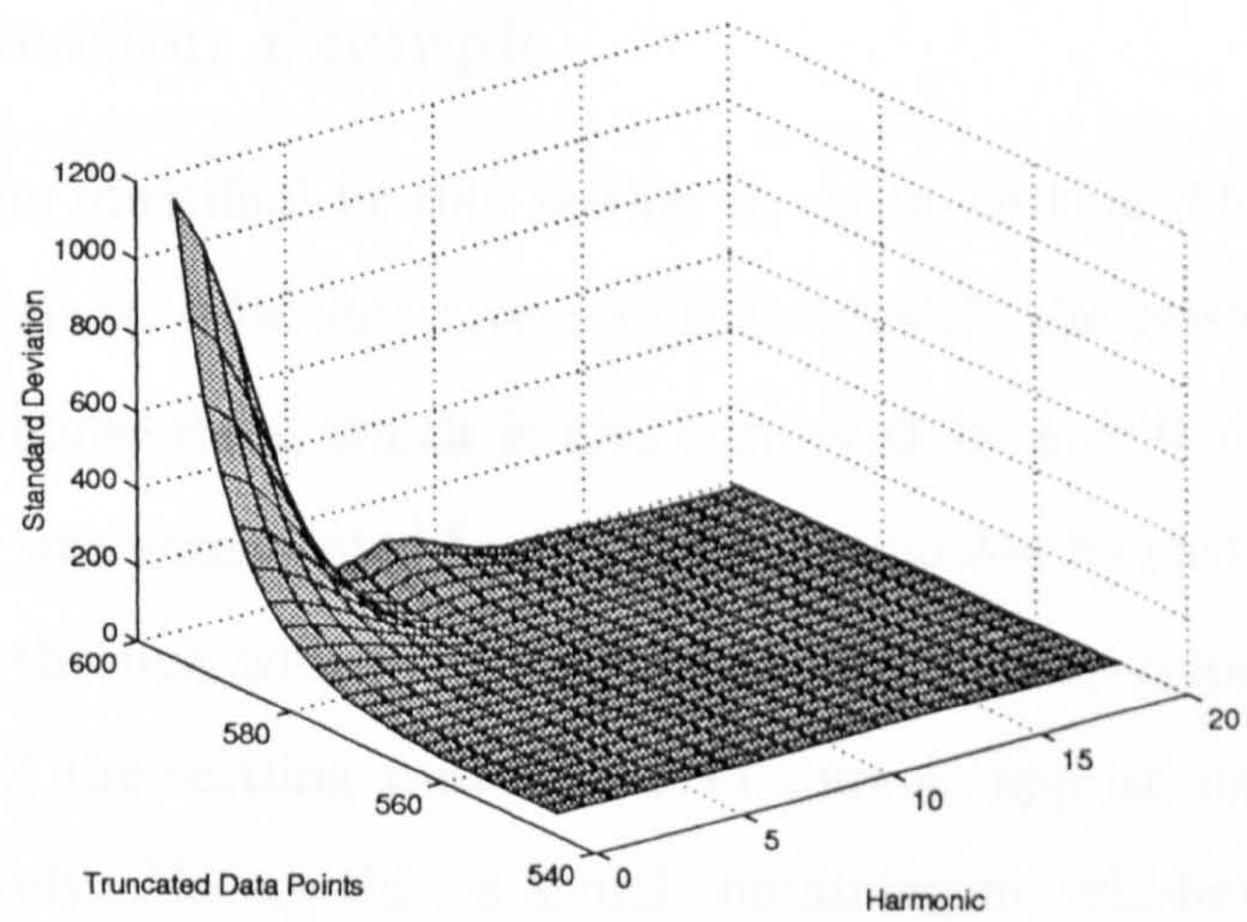

Figure 4.9: Estimated standard deviation of the real part of the complex Fourier coefficients calculated via the IFFT (II).

drift representation in (3.2). The resulting regressor is

$$A' = \left(\begin{array}{cccc|cccc} \Gamma_{1i_1} & \Gamma_{2i_1} & \cdots & \Gamma_{K'_{i_1}} & G_{1i_1} & G_{2i_1} & \cdots & G_{Mi_1} \\ \Gamma_{1i_2} & \Gamma_{2i_2} & \cdots & \Gamma_{K'_{i_2}} & G_{1i_2} & G_{2i_2} & \cdots & G_{Mi_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Gamma_{1i_F} & \Gamma_{2i_F} & \cdots & \Gamma_{K'_{i_F}} & G_{1i_F} & G_{2i_F} & \cdots & G_{Mi_F} \end{array}\right) \tag{4.15}$$

where $G_{m_i} = [\beta_m(i)\ \alpha_m(i)]^T$ from (3.11) with $P = 1$ and $k = \lambda, \cdots, N$. Hence for experimental situations where a drift disturbance affects the measurements, and it is desirable to process the first period of the system response, the Fourier coefficients can be extracted from the data. The inclusion of the polynomial representation for the drift increases the unknown parameters by only a few terms (possibly just one) and so the estimation accuracy will be similar to that obtained with the IFFT algorithm. Use is made of the ability to suppress both drift and transient contributions in the following application example.

4.3.3 Application Example

The process to be identified in this section is, as in Section 3.5.2, the Feedback PT326 process trainer. The objective here is to identify the system using the first period of the response data, which is also corrupted by a drift disturbance. This drift disturbance was accentuated for illustration purposes by partially covering the end of the tube, through which the heated air exits. A step-response test gave an initial estimate of the settling time and dead time as approximately 2 and 0.125 seconds respectively. Noting this, a signal containing all odd-harmonics up to 49 with a fundamental of 0.125 Hz (same as in example of Section 3.5.2) was used to excite the system. The input and output time series for the experiment are shown in Figure 4.10. As can be seen the output data exhibits a significant drift,

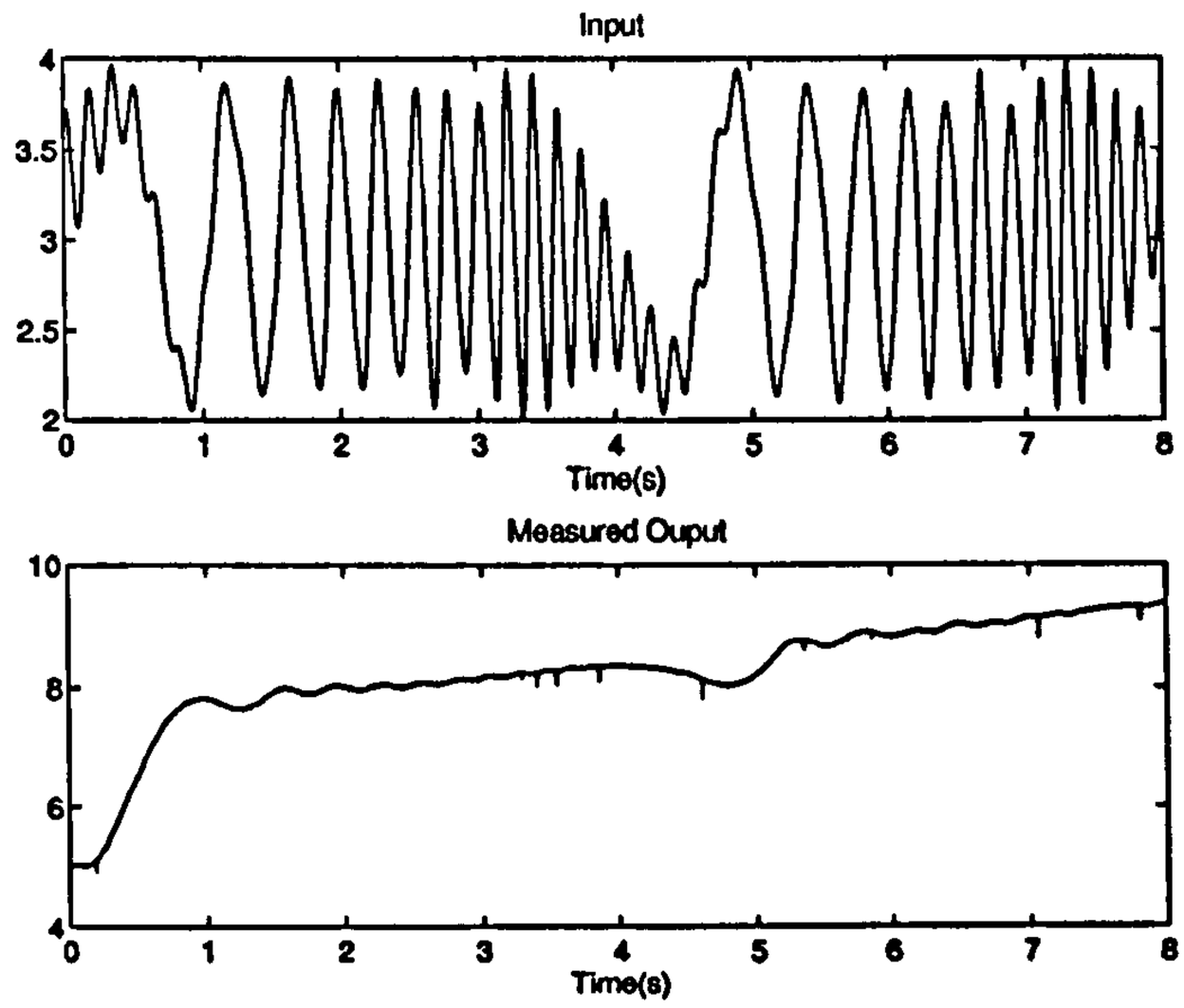

Figure 4.10: Transient and drift corrupted input and output time series obtained from the Feedback PT326 process trainer

while because the experiment involved moving to a different operating level, the output signal contains considerable transient information. Three frequency response functions were estimated from the data:

- Estimate obtained directly from the input and output signals.
- Estimate obtained using IFFT algorithm (Section 4.3) with the initial 2 seconds of data being discarded.
- Estimate obtained with the IFFT algorithm including a linear drift model (Section 4.3.2) with initial 2 seconds of data being discarded.

These estimates are shown in Figure 4.11. Note that the phase response of the estimate obtained from the raw measurements is not shown as it was impossible to interpret. The immediate benefit of removing the transient contribution is clear from the figure, but it can be seen that the estimate at the first harmonic frequency is still significantly in error. The estimate obtained after the drift has been removed,

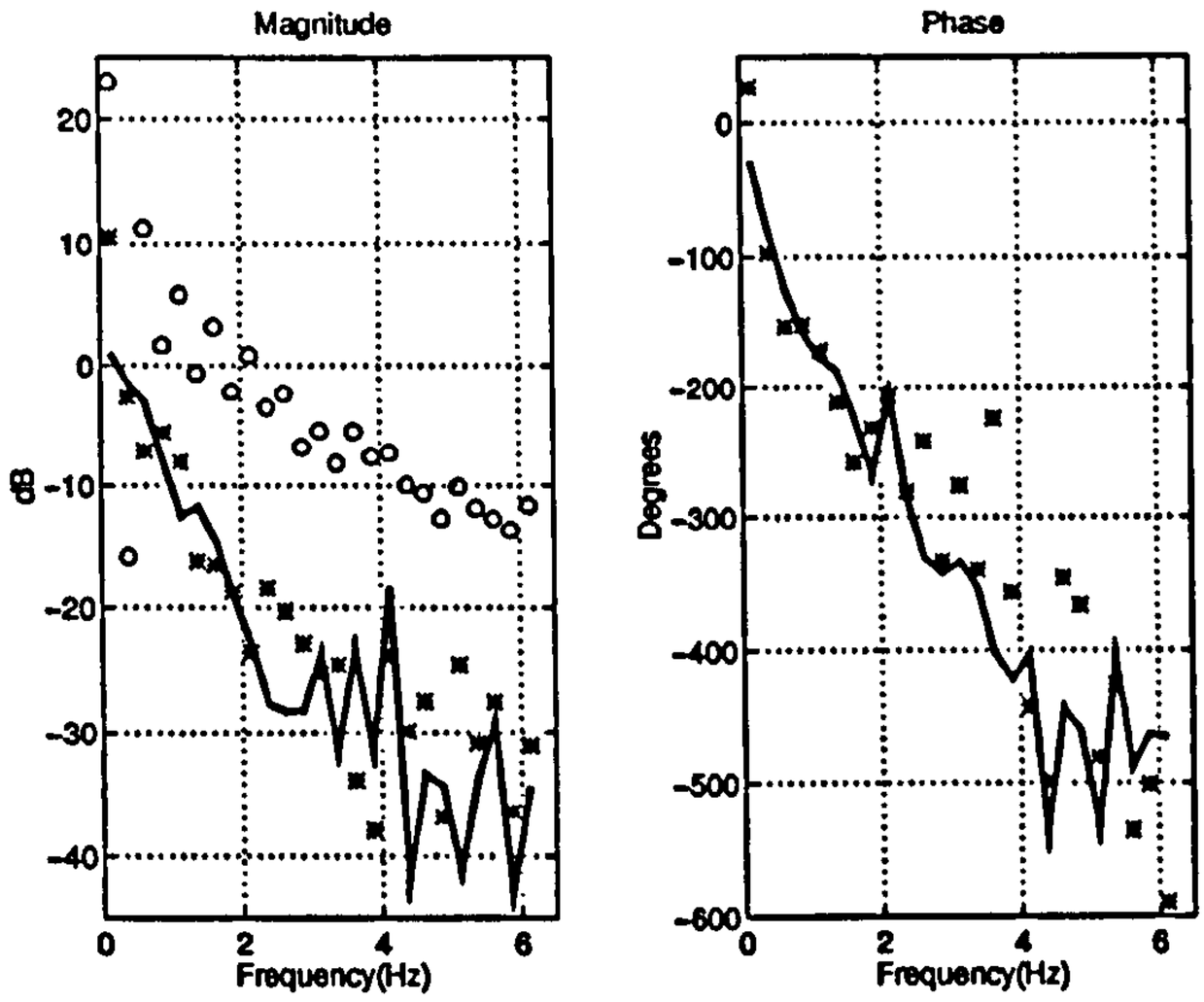

Figure 4.11: The estimate of the frequency response function for the Feedback PT326 process trainer. (o): estimate from raw measurements. (*): estimate removing transient contribution. (-): estimate after both drift and transient contributions have been removed.

shows that this low frequency error has been corrected. This example shows quite clearly that the transient suppression technique works with real experimental data.

4.4 Conclusions

Two practical techniques were presented in this chapter that can suppress the effects of initial transients on the identification results, when a periodic signal is used as the excitation. These find application when it is desirable to process the first period of the response signal, and so are expected to be relevant in identification applications which involve systems with slow dynamics or when experimental time is at a premium.

The first of the methods is designed to work in situations where the system dynamics are dominated by a single long time-constant. It was shown that approx-

imate knowledge of the dominant pole (e.g. obtained from a simple step-response test), can be used to suppress the corresponding transient term.

The second technique is an application of an interpolated fast Fourier transform algorithm, and works solely on the premise that the first period of the response is corrupted by transients for only a fraction of the total duration of the signal. The limits to the method were found to be situations where a signal period equal to three to four times the settling time is used for a consecutive harmonic excitation signal. The signal period could be decreased to about twice the settling time of the system if an inverse-repeat signal is used. Though extending the applicability of the proposed method, system identification experiments seldom use signal periods as low as this since there is considerable scope to miss important features of the system dynamics. The method was shown to be compatible with a drift estimation procedure given in Chapter 3. The signal processing techniques were demonstrated on measurements obtained from a hot-air flow device and were found to give excellent results.

Chapter 5

Periodic Signal Designs for Nonlinear System Identification

5.1 Abstract

The accuracy of nonlinear parametric model estimates is examined for experimental situations which allow the application of a periodic excitation signal. The emphasis of the chapter is on the influence of the excitation properties on the identification results, particularly when the parameter estimation is carried out using algorithms which do not require a special form of excitation.

5.2 Introduction

Parametric approaches to nonlinear system identification are gaining increasing popularity in the literature (Chen, Billings and Luo, 1989; Billings and Zhu, 1991; Patra, 1993). Out of the many possible parametrisations, the discrete-time nonlinear ARMAX (NARMAX) model (Leontaritis and Billings, 1985) and its continuous-time counterpart (Tsang and Billings, 1994) have been developed to a stage where structure detection and parameter estimation can be carried out routinely (Korenberg, Billings, Liu and McIlroy, 1988; Chen et al., 1989). The identification procedures developed to carry out the system parameter estimation are invariably in the time-domain and so, as a result, little attention has been given to the design of the signal used to perturb the system under investigation. Furthermore, since the familiar pseudorandom binary signal is usually unsuitable as an excitation when a nonlinear model is sought, random excitations are applied in the majority of applications.

Although the use of random signals is generally adequate for identification, it is interesting to take note of recent developments in the area of perturbation signal design for *linear* system identification (Godfrey, 1993c; Kollár, 1994; Schoukens et al., 1994), and to determine whether any distinct advantages are to be gained by utilising a periodic excitation in nonlinear system identification. For multi-input

linear system identification, periodic excitations allow the straightforward creation of uncorrelated sets of signals which can, at least, decouple the identification into scalar relationships (Briggs and Godfrey, 1966), thus simplifying the problem considerably. This ability to suppress energy from certain harmonics in the spectrum of the excitation has direct application to the detection of nonlinear effects, where simple processing procedures can highlight their presence (McCormack et al., 1995). One can go further, and use the energy in the response signal, which can only be attributed to nonlinear effects, to identify certain nonlinear models for the system. However, the dependence of the identification strategy on the properties of the input signal is, in these cases, too great to establish whether this type of periodic excitation is useful in the general sense.

The objective here is to examine, through examples, the benefits of using various types of periodic excitations in nonlinear system identification strategies which do not *explicitly* utilise their special properties. Consideration is not given to any possible simplifications which may be possible when using periodic excitations except in the case of continuous-time model estimation, where frequency-domain relations can be used to good effect. To begin with, the properties of periodic signals are discussed in Section 5.3. Following this, the parametric models considered in this chapter are summarised in Section 5.4, together with the identification procedures used for parameter estimation. The application of various periodic signal designs to several nonlinear models is discussed in Section 5.5. Finally in Section 5.6, some conclusions are drawn.

5.3 Periodic Signal Designs

Recent directions in perturbation signal design for linear system identification have concentrated on signals which are capable of realizing an arbitrary set of Fourier

specifications (Godfrey, 1993c; Kollár, 1994; McCormack et al., 1995). The most popular signals are the multisine and discrete-interval binary signals. Due to the unsuitability of binary signals for nonlinear system identification (Leontaritis and Billings, 1987), the properties of multisine signals will be examined here, although multi-level signals (McCormack et al., 1995) are equally applicable. For linear system modelling the complete test signal design procedure involves the selection of a power spectrum, followed by the optimisation of a time-domain realization (Schoukens et al., 1993a). Furthermore, for multiple-input system identification, uncorrelated sets of signals can be designed in a straightforward manner.

Input signal design for nonlinear system identification is still an open problem, but for many of the parametric techniques, the requirements of the excitation signal can be satisfied by using random noise. As with linear system identification, an increased power content is a direct benefit of using periodic signals, due to the availability of powerful periodic signal design procedures (Schoukens et al., 1993a). The ability to control the power spectrum of the excitation also seems to be a potential benefit of using a periodic signal when a nonlinear model is sought. For example, the suppression of harmonics from the power spectrum of the excitation can aid the identification of certain nonlinear models (Barker and Al Hilal, 1985; Schoukens, Pintelon and Renneboog, 1988a). Generally, the rationale behind this is that coherent energy at these frequencies can only be attributed to nonlinear behaviour, and so is of use for parameter estimation. This orthogonality can be directly utilised during the identification of input-nonlinear models (e.g. Hammerstein models), where the suppression of certain harmonics can result in partially uncorrelated signals available for the identification of the linear dynamics. However the benefits, if any, of using these special classes of periodic excitations for parametric nonlinear identification have not been investigated.

To this end, several periodic signal designs will be examined here. The first

design considered is a signal containing the first F harmonics of the fundamental, the second has the same bandwidth but contains zero energy at all even-harmonics, while the third has zero energy at all even-harmonics and every second odd-harmonic. The time-domain properties of these signals are optimised using the time-frequency domain clipping algorithm (Van der Ouderaa et al., 1988a) available as part of the Frequency Domain System Identification Toolbox (Kollár, 1994). The total desired power in each of the signals is the same, with the signals specified to have an approximately equal peak factor (Godfrey, 1993a). Each of these signals will be used to identify a discrete-time nonlinear model for several nonlinear systems. In all cases random-noise excitations will be included for comparison purposes.

5.4 Parametric Nonlinear Models

Polynomial single-input, single-output (SISO) nonlinear models are focussed upon here. These are capable of modelling many systems encountered in practice, but are most applicable when low-order approximations are possible. The polynomial nonlinear ARMAX (NARMAX) model has been shown to be a practical means of modelling nonlinear systems, when used with suitable structure detection and parameter estimation techniques (Billings and Voon, 1986; Chen et al., 1989). The model is of the form,

$$y(k) = F^R(y(k-1), \cdots, y(k-n_y), u(k-1), \cdots, u(k-n_u), e(k-1), \cdots, e(k-n_e)) + e(k) \qquad (5.1)$$

where n_y, n_u and n_e correspond to the maximum lags in the output, input, and noise signals respectively, and F^R represents an R^{th} order nonlinear function. Assuming additive output noise, the NARMAX representation is conceptually similar to the

m-input ARMAX representation

$$A(q)y(k) = B_1(q)u_1(k) + B_2(q)u_2(k) + \cdots + B_m u_m(k) + C(q)e(k) \qquad (5.2)$$

where the signals $u_1(k) \cdots u_m(k)$ are formed from the polynomial expansion of $(u(k), y(k))$. The main distinction between the models of eqn. (5.1) and eqn. (5.2) is the presence of multiplicative noise terms in the former, even when the noise is additive at the output. This means that identification procedures based on the model of eqn. (5.2) will produce biased parameter estimates (Billings, 1985). As is generally the case with discrete-time identification algorithms, an arbitrary excitation may be used as an input signal with the NARMAX representation.

The deterministic continuous-time counterpart of the NARMAX representation is the nonlinear differential equation model, given by

$$\sum_{i=0}^{m} b_i \frac{d^i u(t)}{dt^i} = \sum_{i=0}^{n} a_i \frac{d^i y(t)}{dt^i} + F^R[y^n(t), \cdots, y(t), u^m(t), \cdots, u(t)] \qquad (5.3)$$

which has been examined by various authors (Patra, 1993; Tsang and Billings, 1994), where identification procedures have been developed. In these contributions, no treatment of the noise model is present, and so linear least squares procedures are postulated. The numerical integration identification method postulated by Tsang and Billings (1994) can be used with arbitrary excitations, whereas the use of the discrete Hartley transform by Patra (1993) implicitly requires a periodic excitation. Hence the periodic signal designs considered here are directly applicable to the identification method developed by Patra (1993).

Discrete-time models are focussed upon here, and the parameter estimation is carried out using the System Identification Toolbox (Ljung, 1986). There is expected to be a bias in the parameter estimates through an inadequate noise model (Billings, 1985), however at the noise levels considered in Section 5.5 this is seen to

be insignificant.

5.5 Application Examples

Three representative nonlinear models will be used in this section to examine the performance of the periodic signals described in Section 5.3. The first, a Hammerstein model, consists of a third-order polynomial in series with a linear system, the transfer function of which is given by,

$$H(s) = \frac{1}{(s+1)^3}. \tag{5.4}$$

The input signals used for the identification are zero-order-hold multisines (Schoukens et al., 1994) and were specified to have uniform power in harmonics $1, 2, 3 \cdots 399$ (Signal 1), $1, 3, 5 \cdots 399$ (Signal 2), and $1, 5, 9 \cdots 399$ (Signal 3), with resulting peak factors of 0.9823, 0.9915 and 0.9953 respectively. The maximum value of the amplitude of each input signal was normalised to unity, and the response of the system was disturbed with zero-mean normally distributed noise of variance $(1/40)^2$. The spectrum of each output signal is shown in Figure 5.1. The input and output data were used to obtain a representative NARMAX model of the form

$$\begin{aligned}(1 + a_1q^{-1} + a_2q^{-2} + a_3q^{-3})y(k) = (bl_1q^{-1} + bl_2q^{-2} + bl_3q^{-3})u(k) + \\ (bq_1q^{-1} + bq_2q^{-2} + bq_3q^{-3})u^2(k) + (bc_1q^{-1} + bc_2q^{-2} + bc_3q^{-3})u^3(k) + \\ C(q)e(k) \quad (5.5)\end{aligned}$$

The model identification was repeated 50 times with Table 5.1 showing the mean ($\hat{x}$) and standard-deviation (σ_x) of the resulting parameter estimates. Signal 4 corresponds to a normally distributed random excitation with the same peak-value as the other three signals.

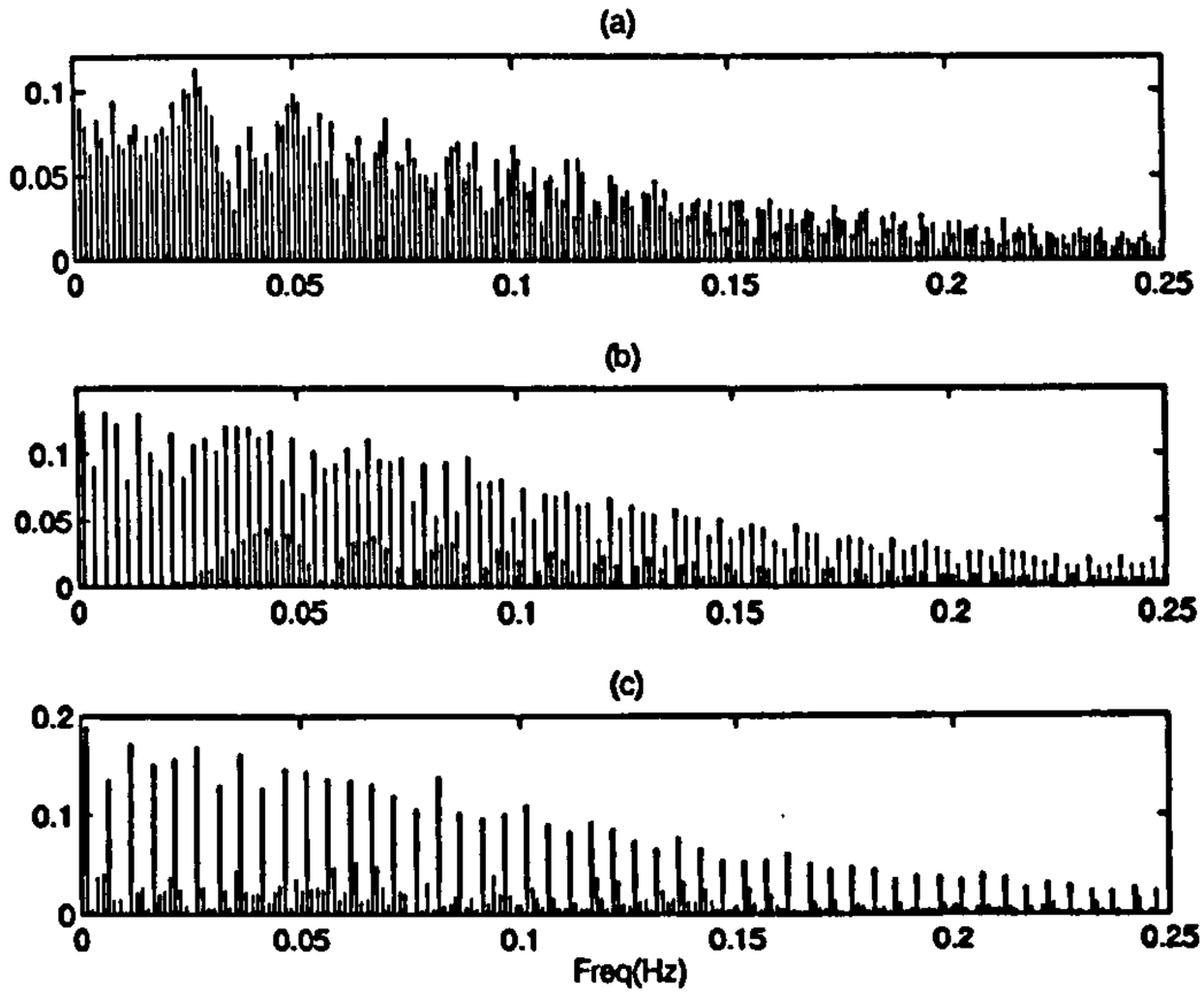

Figure 5.1: Amplitude spectra of response signals for Hammerstein model example

The results from this example indicate that all three periodic signals give parameter estimates having an uncertainty which is typically a third of that obtained with the random signal. However there is no significant difference between the results obtained with the three periodic signals, with any discrepancy being due to the unequal peak factors of the three signals.

The second example uses a nonlinear stiffness model, with a model equation which is given by

$$u(t) = \frac{d^2y(t)}{dt^2} + 2\frac{dy(t)}{dt} + y(t) + y^2(t) + y^3(t). \tag{5.6}$$

The three input signals have the same specified power spectra as in the last example, but now are of the bandwidth-limited type. The peak factors are 1.1418, 1.1458, and 1.1459 for Signal 1, Signal 2 and Signal 3 respectively, with, this time, the maximum value of the input signal amplitudes scaled to 2. The spectra of the three output signals are shown in Figure 5.2. A discrete-time model of the form

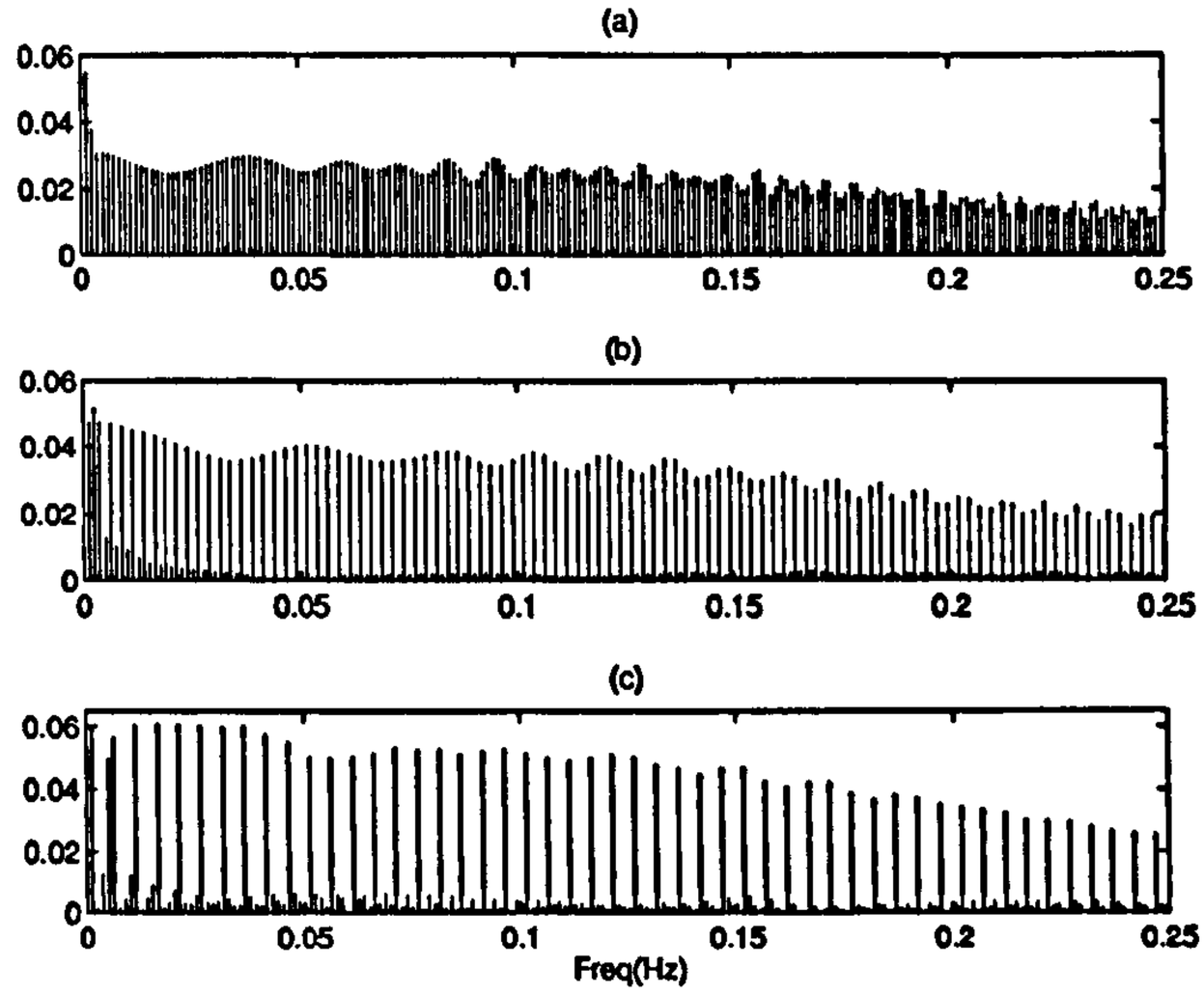

Figure 5.2: Amplitude spectra of response signals for nonlinear stiffness model example

$$(1 + a_1q^{-1} + a_2q^{-2})y(k) = (b_1q^{-1} + b_2q^{-2})u(k) + (aa_1q^{-1} + aa_2q^{-2})y^2(k) + (aaa_1q^{-1} + aaa_2q^{-2})y^3(k) \quad (5.7)$$

was specified, and the parameter estimation carried out 50 times with the resulting mean and standard deviation shown in Table 5.2. Again all three periodic input signals perform equally well, with the uncertainty obtained with the random excitation being very large for the parameters of the nonlinear portion of the model.

The final model is the NARMAX model

$$\begin{aligned}(1 + a_1q^{-1} + a_2q^{-2} + a_3q^{-3})y(k) = (b_1q^{-1} + b_2q^{-2} + b_3q^{-3})u(k) + \\ (ab_1q^{-1} + ab_2q^{-2})u(k)y(k) + aaq^{-2}y^2(k) + aaaq^{-2}y^3(k) + \\ aabq^{-2}y^2(k)u(k) + bbbq^{-2}u^3(k)\end{aligned} \quad (5.8)$$

which was identified by Billings (1985) as representing the input flow rate to the

height of the liquid in a pair of coupled tanks. The input signals described in the first example were used here. The output spectra are shown in Figure 5.3, with the results of 50 identification trials shown in Table 5.3. These results follow the same

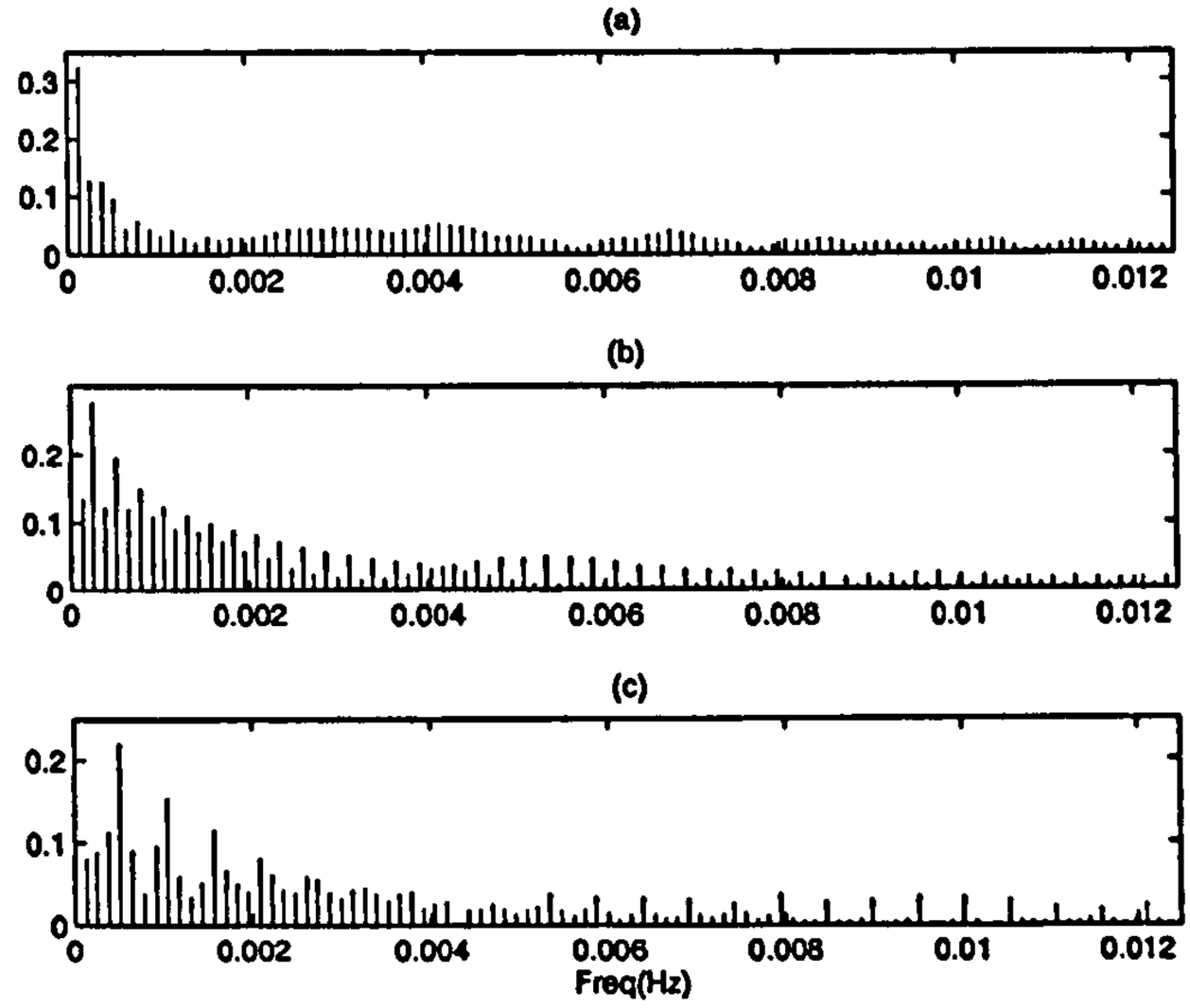

Figure 5.3: Amplitude spectra of response signals for NARMAX model example

pattern as those obtained in the first two examples.

The estimation procedure was repeated for the first two examples, with input signals containing uniform power in harmonics $1, 2, 3 \cdots 49$; $1, 3, 5 \cdots 49$; and $1, 5, 9 \cdots 49$. For the Hammerstein model, this proved to give extremely unreliable parameter estimates, although the output time-series and predicted output signal, shown in Figure 5.4, were in good agreement. With the nonlinear stiffness model, the recorded accuracy of the results obtained with the consecutive harmonic signal decreased by as much as a factor of four, with the accuracy obtained with the non-consecutive odd-harmonic signal remaining virtually unchanged. The inaccuracy of the parameter estimates, although not highlighted by the output time-series is indicative of the need for input signals with greater harmonic content (in the

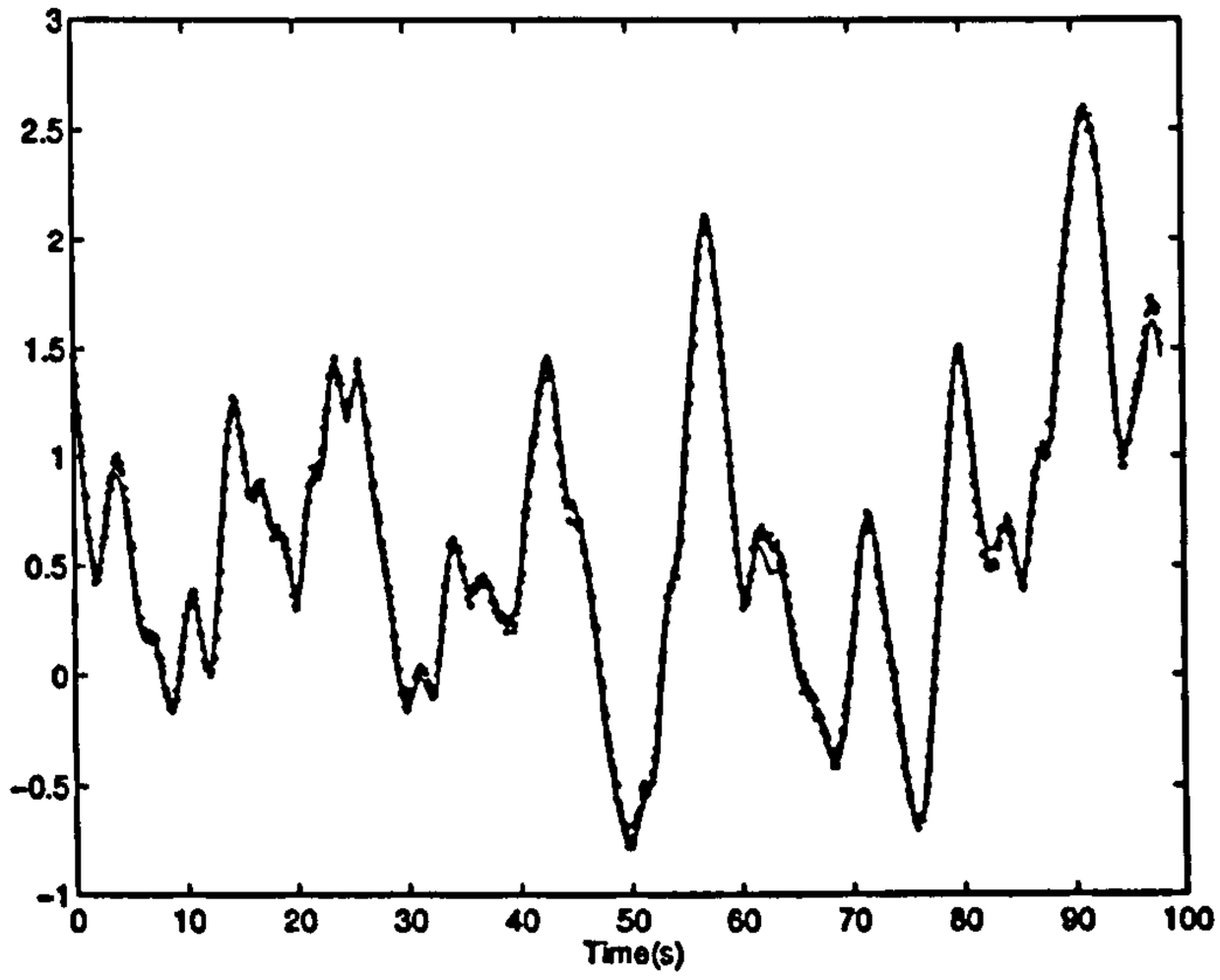

Figure 5.4: Actual and estimated response signal for Hammerstein model with 49 harmonic excitation

Hammerstein example 49 excitation harmonics proved insufficient with 15 free parameters). Hence even if the number of parameters to be estimated is low, the input signal is required to have a very high harmonic content. Although this is an advantage of using random signals, the typical low power content of these signals, when compared to periodic signals, outweighs this advantage. Periodic signals with broadband power spectra are seen to give reliable results in all the examples considered here.

5.6 Conclusions

It is seen that broadband periodic signals with a high harmonic content give superior parameter estimates for a variety of parametric nonlinear models. In this category of excitation signals, inverse-repeat and non-consecutive odd-harmonic signals perform equally well, and so the secondary properties of such signals, e.g. their advantages for nonlinear detection, can be utilised without any loss in accuracy.

P	*Value*	$\bar{x}, \sigma_x$ *Signal 1*	$\bar{x}, \sigma_x$ *Signal 2*	$\bar{x}, \sigma_x$ *Signal 3*	$\bar{x}, \sigma_x$ *Signal 4*
a_1	-1.376	-1.374, 0.035	-1.377, 0.036	-1.379, 0.041	-1.338, 0.144
a_2	0.631	0.629, 0.044	0.633, 0.045	0.635, 0.051	0.583, 0.185
a_3	-0.097	-0.096, 0.015	-0.097, 0.016	-0.098, 0.018	-0.080, 0.065
b_{l1}	0.045	0.045, 0.005	0.045, 0.004	0.045, 0.005	0.045, 0.007
b_{l2}	0.100	0.100, 0.008	0.100, 0.007	0.100, 0.007	0.101, 0.011
b_{l3}	0.014	0.014, 0.007	0.014, 0.007	0.013, 0.007	0.019, 0.017
b_{q1}	0.089	0.089, 0.002	0.089, 0.002	0.089, 0.002	0.090, 0.009
b_{q2}	0.201	0.201, 0.004	0.200, 0.005	0.200, 0.005	0.202, 0.018
b_{q3}	0.028	0.028, 0.010	0.027, 0.010	0.028, 0.011	0.038, 0.036
b_{c1}	0.134	0.133, 0.007	0.132, 0.005	0.133, 0.006	0.133, 0.028
b_{c2}	0.301	0.302, 0.011	0.302, 0.009	0.300, 0.011	0.311, 0.044
b_{c3}	0.041	0.042, 0.015	0.041, 0.014	0.041, 0.015	0.052, 0.062

Table 5.1: Results achieved for the identification of the third-order Hammerstein model

P	*Value*	$\bar{x}, \sigma_x$ *Signal 1*	$\bar{x}, \sigma_x$ *Signal 2*	$\bar{x}, \sigma_x$ *Signal 3*	$\bar{x}, \sigma_x$ *Signal 4*
a_1	-1.543	-1.543, 0.003	-1.543, 0.003	-1.543, 0.003	-1.541, 0.006
a_2	0.595	0.595, 0.003	0.594, 0.003	0.596, 0.003	0.594, 0.005
b_1	0.028	0.028,0.339e-3	0.029,0.441e-3	0.029,0.402e-3	0.029,0.645e-3
b_2	0.024	0.024,0.529e-3	0.024,0.695e-3	0.024,0.641e-3	0.024,0.938e-3
aa_1	-0.028	-0.029, 0.002	-0.029, 0.002	-0.028, 0.002	-0.027, 0.008
aa_2	-0.024	-0.024, 0.002	-0.024, 0.002	-0.024, 0.002	-0.025, 0.008
aaa_1	-0.028	-0.028, 0.002	-0.028, 0.002	-0.028, 0.002	-0.025, 0.030
aaa_2	-0.024	-0.024, 0.002	-0.024, 0.002	-0.024, 0.002	-0.025, 0.030

Table 5.2: Results achieved for the identification of the nonlinear stiffness model

P	*Value*	$\bar{x}, \sigma_x$ *Signal 1*	$\bar{x}, \sigma_x$ *Signal 2*	$\bar{x}, \sigma_x$ *Signal 3*	$\bar{x}, \sigma_x$ *Signal 4*
a_1	-0.436	-0.437, 0.015	-0.440, 0.018	-0.438, 0.010	-0.461, 0.086
a_2	-0.681	-0.681, 0.008	-0.679, 0.010	-0.680, 0.007	-0.673, 0.031
a_3	0.149	0.150, 0.01	0.153, 0.012	0.150, 0.006	0.165, 0.055
b_1	0.396	0.396, 0.002	0.396, 0.002	0.396, 0.002	0.397, 0.004
b_2	0.014	0.013, 0.007	0.012, 0.010	0.013, 0.006	0.004, 0.033
b_3	-0.071	-0.072, 0.003	-0.070, 0.003	-0.071, 0.002	-0.071, 0.006
ab_1	-0.351	-0.350, 0.003	-0.349, 0.004	-0.350, 0.003	-0.348, 0.010
ab_2	-0.135	-0.135, 0.008	-0.133, 0.008	-0.134, 0.006	-0.121, 0.040
aa	-0.034	-0.034, 0.001	-0.034, 0.001	-0.034, 0.001	-0.033, 0.004
aaa	-0.027	-0.027, 8.23e-4	-0.027, 5.57e-4	-0.027, 4.05e-4	-0.026, 0.003
aab	-0.108	-0.107, 0.002	-0.108, 0.003	-0.108, 0.002	-0.103, 0.016
bbb	-0.099	-0.098, 0.005	-0.099, 0.006	-0.099, 0.005	-0.096, 0.013

Table 5.3: Results achieved for the identification of the third-order NARMAX model

Chapter 6

Signal Design and Identification for a Class of Nonlinear Systems

6.1 Introduction

This chapter is directed towards the identification of linear and nonlinear models for certain nonlinear systems. The techniques considered here are thought to be most appropriate when testing systems which are describable by weakly nonlinear models. In this context, low order functional series expansions and block structured systems (single valued) are encompassed. Effectively the linear system description is being augmented with a nonlinear model, rather than completely replaced by it. From the point of view of perturbation signal design, two questions that are of primary importance are:

- Can the accuracy of the linear system description be improved by suitable perturbation signal design ?

- Are there perturbation signal designs which can aid the identification of certain nonlinear models ?

These will be addressed in turn.

The suitability of a periodic excitation when used to identify a linear model for a nonlinear system has been assessed by many researchers in the literature (Godfrey and Briggs, 1972; Buckner and Kerlin, 1972; Godfrey and Moore, 1974; Barker and Davy, 1975; Schoukens, Pintelon, Van der Ouderaa and Renneboog, 1988b; Suki and Lutchen, 1992). This subject is treated in Section 6.2, where both the time- and frequency-domain properties of the excitation are examined. For generality, the discussion does not assume that the Fourier coefficients are used for the parameter estimation, and so the identification procedure used to obtain the system model is not specified.

The identification of a nonlinear model representation is discussed in Section 6.3. To begin with, the requirements of the measurement model are outlined, since they

are significantly different than those required for linear system modelling (Schoukens et al., 1994). In this context, the construction of the measurement channel used in the example of Section 6.3.1 is described. Non-parametric modelling is treated in Section 6.3.2, with attention being limited to the Volterra system description. Some simple identification procedures are described for commonly encountered systems; these are similar to those described in Bendat (1990). The advantages (and disadvantages) of specific designs, that have been described in the literature, are also discussed.

Section 6.3.3 is directed towards the estimation of the parameters of nonlinear differential equation models. Attention is restricted to models that can be written in a form that is linear in the parameters. The identification of a certain class of these models using periodic excitations is discussed in Section 6.3.3, with a realistic experimental example being included to illustrate the method.

The applicability of the continuous-time or discrete-time (Leontaritis and Billings, 1985; Korenberg et al., 1988; Chen et al., 1989) representation must decide the issue over which formulation is more desirable. Furthermore, the decision to process the raw time-domain data, as opposed to the Fourier coefficients, must be made by the user. This is an important choice, since modern continuous-time system identification methods allow systematic modelling errors to be controlled at relatively low sampling frequencies (Sinha and Rao, 1991; Van hamme, 1992). As is the case with linear system identification, time-domain identification methods generally permit the use of an arbitrary excitation signal. Although this is an advantage when external excitation signals are not permitted or impossible to apply to the system, in all other applications, considerable benefits are to be gained from the use of excitation signals which have superior signal-to-noise ratios, and flexible spectral properties. This is discussed further in Chapter 5.

6.2 Linear Modelling of Nonlinear Systems

The frequency-domain response of an R^{th} order polynomial nonlinear system to a periodic excitation with Fourier coefficients A_k, $k \in K$, is given by

$$\begin{aligned} Y(f) = & \sum_{k=-K}^{K} A_k G_1(f_k)\delta(f - f_k) + \sum_{k_1=-K}^{K}\sum_{k_2=-K}^{K} A_k A_{k_1} G_2(f_{k_1}, f_{k_1})\delta(f - f_{k_1} - f_{k_2}) \\ & + \cdots + \sum_{k_1=-K}^{K}\sum_{k_2=-K}^{K}\cdots\sum_{k_R=-K}^{K}\prod_{r=1}^{R} A_{k_r} G_R(f_{k_1}\cdots f_{k_R})\delta(f - f_{k_1}\cdots - f_{k_R}) \end{aligned} \tag{6.1}$$

where $G_r(f_{k_1}, f_{k_2}\cdots f_{k_r})$ is the r^{th} order frequency response of the system. The identification of the linear model, $G_1(.)$, is to be considered here. The identification technique (time/frequency, parametric/non-parametric) is not specified. However identification strategies that require multiple experiments, with signals of different amplitudes (Barker and Davy, 1975; Schoukens, Rolain, Montecelli and De Locht, 1993b), are not considered. The accuracy of the resulting linear system estimate, in the presence of systematic nonlinear effects, is predominantly affected by both the time- and frequency-domain properties of the signal. These are considered separately.

6.2.1 Time-Domain Properties

For the general case, with an excitation signal composed of a pre-determined power spectrum, the only free variables are the harmonic phases, ϕ_k. In practical situations, the presence of strong nonlinear effects (such as limiters) enforces the use of excitation signals with low peak amplitudes and so a low crest factor is essential. However, we wish to examine the benefits of using periodic signals with low crest factors in the presence of weak nonlinear effects. To this end, concentration is focussed upon the effect of the relationship $y = N(x)$ where $N(.)$ is either a

square-law or cubic function. The effect of a square-law characteristic is zero if the excitation is symmetrically distributed (Bendat, 1990). In the general case (see also Section 6.2.2), the excitation signal can be regarded as approximately symmetrically distributed, and so the identification error is expected to be small. This point is discussed further by Schoukens and Pintelon (1991), with measurement examples used to compare the distortions introduced by a nonlinear function on results obtained using both random and periodic excitations. The periodic signals possess much greater immunity to the effect of the nonlinearity.

Considering cubic distortions, Evans, Rees and Jones (1994b) performed simulations and found that when the multisine phases are normally distributed, the error in the linear frequency response estimate is proportional to the square of the crest factor. Hence for Gaussian-like signals, the benefits of a low CF are clear *when the nonlinearity operates upon this signal.* This is of interest if the nonlinearity operates upon a signal which is the result of passing the excitation (which more than likely has a non-Gaussian PDF) through a linear system. Unfortunately, the time-domain properties of an excitation with a low crest factor cannot even be crudely approximated by a Gaussian model, and so this square-law result will not apply to the input nonlinearity case, although it is interesting to note the use of Gaussian-like multisines in Ditmar and Pettitt (1993). However the existence of the square-law relationship for multisines with randomised phases, makes it clear that in general a low crest factor is desirable in order to reduce nonlinear distortions. If linear dynamics separate the input signal and the nonlinear element, then approximate knowledge of the linear dynamics may be used to minimise the crest factor of the relevant signal(s) (Van der Ouderaa, Schoukens and Renneboog, 1988b; Guillaume et al., 1991).

6.2.2 Frequency-Domain Properties

Here we will assume that we have the flexibility to designate an arbitrary excitation spectrum, specified by A_k, $k \in K$, and wish to know if certain choices of K offer potential benefits. A well known choice for K is a set containing no even entries, thus eliminating the effect of an even-order kernel, $G_r(.)$ r even, upon the estimate of $G_1(.)$. Applications of this idea in the literature are numerous (Godfrey and Moore, 1974; Kollár, Pintelon and Schoukens, 1994; Evans, Rees, Jones and Hill, 1994a). Similar statements involving $G_r(.)$ r odd > 1 are difficult since it is impossible to specify a frequency set, K, which causes $G_1(.)$ to be unaffected by other odd-order kernels. This is clear from eqn. (6.1), by realizing that $k \pm k_2 \pm k_3$ can always be made equal to k. Given that there are in the order of L_K^r contributions in the output signal due to an r^{th} order kernel, where L_K is the number of harmonics in K, one might try to reduce the bias by reducing the number of contributions which fall at each frequency. This does not, however, actually reduce the magnitude of the distortion (bias).

The use of a multiharmonic signal with the harmonic set chosen such that the excitation harmonics are affected by the minimum number of contributions is analogous to using a sinewave excitation. Here the energy due to the nonlinearity at the excitation frequencies will have the same phase as the original energy at the excitation frequencies (Evans et al., 1994b; Schoukens and Pintelon, 1994a) (due to $f_k + f_{k_1} - f_{k_1}$). On the other hand, if an excitation with a broadband spectrum is used, then the contribution at a given harmonic will contain an ensemble of contributions each with an independent phase value, thus giving the overall effect of a distortion which is approximately random across the excitation harmonics. In this case, a signal containing a large number of interacting harmonics is expected to produce the best results.

To assess the merits of periodic excitations for linear system identification in the

presence of nonlinearities, attention has been restricted in this research to the influence of second- and third-order nonlinear effects. For these instances, the benefits of a low crest factor have been discussed as well as the advantages of being able to stipulate specific harmonic sets. It has been shown that that a low crest factor is important for good accuracy of the linear model. This may seem obvious since the usual reason for crest factor minimisation is owing to the presence of nonlinear input transducers. Hence once it is realized that the crest factor is an important variable (to maximise input power for a given input range), and care is taken to reduce it, the effect of the crest factor in this range may not be great. However this assumes that low crest factor signals are used in the first place, which is not the case when utilising random excitations. Regarding the selection of the power spectrum, the well known methodology of using inverse-repeat or anti-symmetric signals occurs as the only general principle. Note that the selection of specific harmonic sets for the purpose of higher-order frequency response measurements will be treated further in Section 6.3.2. Using periodic excitations comes with the added advantage of allowing easy selection of the relevant frequencies in the response signal. This is automatically carried out when the Fourier coefficients are used for the identification but can just as easily be used for time-domain methods (Ljung, 1986).

6.3 Identification of Nonlinear Dynamics

This section deals with the non-parametric estimation of the the first three frequency response functions in eqn. (6.1) along with the non-parametric and parametric estimation of a certain class of continuous-time nonlinear systems. The identification will be carried out utilising the spectra of the input and output signals. The parametric models considered in Section 6.3.3, are the class of systems, considered by Tsang and Billings (1994), which enable the model equations to be written in the

frequency-domain, involving the spectra of the input and output signals and their higher powers. The transferral to the frequency-domain enables estimation of the physical parameters of the system, without having to reconstruct the derivatives of the input and output signals, which is necessary with the identification method of Tsang and Billings (1994). To begin with, the requirements of the measurement model and the steps taken to realize it are described.

6.3.1 Measurement Model

With linear system modelling the use of relatively calibrated anti-alias filters is sufficient to realize the bandwidth-limited data. This unfortunately is not sufficient when one needs to reconstruct a nonlinear model from the measurements, since generally

$$AA_2(\jmath\omega)Y(\jmath\omega) = F(AA_1(\jmath\omega)U(\jmath\omega)) \neq AA_1(\jmath\omega)F(U(\jmath\omega)) \tag{6.2}$$

where $F(.)$ is a nonlinear function and $AA_1(.)$, $AA_2(.)$ respectively represent the characteristics of the input and output anti-alias filters. To model the nonlinear system, $AA_1(.)$ (and consequently $AA_2(.)$) must commute with $F(.)$, and so equalisation of the anti-alias filters will generally be necessary. This, though time consuming, may be carried out with the methods developed by Pintelon, Rolain, Vanden Bossche and Schoukens (1990), Kollár, Pintelon, Rolain and Schoukens (1991) and Schoukens and Pintelon (1991), with the complete design possible using the identification techniques implemented by Kollár (1994), for an arbitrary filter characteristic. With periodic excitations it is sufficient to model the frequency characteristic of the anti-alias filter(s), and to apply a simple correction procedure off-line. The measurements obtained in this work were obtained using an Advantest R9211C servo analyser (Advantest Corporation, 1989), connected to a PC using a

GPIB interface. The analyser has anti-alias protection on both channels (which may be switched off), but the frequency response of these filters is not specified, and is not measurable to a high accuracy given the configuration of the hardware. For this reason, external anti-alias filters were used (eight-order Barr and Stroud), with the Advantest filters switched off. The bandwidth of interest will be in the range 0-≈ 200 Hz (see Section 6.3.3) and so the cut-off frequency of the filter is set to 200 Hz. The results of equalising the output channel, in this case, are shown in Figure 6.1. It is sufficient for the compensated characteristic to have a uniform amplitude

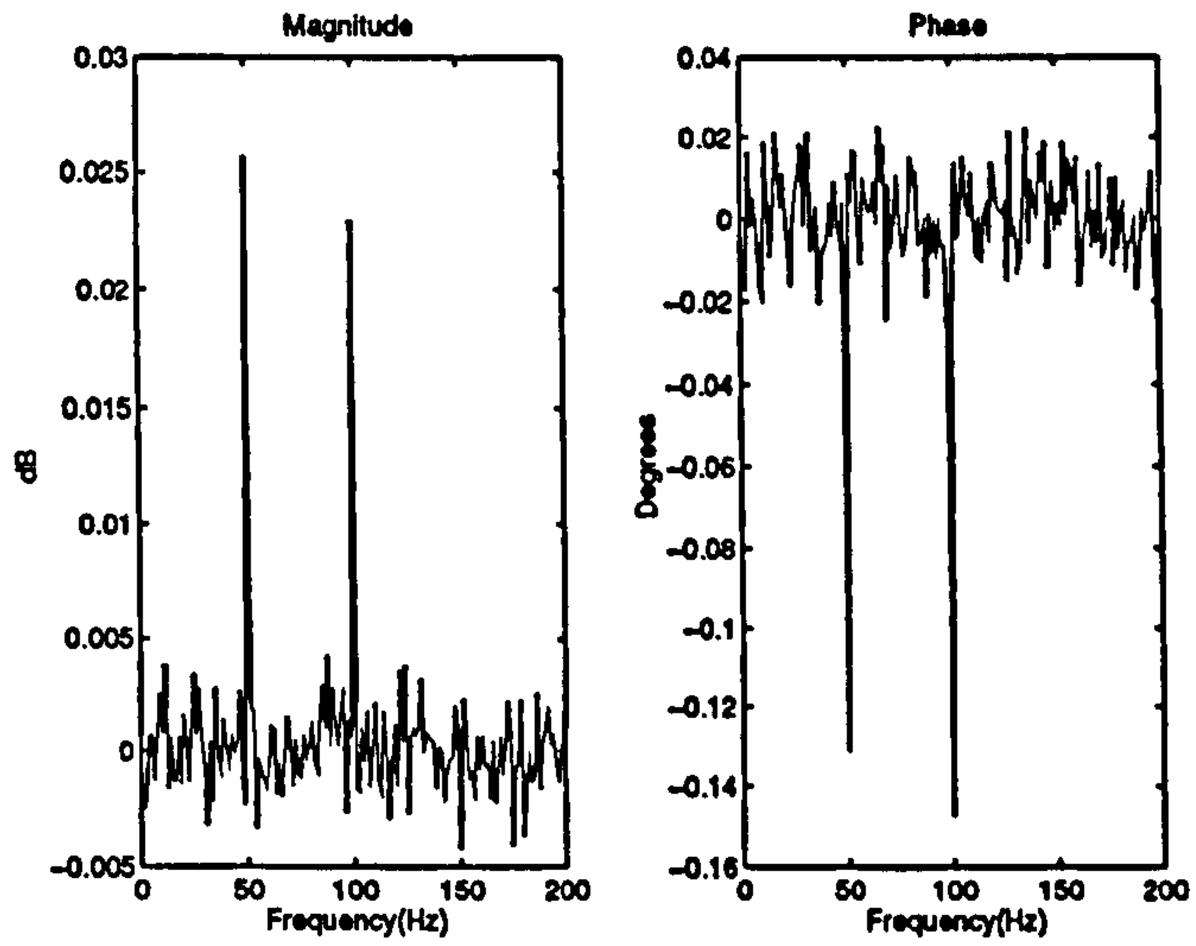

Figure 6.1: Equalised frequency response of anti-alias filter

characteristic and linear phase, but because the compensation is being carried out off-line, with periodic signals, we can obtain the characteristic shown in Figure 6.1. Note that only output noise will be considered here, and hence the measurement system requires anti-alias protection on the output channel only.

6.3.2 Higher-Order Frequency Response Measurement

The direct measurement of the higher-order frequency response functions, $G_r(.)$, $r > 1$, using periodic signals is now treated. The use of sparse-harmonic sets for this purpose has received attention in the literature (Victor and Shapely, 1980; Lawrence, 1981; Boyd, Tang and Chua, 1983), with attention usually being restricted to the second- and third-order functions, $G_2(f_{k_1}, f_{k_2})$ and $G_3(f_{k_1}, f_{k_2}, f_{k_3})$, respectively. Throughout the present discussion, it is assumed that the system to be identified is represented by R frequency kernels and not just the R^{th} frequency kernel. For R equal to three, the nonlinear system is shown in Figure 6.2. To begin with, the case of $R = 2$ is considered.

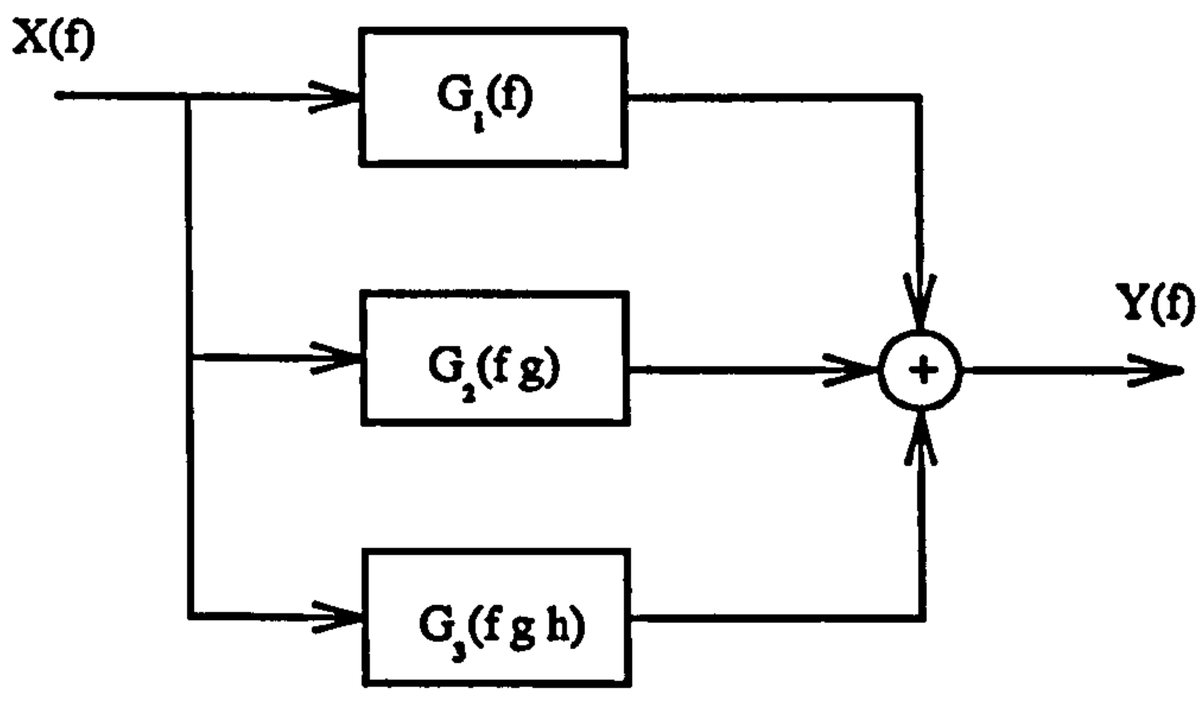

Figure 6.2: The general third-order nonlinear model

6.3.2.1 First- and Second-Order Frequency Response Measurement

The system is assumed to be represented by the input-output relationship shown in Figure 6.2 with $G_3(.)$ equal to zero. To obtain correct measurements of $G_1(.)$ and $G_2(.)$, it is clear that the $Y(f_{k_1} + f_{k_2})$ must be uncorrelated with $Y(f_k)$ and also that $Y(f_{k_1} + f_{k_2})$ results from a unique combination $(\pm f_{k_1}, \pm f_{k_2})$. If this is the case, the frequency response functions can be measured using (noting that $X(-f_{k_a}) =$

$X^*(f_{k_a})$)

$$G_1(f_k) = \frac{Y(f_k)}{X(f_k)} \qquad \forall k \in K \tag{6.3}$$

and

$$G_2(f_{k_1}, f_{k_2}) = \frac{Y(f_{k_1} + f_{k_2})}{Sc_2 X(f_{k_1}).X(f_{k_2})} \quad \forall k_1, k_2 \in K \quad Sc_2 = \begin{cases} 1 & k_1 = k_2 \\ 2 & \text{otherwise.} \end{cases} \tag{6.4}$$

A harmonic set which satisfies the above condition, exactly, is given in eqn. (6.5), which, when used to excite a system that can be described by a second-order Volterra system, will allow 400 measurements of $G_2(f_2, f_2)$ to be made

$$K = 11, 13, 17, 25, 57, 73, 137, 189, 209, 297, 375, 441, 533,$$
$$629, 797, 935, 1075, 1177, 1287, 1287, 1517. \tag{6.5}$$

The ratio of the lowest harmonic to the highest harmonic is approximately 150 for this signal, and so, in theory, this signal could be used to to provide a useful estimate of the system dynamics. The price to be paid is the necessity to have the lowest harmonic in the set equal to eleven, which increases the signal period by at least the same factor. If the first harmonic is specified in the twenty harmonic set, the equivalent ratio of the highest to lowest harmonic is 1649, which is thought to be too great for most practical applications.

Given that the use of the harmonic set given in eqn. (6.5) allows 400 estimates of the second-order frequency response function to be made, at the expense of a long experimental time, it seems logical to attempt a compromise solution. The harmonic set (number of harmonics, N_K, equal to 20), given in eqn. (6.6) is one such solution:

$$K = 5, 9, 17, 29, 45, 65, 89, 117, 149, 185, 225, 269$$

$$317, 369, 425, 485, 549, 617, 689, 765 \tag{6.6}$$

drawn from the recursion

$$K(i+1) = K(i) + 4.(i) \tag{6.7}$$

with $i = 1, \cdots, N_K$ and $K(1)$ equal to 5. When this is compared to a harmonic set, specified to produce no interaction between the harmonic frequencies (i.e. the counterpart of eqn. (6.5)):

$$K = 5, 7, 11, 31, 39, 79, 131, 161, 219, 275, 335, 439$$
$$551, 645, 699, 797, 1085, 1271, 1461, 1537 \tag{6.8}$$

it is seen that the use of the harmonics in eqn. (6.6) allows an approximate halving of the signal period. The compromise in using a harmonic set, calculated using eqn. (6.7), is the reduction in the number of estimates of $G_2(.)$ which can be made: for eqn. (6.6) the number of possible measurements is 289.

To illustrate a second-order frequency response measurement, Figure 6.3 shows the magnitude of the calculated $G_2(.)$ for a Wiener model consisting of a first-order linear system (time constant 10ms, unity gain) and a square-law nonlinearity, when excited with a signal containing the harmonics of eqn. (6.6), with a signal period 7.65 seconds. Figure 6.4 shows the corresponding estimate obtained when the system is excited with the harmonic set of eqn. (6.8), with a period of 15.37 seconds now being required. The utilised energy, up to the excitation signal bandwidth, for the two responses is shown in Figure 6.5. This illustrates that although significant energy is being discarded from the response to the "faster" excitation, good coverage of the dynamic range is being achieved.

The use of sparse harmonic sets to measure the first- and second-order frequency

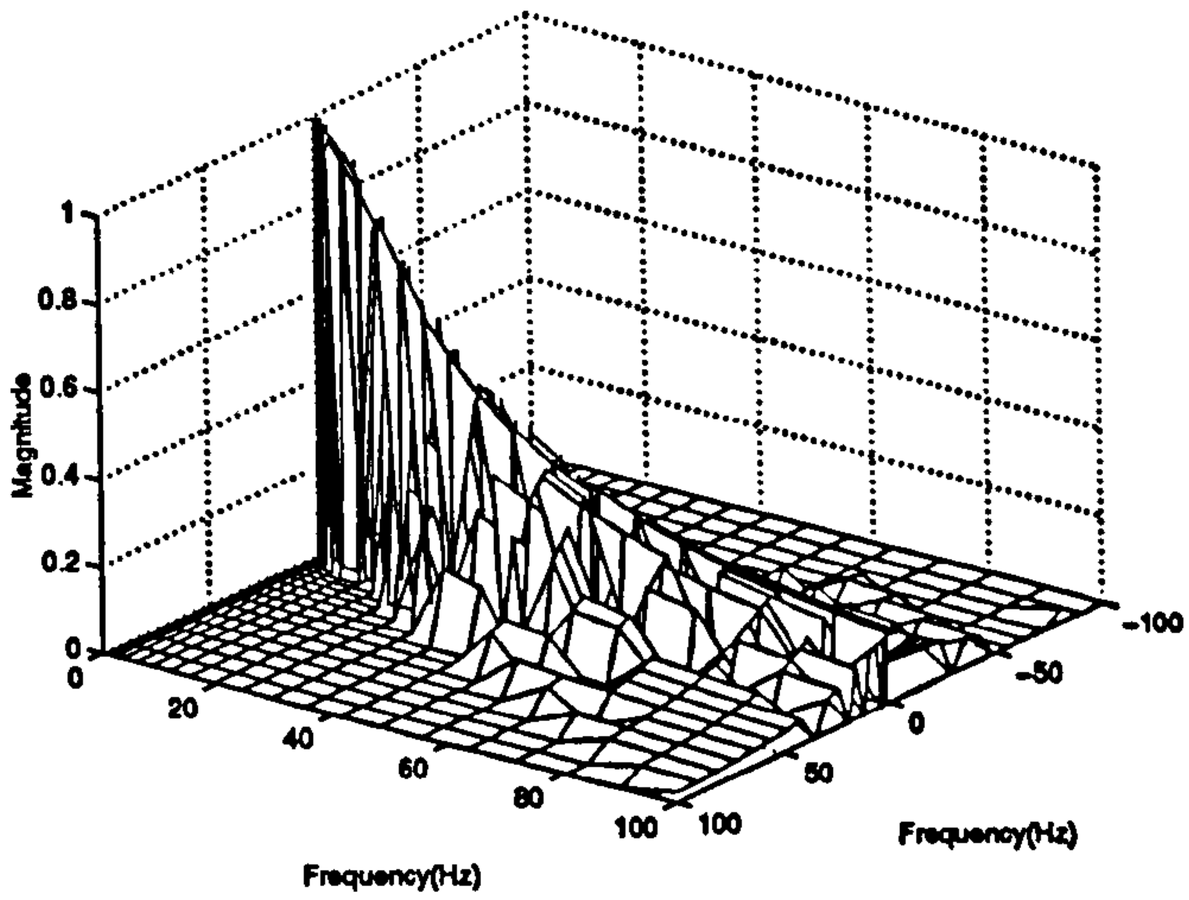

Figure 6.3: Second-order frequency response magnitude obtained using signal based on the harmonic set obtained using the recurrence formula

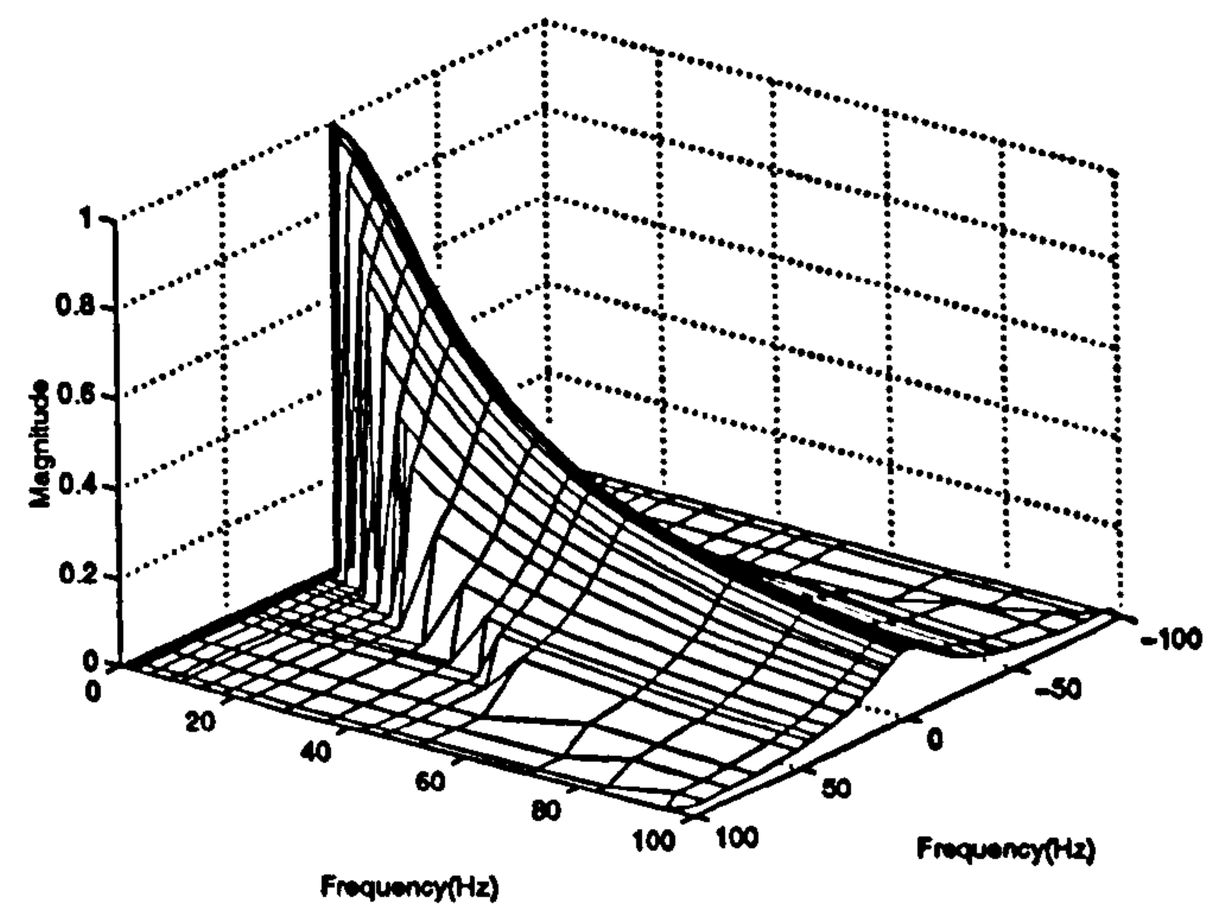

Figure 6.4: Second-order frequency response magnitude obtained using signal based on harmonic set designed to yield a maximum number of $G_2(.)$ measurements

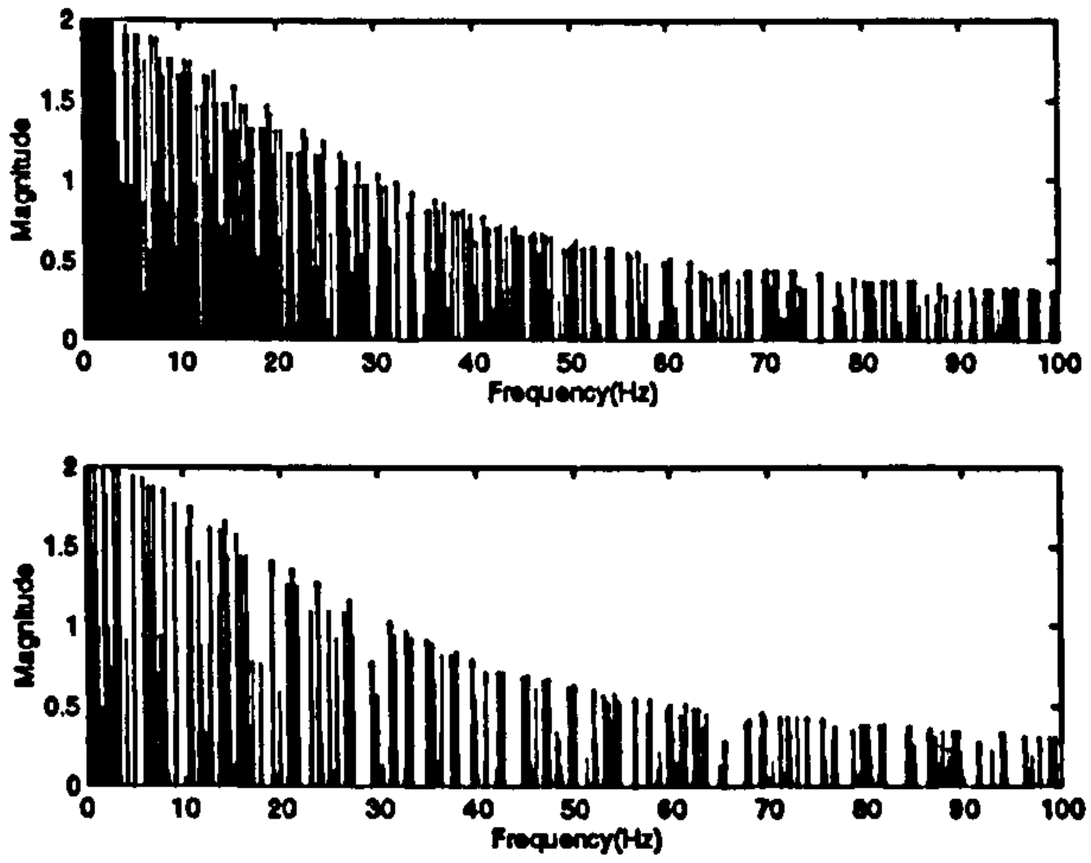

Figure 6.5: Upper: spectrum of output signal using input harmonic set with each harmonic yielding a second order kernel estimate. Lower: equivalent using input harmonic set which yields 289 estimates of $G_2(.)$. Graphs show utilised harmonics only.

response functions is therefore most suited to systems which exhibit relatively fast dynamics. If the time-constants of the system approach the order of seconds, the signal period quickly approaches large values (hours). However, for systems which exhibit fast responses, the method is feasible. The simultaneous measurement of the third-order frequency response function is discussed in Section 6.3.2.2.

6.3.2.2 Third-Order Frequency Response Measurement

The measurement of the third-order frequency response function is included here for completeness. An immediate problem is that an independent measurement of $G_1(f_k)$ cannot be made, as discussed in Section 6.2. Along the same lines as Section 6.3.2.1, $G_3(f_{k_1}, f_{k_2}, f_{k_3})$ can be measured directly if the measured response $Y(f_m)$ at the frequency f_m results from a unique combination of any three of the excitation

frequencies $(\pm f_{k_1}, \pm f_{k_2}, \pm f_{k_3})$. If this is the case, $G_3(.)$ can then be calculated using

$$G_3(f_{k_1}, f_{k_2}, f_{k_3}) = \frac{Y(f_{k_1} + f_{k_2} + f_{k_2})}{Sc_3.X(f_{k_1}).X(f_{k_2}).X(f_{k_2})} \qquad \forall k_1, k_2, k_3 \in K$$

$$Sc_3 = \begin{cases} 1 & k_1 = k_2 = k_3 \\ 6 & k_1 \neq k_2 \neq k_3 \\ 3 & \text{otherwise.} \end{cases} \tag{6.9}$$

The task now is to find suitable harmonic sets which allow direct measurement of $G_3(.)$. The harmonic sets which allow the maximum number of points on $G_2(.)$ to be made, such as those given in eqn. (6.5) and eqn. (6.8), were examined initially to see if they offer any advantages when it comes to the measurement of $G_3(.)$. Considering a ten harmonic set, with the same properties as that given in eqn. (6.5), although it allows 79 measurements of $G_3(.)$ to be made, not one of these measurements falls within the bandwidth of the excitation. For the compromise signal, calculated using the recursion of eqn. (6.7), the situation is worse (35 measurements of $G_3(.)$, all outside the excitation bandwidth). Hence different harmonic sets are necessary.

One approach that is fruitful (but not necessarily practical) is to specify only those harmonics which do not contribute to the third-order harmonic frequencies, i.e. for an excitation $X(k)$,

$$X(k_1, k_2, \cdots, k_R) = \begin{cases} 0, & k_1 \pm k_2 \pm \cdots \pm k_R = R.k \\ A_k, & k_1 \pm k_2 \pm \cdots \pm k_R \neq R.k \end{cases} \quad \forall k \in K. \tag{6.10}$$

A set of frequencies displaying this property is

$$K = 1, 7, 25, 31, 97, 103, 121, 127, 385, 391. \tag{6.11}$$

Using an excitation containing these frequencies allows 61 measurements of $G_3(.)$ to be made within the excitation bandwidth (100 measurements in total), along with 66 measurements of $G_2(.)$. The use of signals with $\min\{K\} >> 1$ is again beneficial for practical coverage of the dynamic range of the system, but results in long signal periods. The use of a signal containing the harmonics given in eqn. (6.11) or a set specified along similar lines would only be justified if the experimental time was not excessive and no structural knowledge was available. In these circumstances they are an alternative to the random excitations considered by Bendat (1990).

This discussion of frequency-domain non-parametric higher-order frequency response estimation, shows that the first three kernels can be estimated using simple analysis procedures. However the benefits of direct measurement comes with the potentially considerable disadvantage of employing signal periods which are several orders of magnitude greater than the dominant time-constants of the system. This point has already been discussed by Lawrence (1981). However if this long signal period is not unworkable e.g. for systems with fast dynamics, employing specially designed signals does simplify the identification procedure. This simplification is sometimes useful in biological applications (Victor and Shapely, 1980) where parametric models are often difficult to postulate. However for many practical applications, this direct measurement is not feasible due to the long experimental time necessary. For this reason, the identification of certain commonly encountered parametric models is considered in the following sections.

6.3.2.3 Hammerstein and Wiener Models

The attention that has been given to Hammerstein and Wiener model structures (Billings and Fakhouri, 1977; Billings and Fakhouri, 1979; Greblicki and Pawlak, 1986; Hunter and Korenberg, 1986) illustrate the common appearance of these models in practical applications. They are included here to illustrate the advantages of

using periodic signals with specified spectra, and also the limitations of using such ideas.

The higher-order frequency response functions of the Hammerstein, $G_{Hr}(.)$, and Wiener, $G_{Wr}(.)$, models are, respectively,

$$G_{Hr}(f_{k_1}, f_{k_2} \cdots f_{k_r}) = a_r G_1(f_{k_1} + f_{k_2} \cdots + f_{k_r}) \tag{6.12}$$

and,

$$G_{Wr}(f_{k_1}, f_{k_2} \cdots f_{k_r}) = a_r G_1(f_{k_1}).G_1(f_{k_2}) \cdots G_1(f_{k_r}) \tag{6.13}$$

and are thus special cases of the general case-1 and case-2 nonlinear models considered by Bendat (1990), and depicted in Figure 6.6. For the model of Figure 6.6(a),

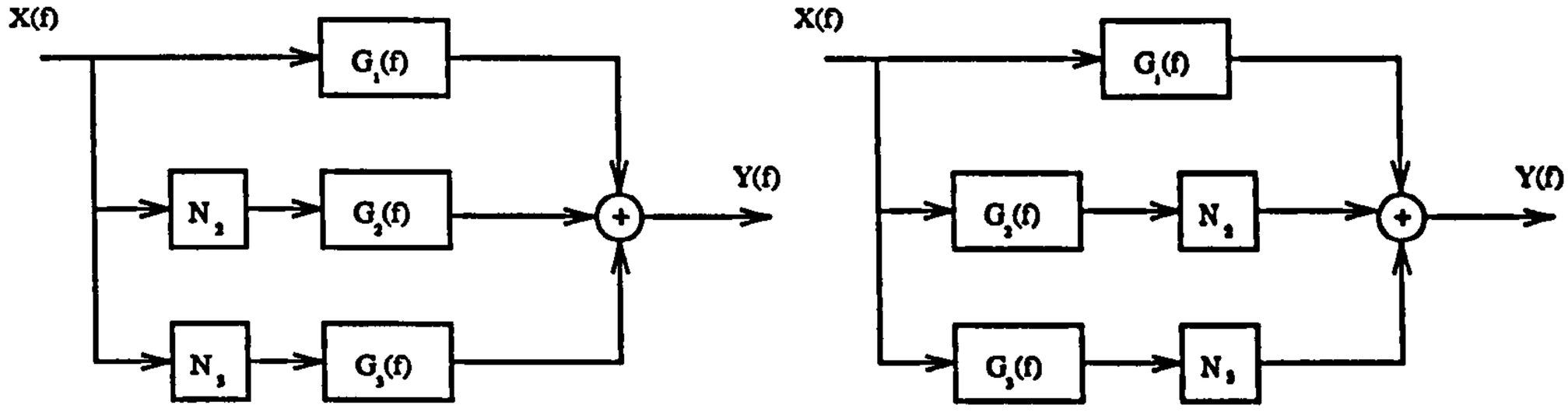

Figure 6.6: (a): General nonlinear input model. (b): General nonlinear output model

the use of non-parametric multiple-input single-output identification procedures, with uncorrelated signals, has been reported for random excitations (Bendat, 1976), and periodic excitations (Barker and Al Hilal, 1985; McCormack et al., 1994b). For the periodic case, the use of an inverse-repeat signal with certain odd-harmonics of the fundamental set equal to zero, allows the identification of the linear and nonlinear subsystems. Note that the Hammerstein model is considered again in Section 6.3.3.

With the Wiener model, when R equals 2, the use of an inverse-repeat signal

allows the direct measurement of the linear and nonlinear characteristics. However when $R = 3$, the correlation at $Y(f_k)$, between the linear and cubic contributions, allows no such procedures to be used. At this point iterative estimation procedures are needed which may be either in the time-domain (Billings and Fakhouri, 1977; Hunter and Korenberg, 1986) or in the frequency-domain. A frequency-domain method similar to that used by Hunter and Korenberg (1986) can be written as follows, with initial conditions calculated according to Billings and Fakhouri (1977),

1. Calculate

$$y^i(r) = \mathcal{F}^{-1}(\hat{G}^{(i)}(j\omega).X(j\omega)) \tag{6.14}$$

where $\hat{G}^{(i)}(j\omega)$ is the FRF of the estimated model $\hat{G}^{(i)}(s)$.

2. Calculate the energy expected in the frequency set $\{H\}$.

$$\hat{Y}^{(i)}(h) = \sum_{n=0}^{N-1} (y^{(i)}(r))^3 \exp(-j\omega hn/N) \tag{6.15}$$

3. Estimate the corresponding value for $\hat{c}_3^{(i)}$

$$\hat{c}_3^{(i)} = \text{sign}(\hat{c}_3^{(0)}).\frac{1}{N_H}\sum_h \frac{|Y(h)|}{|\hat{Y}^{(i)}(h)|} \tag{6.16}$$

where $\hat{c}_3^{(0)}$ is the initial estimate of c_3, and $Y(h)$ is the measured Fourier coefficients in H for which a high coherence is estimated (see Chapter 1 and McCormack et al. (1995)).

4. Compensate the linear output signal

$$y^{(i+1)}(r) = y_k(r) - (\hat{c}_3^{(i)}.(y_k^{(i)}(r))^3) \tag{6.17}$$

where, $y_k(r)$ is the measured output signal with all non-excitation frequencies

removed.

5. Estimate a new model $\hat{G}^{(i+1)}(s)$.

6. Iterate a set number of times, or until $\|\hat{c}_3^{(i+1)} - \hat{c}_3^{(i)}\| < \varepsilon$.

After calculation of $\hat{c}_3$, the corrected linear output can be used to calculate $\hat{c}_2$.

The accuracy of the above method, and the non-parametric Hammerstein algorithm is illustrated using the example in Section 6.3.2.4.

6.3.2.4 Hammerstein and Wiener Identification Example

The identification methods for the Hammerstein and Wiener models are tested on a system whose linear element is given by,

$$G(s) = \frac{b_0}{a_3 s^3 + a_2 s^2 + a_1 s + a_0} \tag{6.18}$$

with the parameter vector, $P_l = [b_0, a_1, a_2, a_3] = [1, 3, 3, 1]$ (note that a_0 has been fixed to unity). For both models, the static nonlinear element consists firstly of a third-order polynomial and secondly a saturation element, where in both cases the nonlinear element has parameters, $P_{nl} = [c_2, c_3]$. The excitation signal is a five-level signal, (see Chapter 1 and McCormack et al. (1995)), specified to have equal power in harmonics,

$$K = 1, 5, 7, 11, 13, 17, 19, 23, 25, 29, 31, 35, 37, 41 \tag{6.19}$$

and with all multiples of harmonics two and three suppressed. The mean and standard deviation, $[\hat{(.)}, \hat{\sigma}]$, of the parameters are calculated from 100 sets of data, each containing five periods of the response to the excitation, excluding transients. The input signal amplitude was normalised to unity, and the variances of the output Fourier coefficients were obtained from an analysis of the individual periods. The

System	*Noise*	*Estimate,* $[(.), \hat{\sigma}]$					
		Linear				Nonlinear	
		b_0	a_1	a_2	a_3	c_2	c_3
Hammerstein Polynomial	True Value	1	3	3	1	2,5	3.5
	Noise Free	1	3	3	1	2.5	3.5
	$\sigma_y^2 = 0.1$	[1.02,0.02]	[3.0,0.01]	[3.0,0.01]	[1.0,0.03]	[2.51,0.13]	[3.51,0.18]
Hammerstein Saturation	True Value	1	3	3	1	0	-0.58
	Noise Free	1.08	2.95	2.95	2.96	0	-0.31
	$\sigma_y^2 = 0.1$	[1.11,0.1]	[2.99,0.08]	[3.01,0.06]	[0.98,0.14]	0	[-0.32,0.05]
Wiener Polynomial	True Value	1	3	3	1	1.5	2
	Noise Free	1	3	3	1	1.5	2
	$\sigma_y^2 = 0.1$	[0.99,0.03]	[2.99,0.03]	[3.00,0.03]	[0.99,0.04]	[1.52,0.1]	[2.08,0.30]
Wiener Saturation	True Value	1	3	3	1	0	-0.9492
	Noise Free	1.12	3.07	2.97	1.04	0	-0.9492
	$\sigma_y^2 = 0.1$	[1.12,0.03]	[3.07,0.08]	[2.97,0.06]	[1.06,0.09]	0	[-0.95,0.02]

Table 6.1: Results obtained from the identification of the Hammerstein and Wiener models

results are summarised in Table 6.1. For the noisy cases, typical coherence functions obtained for the Hammerstein (polynomial) and the Wiener (saturation) trials are shown in Figures 6.7 and 6.8 respectively.

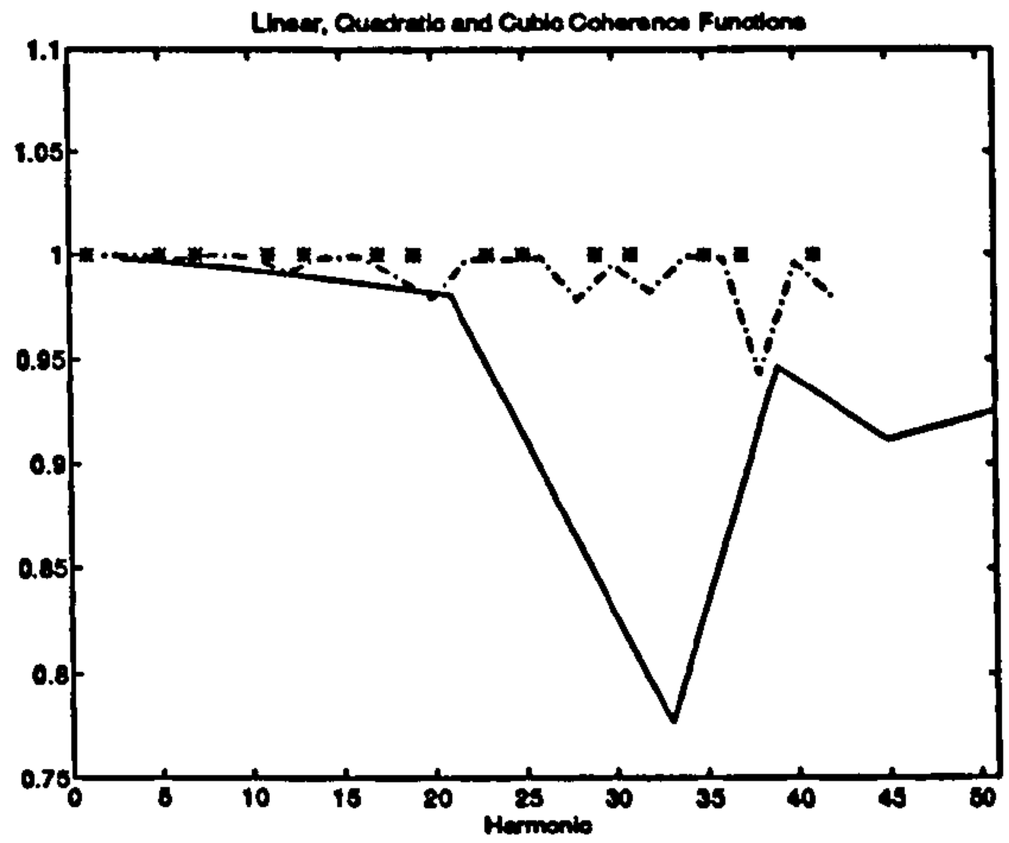

Figure 6.7: Coherence functions estimated for the Hammerstein model: linear (*), quadratic (-) and cubic (-.)

The results shown in Table 6.1 illustrate that the use of non-parametric-like methods can aid the identification of the Hammerstein model, and give satisfactory results. For the Wiener model, the results obtained also show good accuracy. How-

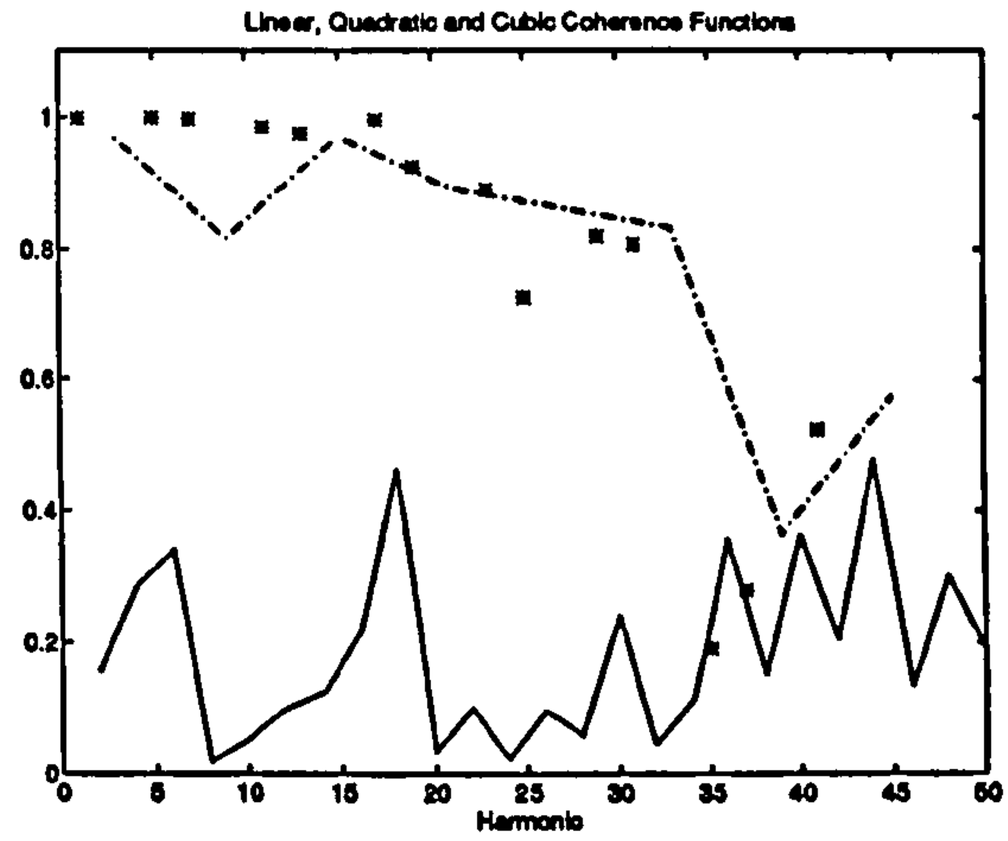

Figure 6.8: Coherence functions estimated for the Wiener model: linear (*), quadratic (-) and cubic (-.)

ever for the Hammerstein model the identification method is not easily extendible beyond a third-order polynomial approximation of the nonlinearity. This is due to the orthogonality conditions between the input signal and its higher powers being inherently limited. Similarly for the Wiener model, the use of the Fourier coefficients for the identification of the nonlinear characteristic is limited to a single monomial coefficient.

6.3.3 Identification for a Class of Nonlinear Systems

The nonlinear models examined in this section are a subset of the class of systems describable by a model of the form,

$$\sum_{i=0}^{n} a_i \frac{d^i y(t)}{dt^i} + F^R[y^n(t), \cdots, y(t), x^m(t), \cdots, x(t)] = \sum_{i=0}^{m} b_i \frac{d^i x(t)}{dt^i} \qquad (6.20)$$

where, F^R is an R^{th}-order polynomial nonlinear function of the input and output signals, and their higher-order derivatives. This model has been treated by Patra (1993) and Tsang and Billings (1994). Due to the linear in the parameters nature of

eqn. (6.20), the orthogonalised linear regression techniques, which were developed for discrete-time models (Korenberg et al., 1988; Chen et al., 1989), are used by Tsang and Billings (1994) for the parameter estimation and structure detection. The decision to identify the parameters of the continuous-time system as opposed to the discrete-time counterpart (Leontaritis and Billings, 1985), is based on the need for physical interpretation of the parameters.

Although the time-domain approach developed by Tsang and Billings (1994) is applicable to the whole class of systems described by eqn. (6.20), the high over-sampling needed to numerically approximate the integrating filters (a factor of 400 is evident in the results section of Tsang and Billings (1994)), may prevent the application of the approach in many real-life situations. The approach adopted here utilises the Fourier coefficients of the input and output signals, their higher powers, and if necessary, their cross products. Therefore the following model is appropriate,

$$A(\jmath\omega)Y(\jmath\omega) + \sum_{r=2}^{R} G_r(\jmath\omega)Y^r(\jmath\omega) = B(\jmath\omega)X(\jmath\omega) + \sum_{r=2}^{R} G_r(\jmath\omega)X^r(\jmath\omega) + \sum_{r=1}^{R} G_r(\jmath\omega)Z^r(\jmath\omega) \tag{6.21}$$

where $X^r(\jmath\omega)$, $Y^r(\jmath\omega)$ and $Z^r(\jmath\omega)$ correspond to the spectra of $x^r(t)$, $y^r(t)$ and $(x(t)y(t))^r$ respectively. The main advantage of using the Fourier coefficients, within the model equation of eqn. (6.21), is the applicability of the method at relatively low sampling intervals (theoretically no over-sampling is necessary), as well as the exact relationship in the frequency-domain when periodic signals are used with synchronised sampling.

Systems which may be modelled by the general model of eqn. (6.21) include the

Hammerstein model, written in frequency-domain transfer function representation,

$$Y(j\omega) = \frac{B(j\omega)}{A(j\omega)}(X(j\omega) + \sum_{r=2}^{R} h_r X^r(j\omega)), \tag{6.22}$$

the nonlinear feedback system, written in frequency-domain transfer function representation,

$$Y(j\omega) = \frac{B(j\omega)}{A(j\omega)}(X(j\omega) - \sum_{r=2}^{R} h_r Y^r(j\omega)) \tag{6.23}$$

and the nonlinear damping system

$$\sum_{i=0}^{n} a_i \frac{d^i y(t)}{dt^i} + d_3 y^2(t) \frac{dy(t)}{dt} = \sum_{i=0}^{m} b_i \frac{d^i u(t)}{dt^i}, \tag{6.24}$$

with frequency-domain transfer function

$$Y(j\omega) = \frac{1}{A(j\omega)}(B(j\omega)X(j\omega) - D(j\omega)Y^3(j\omega)), \tag{6.25}$$

where h_r, f_r and d_3 are scalar coefficients and $D(j\omega)$ is a differentiator.

The frequency-domain transfer functions are nonlinear in the parameters, and so the linearised solution of the estimation problem (similar to using the method of Levi (1959) for linear systems) is expected to produce poor results in situations containing even a moderate amount of noise. Iterative schemes that have been developed to produce improved estimates of the transfer function coefficients of linear systems (Sanathan and Koerner, 1963; Lawrence and Rogers, 1979) could, in principle, be modified accordingly, and used with the problem at hand. However with easy access to powerful numerical minimisation techniques (Grace, 1992), the original least-squares cost function can be minimised. This is by no means the best way to obtain parameter estimates for the model equations, but it is expected to at least generate usable results. The start values for the optimisation can be obtained

from the linearised cost function.

6.3.4 Nonlinear Feedback Example

To illustrate the identification technique, the system shown in Figure 6.9 is used. The saturation element was used in favour of polynomial operators owing to the latter not being available. However its use does give an indication of the accuracy of the identification method, and the performance in the presence of modelling errors. The input signal is a multisine with frequencies $(1, 5, 7, 11 \cdots 53) * 500/400$ Hz, i.e.

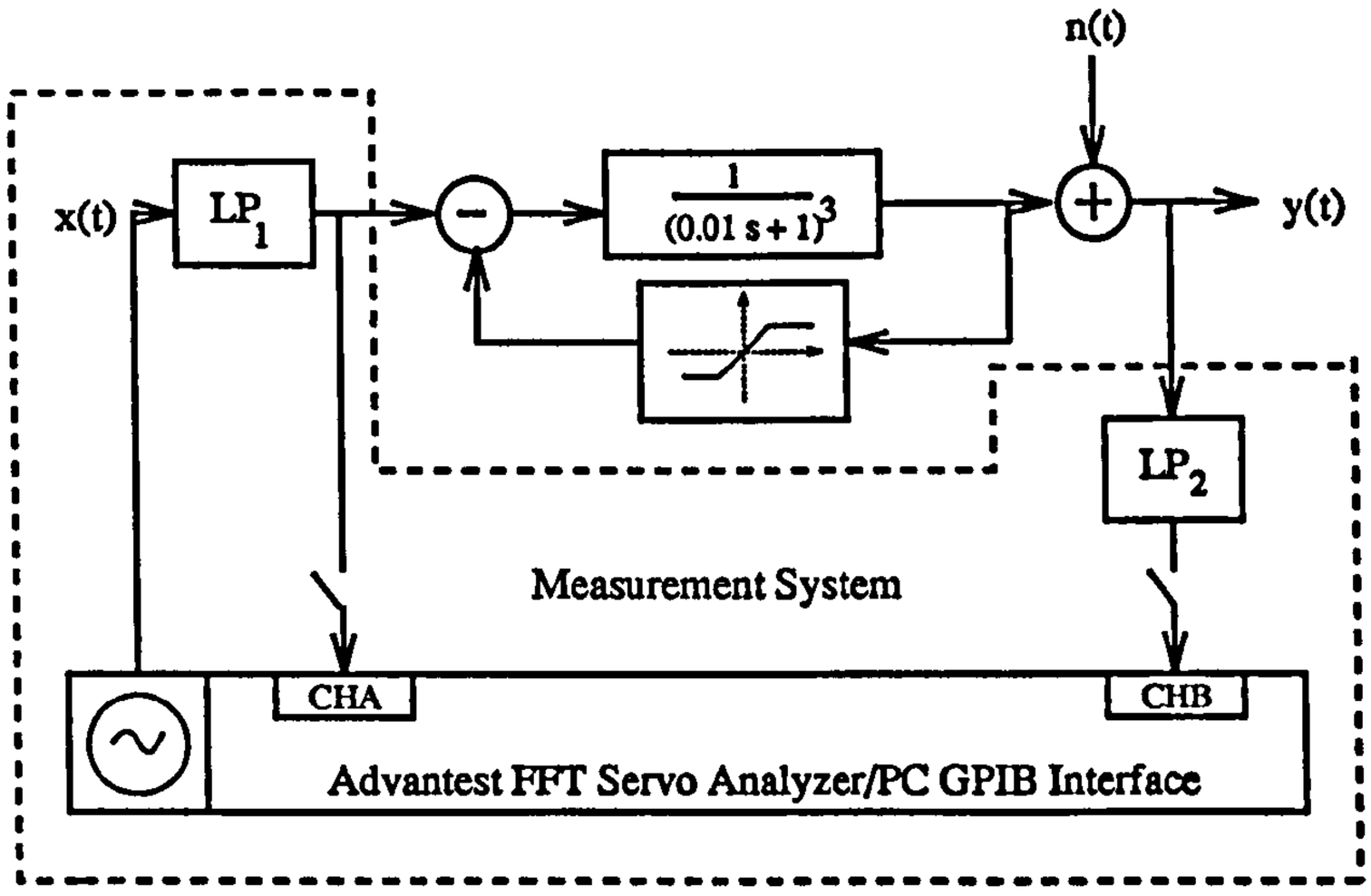

Figure 6.9: Nonlinear system under investigation with measurement arrangement shown.

all multiples of harmonics two and three are suppressed from the input spectrum. In Figure 6.9, LP_1 represents an eighth order reconstruction filter (cut-off frequency: 60Hz) to ensure spectral purity of the excitation, and LP_2 is the (equalised) anti-alias filter whose characteristic is shown in Figure 6.1. The noise source, $n(t)$, was obtained from a Hewlett-Packard Gaussian noise generator, and was specified to have 0.2 V_{rms} in the frequency range up to 500 Hz. To obtain an indication of the

accuracy of the results, 19 periods of the steady-state response, $y(t)$, were measured. One period of the time series and the corresponding amplitude spectra are shown in Figures 6.10 and 6.11 respectively.

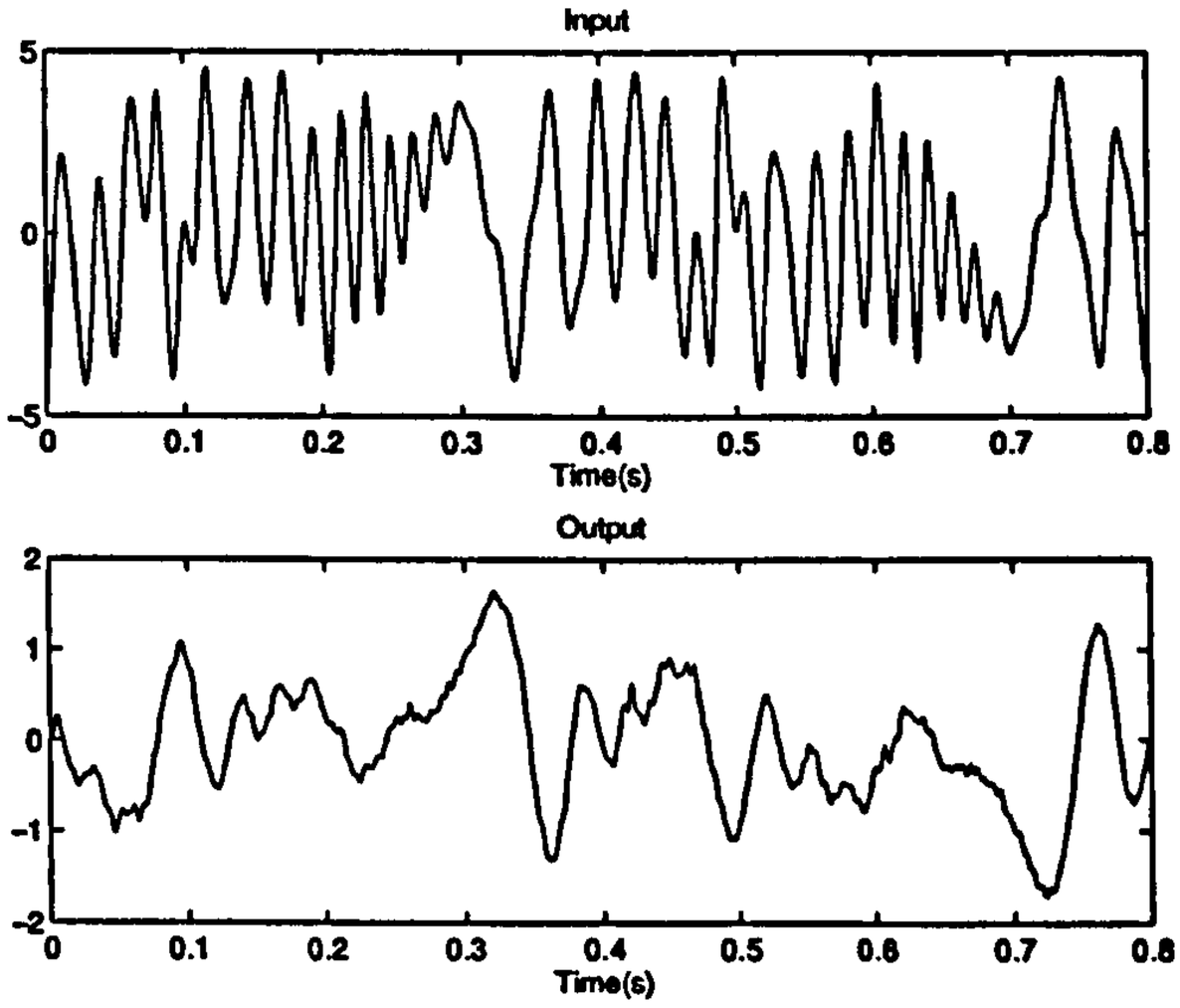

Figure 6.10: Input and output time series for the nonlinear feedback example

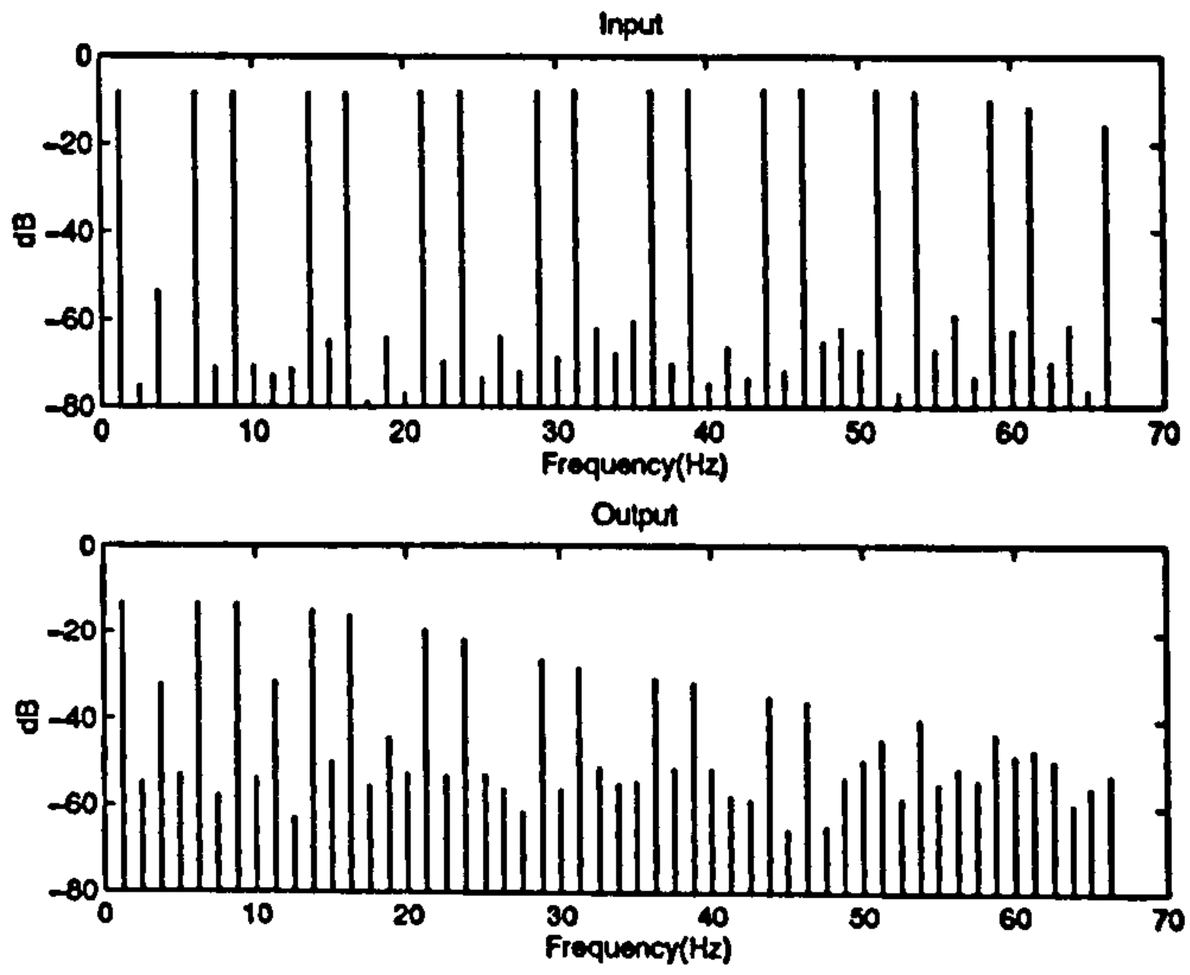

Figure 6.11: Input and output amplitude spectra for the nonlinear feedback example

The model equation for the system shown in Figure 6.11 is,

$$Y(j\omega) = \frac{1}{A(j\omega)}\{X(j\omega) - NL(Y(j\omega))\} \tag{6.26}$$

where, $NL(.)$ represents the saturation element in the feedback loop, and $1/A(j\omega)$ represents the linear dynamics. The saturation element can be crudely approximated by a combined linear and cubic function, i.e. writing eqn. (6.26) as

$$Y(j\omega) = \frac{1}{A'(j\omega)}\{X(j\omega) - c_3 Y^3(j\omega)\} \tag{6.27}$$

where $Y^3(j\omega)$ are the Fourier coefficients of $y^3(t)$ and $1/A'(j\omega)$ represents the linear dynamics in the linear feedback loop. The result of the 19 identification trials are shown in Figures 6.12 and 6.13.

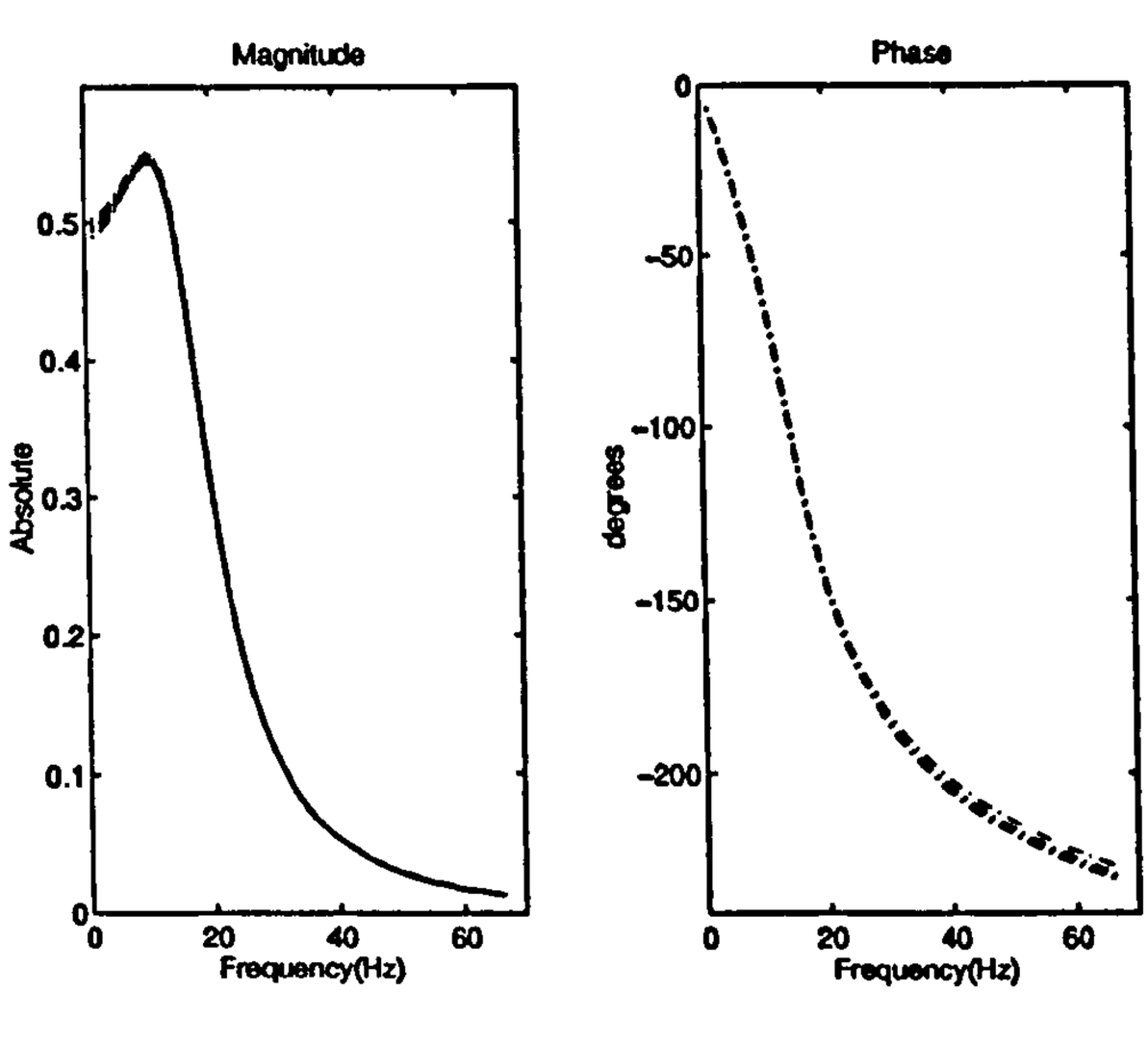

Figure 6.12: Estimated linear model frequency response functions for the nonlinear feedback example

It can be seen that the nonlinear least-squares procedure, based on the model of eqn. (6.27), has provided accurate estimates of the linear and nonlinear charac-

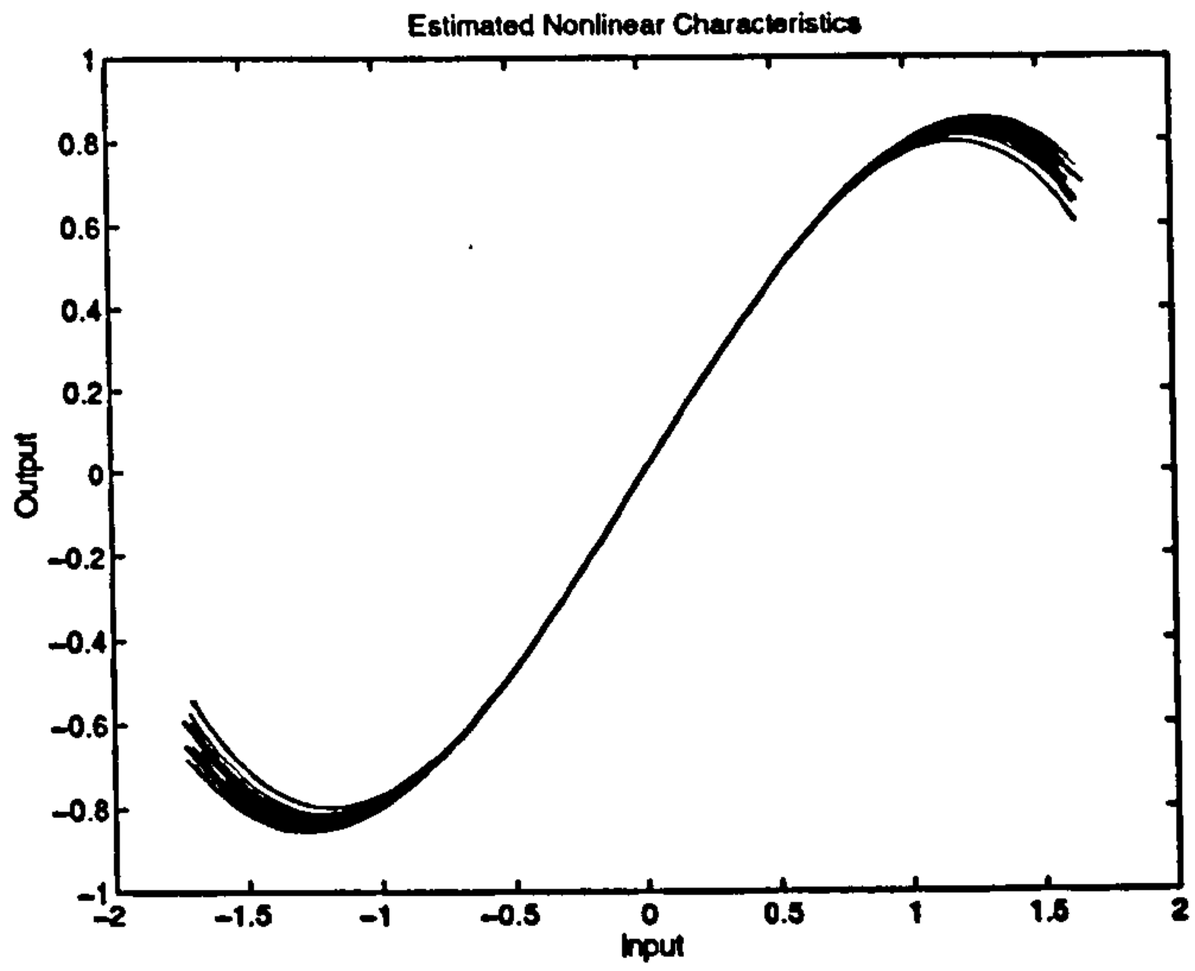

Figure 6.13: Estimated nonlinear characteristic for the nonlinear feedback example

teristics. The use of the period excitation has ensured that the model parameters are well defined, and that high accuracy has been achieved. Note that the presence of multiplicative noise terms (see Chapter 5) means that this consistency is expected to break down at higher noise levels. It should also be noted that the estimates obtained from the linearised cost function were unusable in this example, but proved adequate for initial conditions. Hence the use of the frequency-domain measurements has proved advantageous for the estimation of the continuous-time parameters in the nonlinear model of this example.

The use of the known structure of the model may be seen to be an idealisation in practice, and owing to the iterative nature of the parameter estimation method, the orthogonal least-squares procedures (Korenberg et al., 1988; Chen et al., 1989), which aid structure detection for discrete-time models, are not directly applicable, except to the linearised model equations. The similarity of the identification method to deterministic methods of linear system identification in the frequency-domain, surveyed by Pintelon, Guillaume, Rolain, Schoukens and Van hamme (1994a), is

immediately apparent. For linear system identification, structure selection based on linear estimation methods is not advised (Pintelon et al., 1994a), and so is expected to be even more hazardous when the model is nonlinear. Further work is needed in this area if the benefits of frequency-domain methods are to be of use in nonlinear system identification.

6.4 Conclusions

Linear and nonlinear modelling of systems, having commonly encountered structures, has been treated in this chapter. The representation of the system using a linear model is, at the present time, the most widespread assumption, and it has been shown that periodic excitations possess two properties which are useful when nonlinearities are likely to effect the measurements; a low crest factor and, if desired, a perfectly symmetric amplitude PDF. The combination of these features eliminates and reduces, respectively, the effects of even- and odd-order nonlinear functions on the linear representation of the system. This material is included mainly to refute the recommendation by Advantest Corporation (1989) and Hewlett Packard Co. (1989), that a random excitation should be used when a linear model is to be identified for a nonlinear system.

The main results in this chapter are from the area of nonlinear system identification. The results presented, and the identification algorithms developed, rely completely on the utilisation of a periodic excitation. Therefore, one must assess whether the potential benefits actually justify this restriction. In this area one must always be careful not to place too much emphasis on what is, after all, only one part of the complete identification strategy.

To begin with, it was shown that specific harmonic sets can be used to measure the second- and third-order frequency response functions. The measurement of the

second-order frequency response is something that is useful since its visualisation is straightforward using suitable computer graphics. Also an adequate number of points may be measured using a single application of a perturbation signal, albeit with a long signal period. A set of harmonics was given that allows a reduction in the experimental time, at the expense of a slightly reduced number of measurements.

The situation for the third-order frequency response measurement is different in that even greater restrictions are enforced on the input signal, and, even with this, several experiments would have to be conducted to adequately cover the frequency space. However, if one wishes to estimate such a quantity in the first place, without estimating a parametric model (with its associated difficulties: model/structure selection), the problem of a long experimental time is not specific to periodic excitations. In this way, the signals considered in Section 6.3.2 are an alternative to the random excitations considered by Bendat (1990).

The second class of nonlinear models considered were those that can be represented by a linear element in series with a single static nonlinear characteristic. In this case, non-parametric methods (or a similar parametric method) could be used to identify the Hammerstein model, in a simple manner, if the nonlinear characteristic could be represented by a third-order polynomial. This identification method makes direct use of the energy in the output signal, at frequencies not present in the input signal, and so is the most intuitive utilisation of perturbation signals which have gaps in their power spectrum. However, this simplification does not extend to even a Wiener model (except in the case of a purely quadratic nonlinearity), and so the advantages of leaving out harmonics for nonlinear identification (as opposed to detection, see Chapter 1 and McCormack et al. (1995)) are not as obvious. Nonetheless, an identification strategy for the Wiener model was described which utilised the "harmonic gaps". Using this, the parameters of a continuous-time linear system and a polynomial model of the nonlinearity may be identified, but this technique

is limited to a third-order approximation of the nonlinear system. Hence the only advantage, *generally*, in using a periodic signal, will be the high signal-to-noise ratio achieved and the possibility to identify the continuous-time parameters, since any other advantages will generally be coupled to a particular identification strategy.

The final part of the chapter considered a subset of the class of nonlinear models which may be written as linear in the parameter differential equations. Identification of these models was considered in the frequency-domain since the continuous-time representation is a direct extension of the linear case. It was shown that periodic signal designs, along with a nonlinear least-squares procedure, can produce accurate estimates of the continuous-time parameters. The use of periodic signals is also implicitly assumed by Patra (1993), although a different identification strategy is used. The big question is whether the transferral of the periodic data (assuming a periodic signal is used for its high power content) to the frequency-domain is worthwhile when a time-domain identification method (such as that developed by Van hamme (1992)) may be more applicable. The answer to this question is probably no. However, at least for the models considered in Section 6.3.3, there is no disadvantage, and given the relative simplicity, it may be useful in practice. The extension of the time-domain modelling procedures developed by Van hamme (1992) to the model of eqn. (6.20), and a comparison of the accuracy achieved when applied to the models considered here, will be an area of future research.

Chapter 7

System Identification and Rule Based Controller Design for Autotune Control

7.1 Introduction

One of the main reasons for the popularity of PID controllers in industry is the simplicity involved in tuning the controller parameters. Another reason is the fact that PID controllers can be used to obtain satisfactory performance from typical industrial plant (Åström, Hägglund, Hang and Ho, 1993; Institution of Electrical Engineers, 1994). Traditionally, the tuning procedure has been achieved using rules such as the Ziegler-Nichols tuning formulae, which require only a limited knowledge of the process characteristics. The recent resurgence of interest in PID parameter tuning has recognised this acceptance of the PID controller, and also the need to keep the required knowledge of the process plant to a minimum (Hang, Åström and Ho, 1991; Zhuang and Atherton, 1993; Voda and Landau, 1995). Although some authors have considered the tuning of the controller parameters from a model based viewpoint (Schei, 1994), this seems to leave behind the strong asset of simplicity which comes with the PID structure.

This chapter addresses the problem of *automatically* tuning (autotuning) the PID parameters, and assumes that a rule based controller design is to be used. Many of the more recent developments in autotuning have adopted the same rationale as developed by Ziegler and Nichols (1943), meaning that either an open-loop step response test is carried out, or an estimate of the system frequency response is obtained. A popular method for obtaining the latter is to replace the system controller with a relay element, and allowing the process output to limit cycle for a short duration, as introduced by Åström and Hägglund (1984). Although this method of system identification has the advantage that a suitable input signal is generated without the need for *a priori* knowledge of the process characteristics, the replacement of the controller by a relay element can seem overly harsh.

Since closed-loop identification is all that is required for an autotune controller,

this work considers the frequency response estimation problem, assuming that a low-amplitude perturbation signal can be applied to the plant for the purposes of identification. In this way, normal control is maintained during the tuning procedure, with just a small disturbance affecting the setpoint. This is believed to be more favourable in practice than relay excitation, especially if regular retuning is necessary, e.g. when the process is nonlinear over its operating region.

The chapter is organised as follows. Section 7.2 gives a summary of the (known) rule based controller designs which have been postulated in the literature, together with the process characteristics which are required to implement them. Signal design issues are discussed in Section 7.3, with these very closely linked to the specific rule based design used for controller tuning. A comparison of several approaches to the frequency response estimation problem is given in Section 7.4. The identification results and the performance of the controller designs are assessed in Section 7.5, with a laboratory scale system used as a bench-test. Finally in Section 7.6, some conclusions are drawn.

7.2 PID Tuning Rules

The tuning rules that are compared in the examples of Section 7.5 are the designs based on: Ziegler-Nichols (Ziegler and Nichols, 1943), refined Ziegler-Nichols (Hang et al., 1991), ISTE Criteria (Zhuang and Atherton, 1993), and the KLV method (Voda and Landau, 1995). Table 7.1 lists the information required, and the tuning formulae for all the designs. These methods are suitable for direct comparison since no design variable based on desired process characteristics is used. Controller designs such as those based on specified phase and amplitude margins (Åström and Hägglund, 1984), and the M-circle design criteria (Åström and Hägglund, 1988), need also to be compared, but these are dependent upon the specified loop perfor-

mance. The evaluation of other controller designs is to be carried out as a part of future research. In the table, $1/K_u$ and $1/T_u$ denote the system gain and frequency

Rule	*Process Information*	K_p	T_i	T_d
Z-N	K_u, T_u	$0.6K_u$	$0.5T_u$	$0.125T_u$
RZ-N	K_u,T_u,K	$0.6K_u$	$0.5T_u^{(1)}$, $\frac{2}{9}\kappa T_u^{(2)}$	$0.125T_u$
ISTE	K_u, T_u, K	$0.509K_u$	$0.051(3.302\kappa+1)T_u$	$0.125T_u$
KLV	K_{135},ω_{135}	$\frac{4+B}{4}\cdot\frac{B}{2\sqrt{2}K_{135}}$	$\frac{4+B}{B}\cdot\frac{1}{\omega_{135}}$	$\frac{B}{4+B}\cdot\frac{1}{\omega_{135}}$
(1) $\beta=\frac{15-\kappa}{15+\kappa}$; $2.55<\kappa<15$				
(2) $\beta=\frac{8}{17}(\frac{4}{9}\kappa+1)$; $1.5<\kappa<2.55$				

Table 7.1: Summary of rule based controller designs

at which the phase is -180 degrees. Similarly K_{135} and $\omega_{135}/2\pi$ denote the system gain and frequency at which the phase is -135 degrees. The d.c. gain of the system is denoted by K, with $\kappa = K.K_u$. The variable B in the KLV rule is an acceleration factor which can take a value between 1 and 2 and is assumed to be 1.5 in this work. The variable β in the refined Ziegler-Nichols method is a weight applied to the setpoint in the proportional term of the controller.

7.3 Excitation During Tuning

This section treats the frequency-domain properties of periodic excitation signals, which can be used to identify the open-loop frequency response of a system. There is large scope for choice here since the controller design rules differ significantly in the information that they require. Effectively the main decision to be made is whether to use a narrow-band or broadband excitation. If a single estimate of the system frequency response is required, e.g. as for the Ziegler-Nichols method (Ziegler and Nichols, 1943), and the KLV method (Voda and Landau, 1995), a narrow-band excitation centred around the frequency point of interest would be advantageous. However, if some knowledge of the low-frequency behaviour of the system is also

required, e.g. as for the refined Ziegler-Nichols method (Hang et al., 1991) and the ISTE method (Zhuang and Atherton, 1993), a broadband excitation would be necessary. These two situations are depicted in Figures 7.1 and 7.2 respectively.

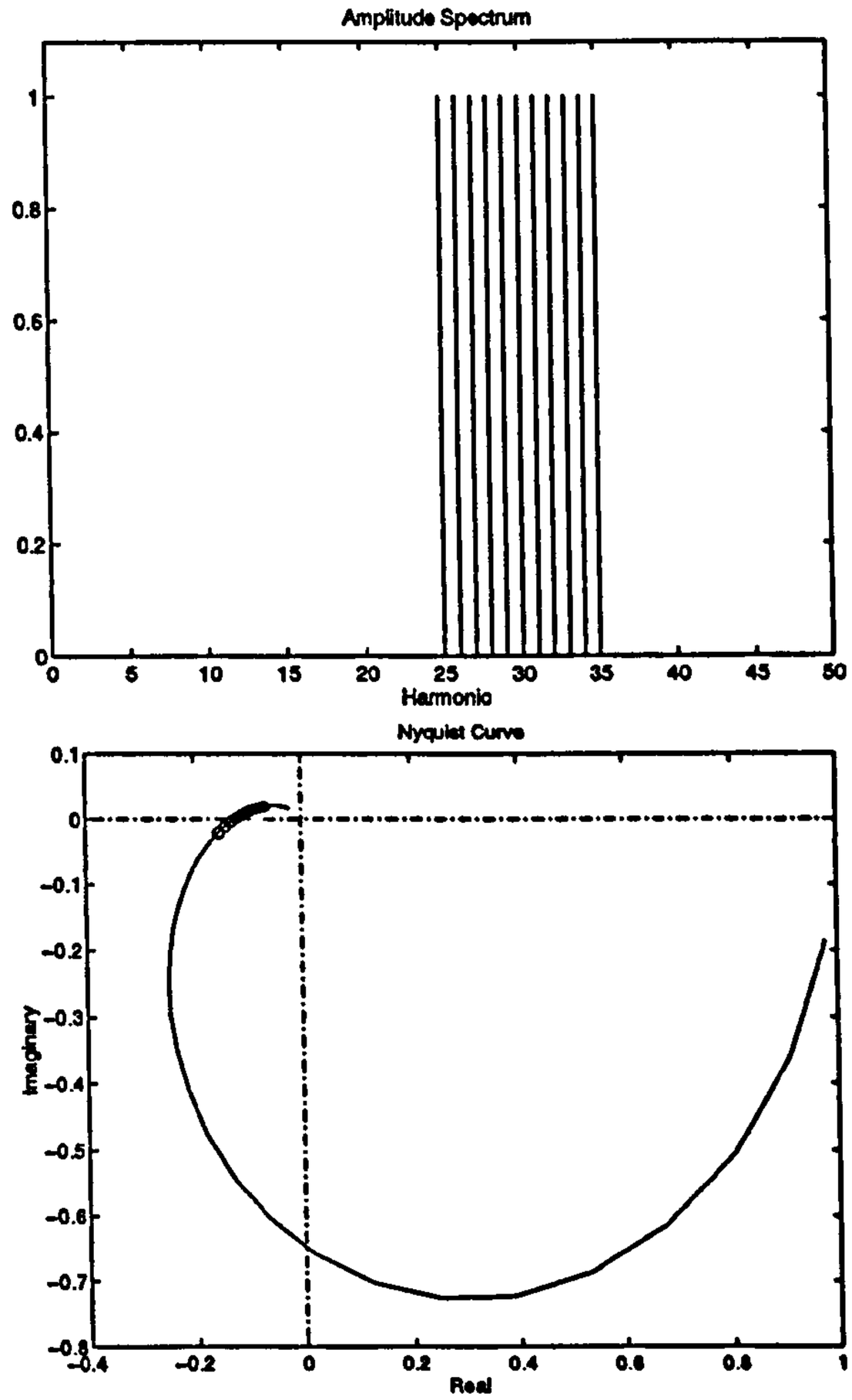

Figure 7.1: Upper: amplitude spectrum of narrow-band excitation signal. Lower: system Nyquist curve (solid) and portion estimated using narrow-band signal (circles).

Naturally, there are intermediate choices, one of which is shown in Figure 7.3. Using the spectra shown in Figure 7.3, both the low- and high-frequency behaviour of the system can be estimated, with energy present in the excitation only in areas where an estimate is needed.

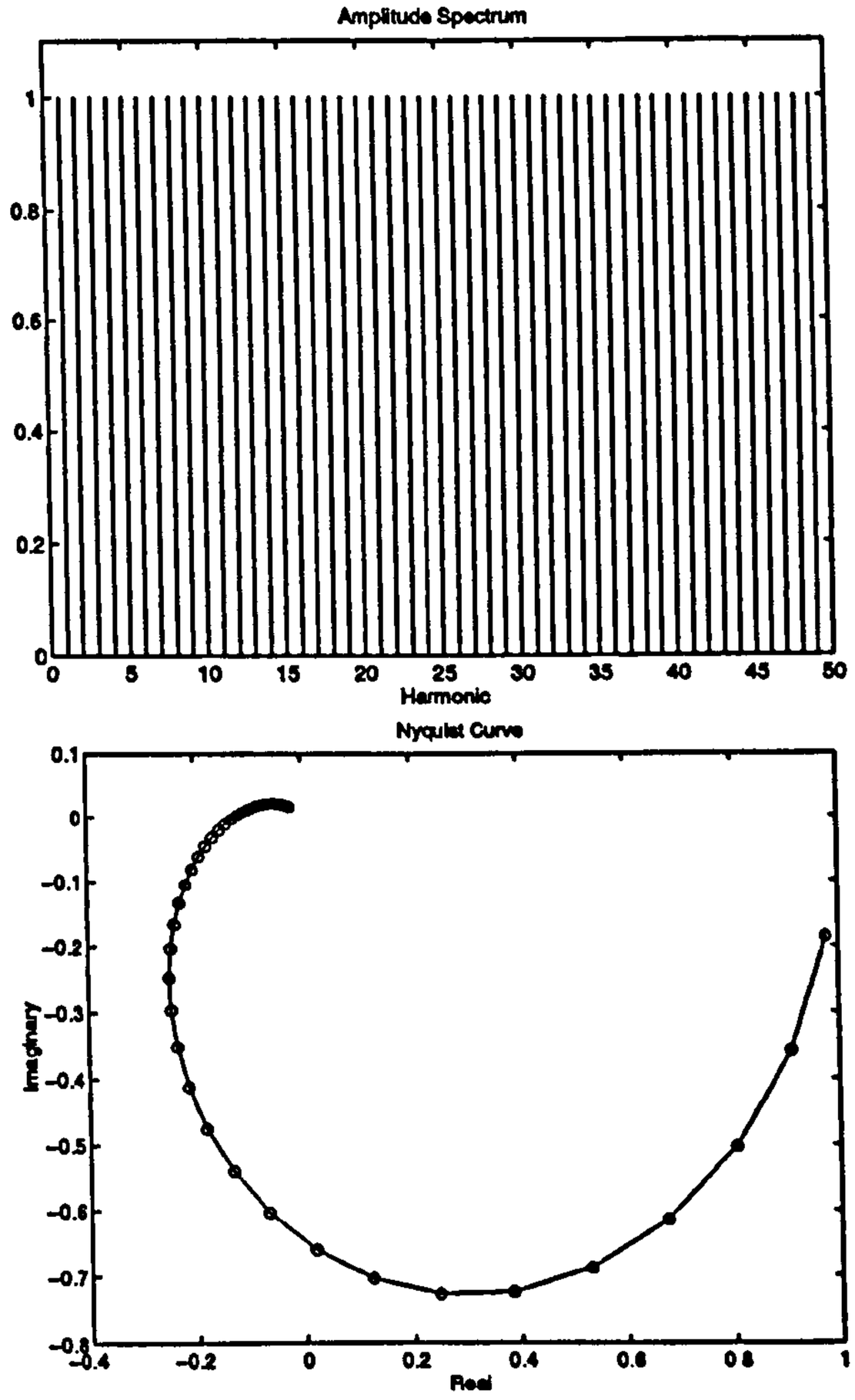

Figure 7.2: Upper: amplitude spectrum of broadband excitation signal. Lower: system Nyquist curve (solid) and portion estimated using broadband signal (circles).

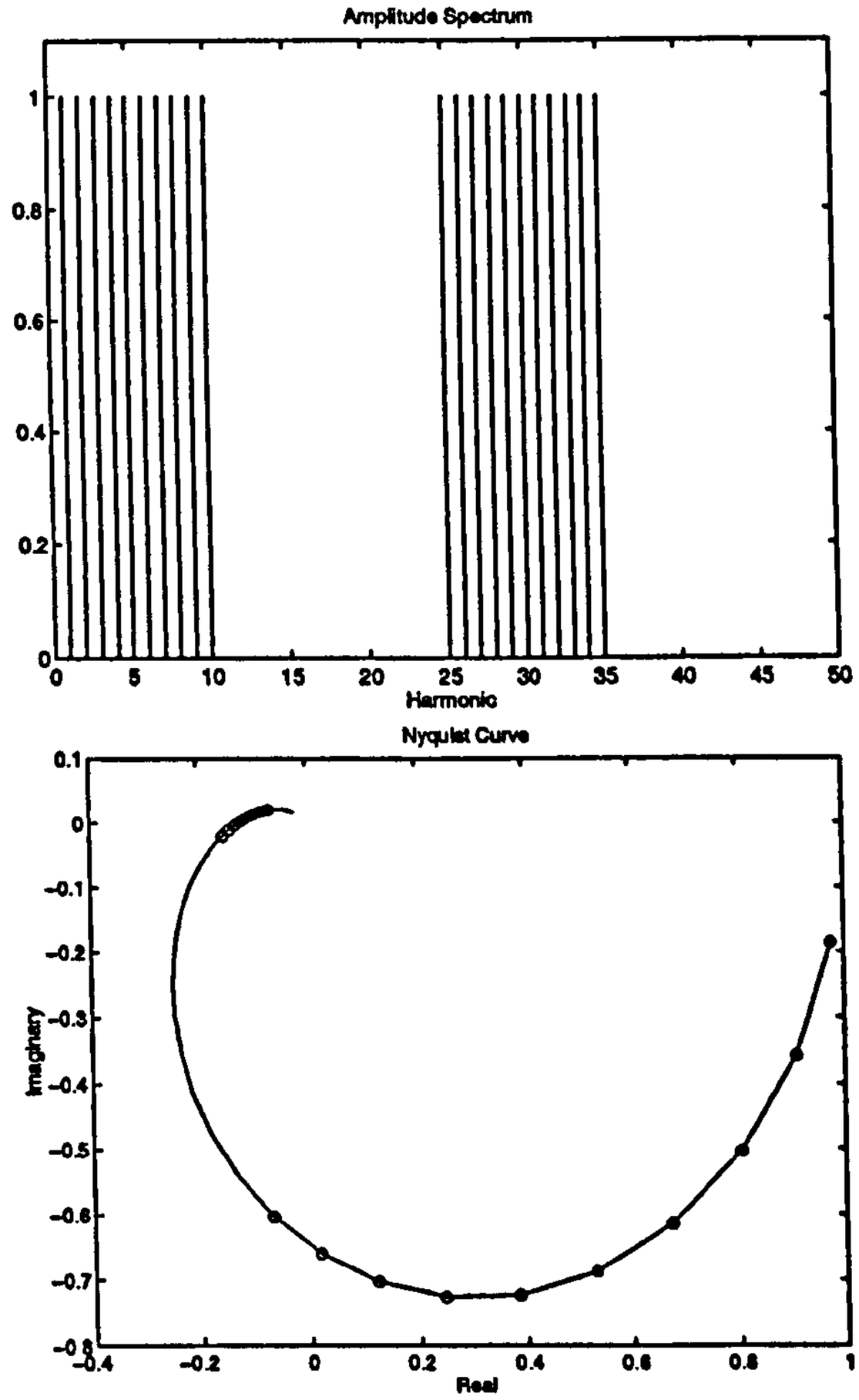

Figure 7.3: Upper: amplitude spectrum of stop-band excitation signal. Lower: system Nyquist curve (solid) and portion estimated using stop-band signal (circles).

The main novelty of the proposed autotune method is the ease with which the spectra shown in Figures 7.1, 7.2 and 7.3, can be realised with the periodic signals used throughout this thesis. The three classes of periodic signals, the binary (Van den Bos and Krol, 1979), the m-level (McCormack et al., 1995), and the multisine (Van der Ouderaa et al., 1988a), all have the ability to realise these spectra in time-domain forms with a very high power content in the desired harmonics. The binary signal has the advantage of yielding a signal with a very high power content in a time-domain form consisting of just two amplitude levels. However, and similarly for the m-level signal, a band-limited design is impossible, thus making over-sampling necessary to reduce the effects of aliasing. In this regard, the multisine with snow (Guillaume et al., 1991) has the property of a very high power content, and can also be specified to be band-limited. The signal spectra shown in Figure 7.4 illustrate results produced with the binary, the m-level ($m = 4$) and the multisine with snow signals, all designed to meet the specification shown in Figure 7.3.

It must be noted that for a given maximum frequency, e.g. the critical frequency, the spectra shown in Figures 7.1, 7.2 and 7.3, will require signal periods of approximately the same length. Therefore the spectrum shown in Figure 7.3 is a good compromise, since the accuracy of the most important points on the frequency response is being accentuated. However, constraints on the tuning time would generally mean that a fixed ratio of the signal period to the settling time of the system must be used. If this is the case, a broadband signal is the only viable option, since constraining the high-frequency harmonics in Figure 7.3 to be at a certain frequency may dictate an excessively long signal period.

However, even if a broadband excitation is used, considerable flexibility is available. In this application, the main benefit of the using periodic excitations will be gained if non-parametric identification can be used (similar to relay excitation). To

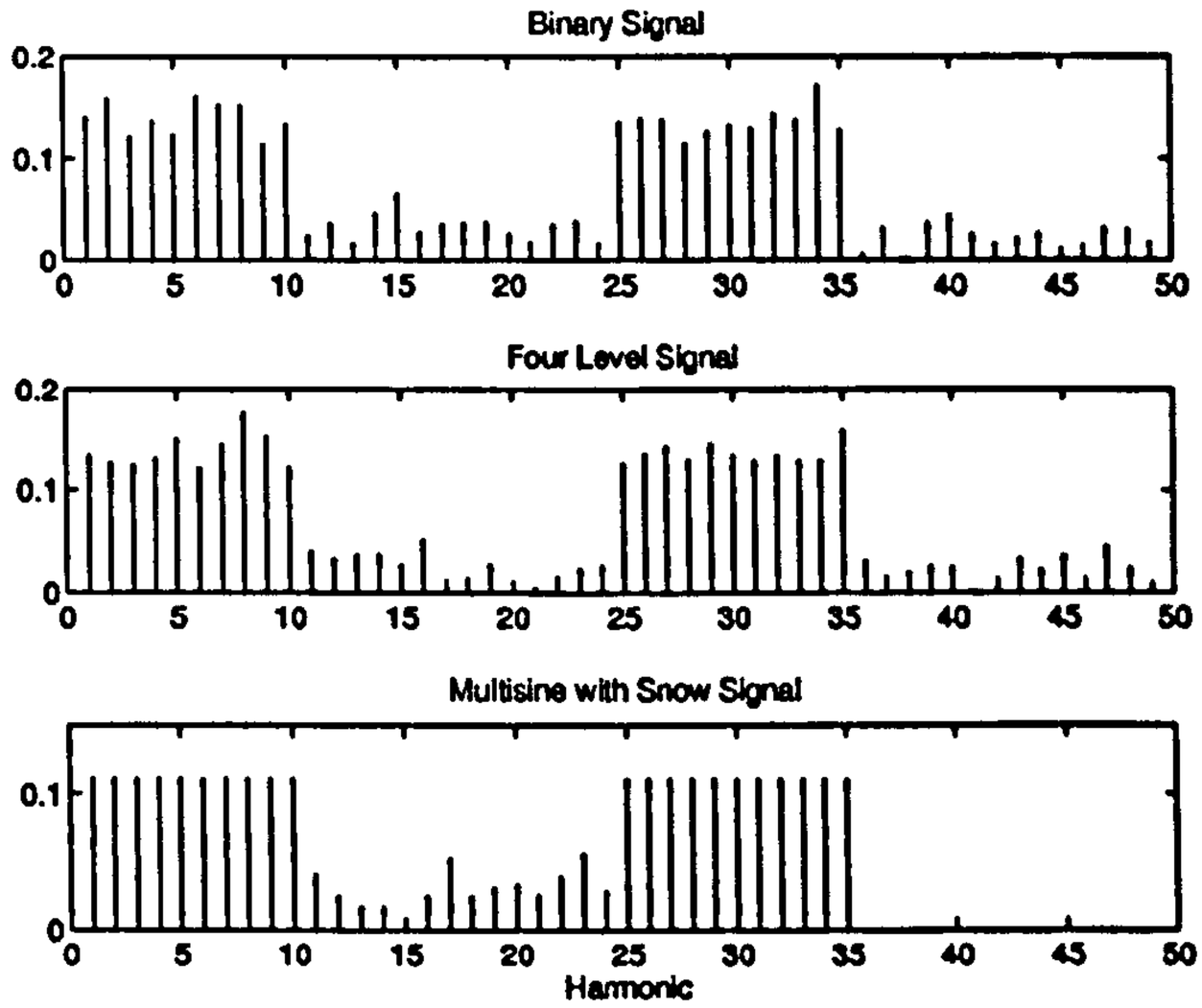

Figure 7.4: Amplitude spectra realized for stop-band spectra with a binary, four-level and multisine with snow signal.

this end, the excitation signal will be designed to increase the accuracy of the non-parametric frequency response function. Since no knowledge of the system frequency response can be assumed *a priori*, the only general way to achieve an increase in accuracy is to compress the power in the signal into a smaller number of harmonics. The broadband signal used from here on is shown in Figure 7.5. As can be seen, a consecutive harmonic signal is used, with the first fifteen harmonics specified to have equal power. Snow is allowed at harmonics sixteen to forty, so as to increase the power in the desired harmonics. This is believed to allow good coverage of the dynamic range of typical low-pass systems, while also allowing a high power content in a comparatively small number of harmonics. The application of this signal is considered in Section 7.5.

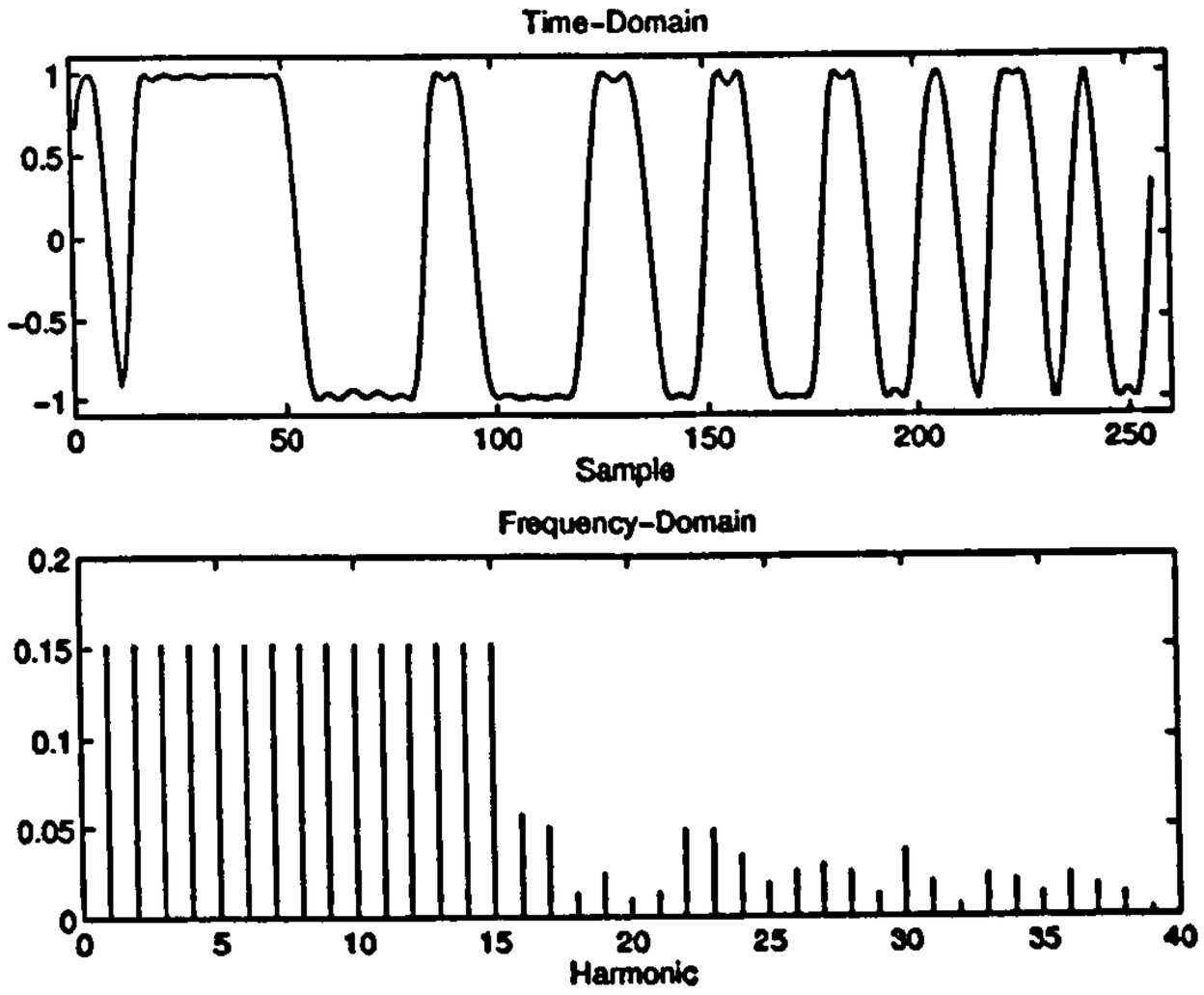

Figure 7.5: Upper: time-domain realization of multisine with snow signal specified to have a broadband spectrum. Lower: amplitude spectrum.

7.4 Frequency Response Estimation

Different classes of frequency response estimators, which are appropriate for this application, are considered in this section. A conceptual feedback loop is shown in Figure 7.6. In Figure 7.6, either of the sources, r_1 and r_2, can represent the

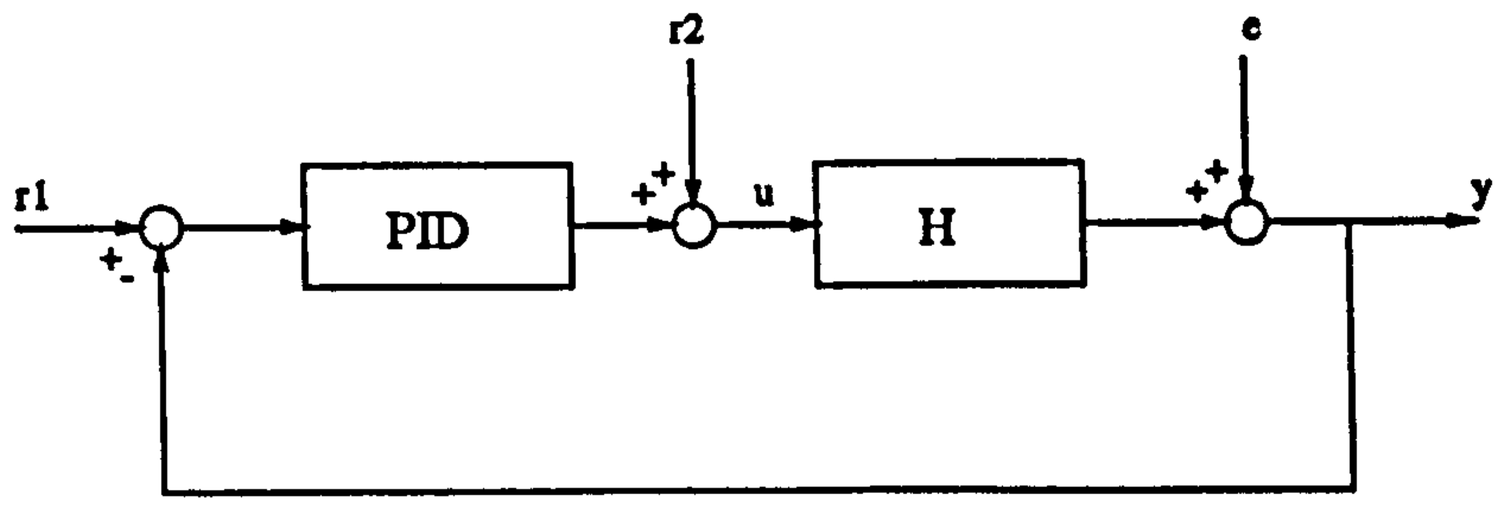

Figure 7.6: Feedback configuration assumed for identification of process dynamics

multiharmonic excitation signal, while e represents a disturbing noise source.

7.4.1 Non-Parametric Estimation

To obtain a non-parametric estimate of the forward path frequency response function, the following estimator can be used (Wellstead, 1981)

$$H(j\omega) = \frac{\hat{S}_{ry}(j\omega)}{\hat{S}_{ru}(j\omega)} \tag{7.1}$$

where $\hat{S}_{ry}$ and $\hat{S}_{ru}$ represent the cross-spectra between $[r_{1,2}(t), y(t)]$ and $[r_{1,2}(t), u(t)]$, respectively. However, to take advantage of the properties of this estimator, averaging of the quantities $\hat{S}_{ry}$, and $\hat{S}_{ru}$ would be necessary, meaning that several periods of the excitation signal would have to be applied to the system. This may not be practical in many situations.

Note also that the non-parametric closed-loop identification framework developed by Schrama (1991) is similar to using eqn. (7.1), but alleviates the need to measure $r_1(t)$. There is no advantage evident in this for our particular application. Due the correlation between $u(t)$ and $y(t)$, the application of the direct method of identification (using eqn. (7.1) with one period of the excitation), is dubious if the noise signal, $e(t)$, is significant. However, since the excitation has been designed to have a high power content at selected harmonics, the accuracy of the estimates may not be unduly affected. This will be assessed in Section 7.5.

7.4.2 Parametric Identification

Even without the presence of the feedback path, the non-parametric estimate of the frequency response function may not yield the required accuracy. Schoukens et al. (1993a) have shown that if a flat spectrum multisine is used for a non-parametric estimate, and a multisine with an optimal spectrum is used for a parametric estimate, the ratio of the measurement uncertainties for the resulting frequency response

functions is

$$\frac{\sigma_{Hnp}}{\sigma_{Hp}} = \sqrt{\frac{2F}{n_p}} \tag{7.2}$$

where σ_{Hnp} and σ_{Hp} are the uncertainties of the non-parametric and parametric estimates respectively, F is the number of frequencies in the excitation, and n_p is the number of free parameters in the parametric model. Although this large reduction in the uncertainty is unlikely to be achieved in practice, the parametric method can improve the accuracy of the results significantly in noisy situations.

However, closed-loop issues are again raised, possibly to an even greater extent than for the non-parametric case. A maximum likelihood method has been developed by Pintelon, Guillaume, Rolain and Verbeyst (1992), which is based on knowledge of the input and output spectra, $U(f)$ and $Y(f)$, together with the stochastic disturbances (with their covariance) in Figure 7.6.

The open-loop version of this algorithm, as given by Schoukens and Pintelon (1991), has been implemented as part of the autotune controller developed in this work. Since this estimator, and other frequency-domain methods are off-line methods, the computations have to be split up so that one iteration of the optimisation procedure is carried out at each sampling interval. Given good computing power, the full closed-loop estimator is a feasible option.

It is felt that time-domain methods are better in this application because of the ability to process the measurements in a true on-line manner. The parametric identification of systems operating in feedback loops has received a great deal of attention in the literature recently (Gevers, 1991; Van den Hof and Scrama, 1994). Current issues center on *approximate* identification of the forward path dynamics, and so are deemed less restrictive than prediction-error methods (Ljung, 1987). An assessment of time-domain approaches to closed loop identification needs to be carried out, and is to be an area of future research.

Very much in line with feedback issues is the assessment of model errors in parametric identification (Goodwin, Gevers and Ninness, 1992). How important under-modelling is in autotune applications needs to be assessed, but if it is found to significant, the use of the periodic excitations discussed in Section 7.3 does simplify the quantitative assessment of model errors in *open loop* system identification (Schoukens and Pintelon, 1994b). Unfortunately no similar results are available at the present time for feedback systems.

To end this section, which began with the simple task of obtaining an estimate of the open-loop frequency response function, and then went on to approximate identification in closed-loop, and the assessment of model errors in parametric identification, it must be said that these issues may prove to be academic with the type of systems which are normally under autotune PID control. They are certainly not considered when identification is carried out in other autotune methods (Hägglund and Åström, 1991; Schei, 1994). For the moment, standard frequency response techniques (albeit with a specially designed excitation signal), and open-loop parametric identification has been implemented so far, and is illustrated in the examples of Section 7.5.

7.5 Application of Autotune Control

The frequency response estimation methods, and the performance of the controller designs are now tested with the Feedback PT326 process trainer serving as an illustrative system. This system has many of the features of typical process plant, including an inherent dead-time, and a natural noise component due to turbulence within the tube attached to the output of the heater. It also has a limited linear range, as noted in Chapter 2. A nominal model structure, recommended by the manufacturers, consists of a second-order denominator with a pure time-delay. The

system has a settling-time of approximately two seconds, so the period of the excitation signal is set to ten seconds, giving a frequency range of $0.1 - 1.5$ Hz. An open-loop frequency response function was estimated, with the results of twenty periods being averaged. The mean and standard deviations of the estimates are shown in Figures 7.7 and 7.8. The results indicate that the estimates at the lower harmonics contain significant noise contribution, thus making a "smoothed" estimate desirable.

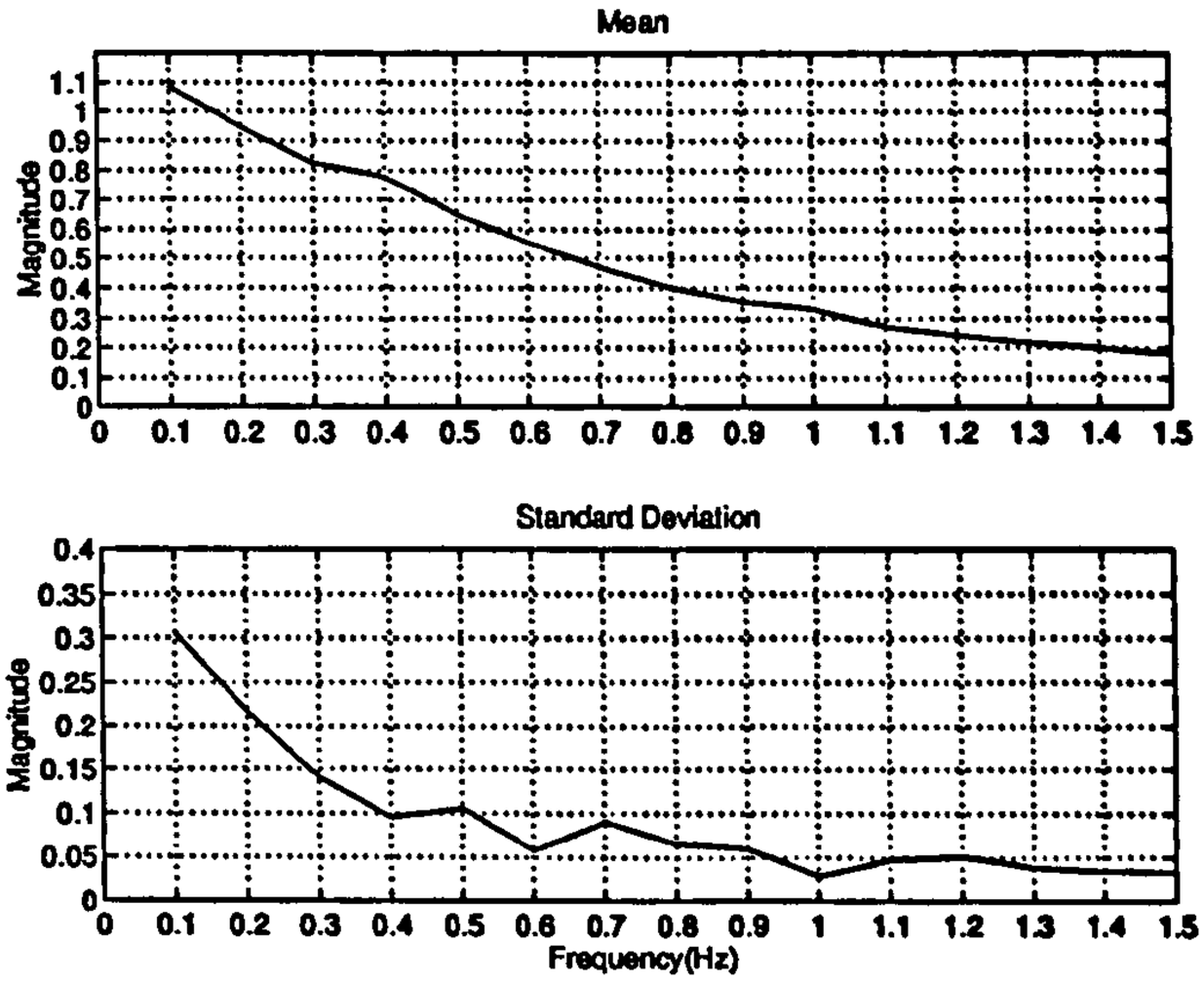

Figure 7.7: Mean and standard deviation of the empirical frequency response function magnitude, estimated for the hot-air flow device. Open-loop configuration.

The frequency response estimation was then repeated in closed-loop, with two periods of the excitation signal disturbing the setpoint during each tuning cycle. The disturbance to the output is shown in Figure 7.9. The spectra of $u(t)$ and $y(t)$ were calculated recursively during the application of the second period. The mean and standard deviations of the resulting frequency response functions (20 trials) are shown in Figures 7.10 and 7.11. It can be seen that there is no deterioration of the results when the feedback loop is present. However the large uncertainty of the

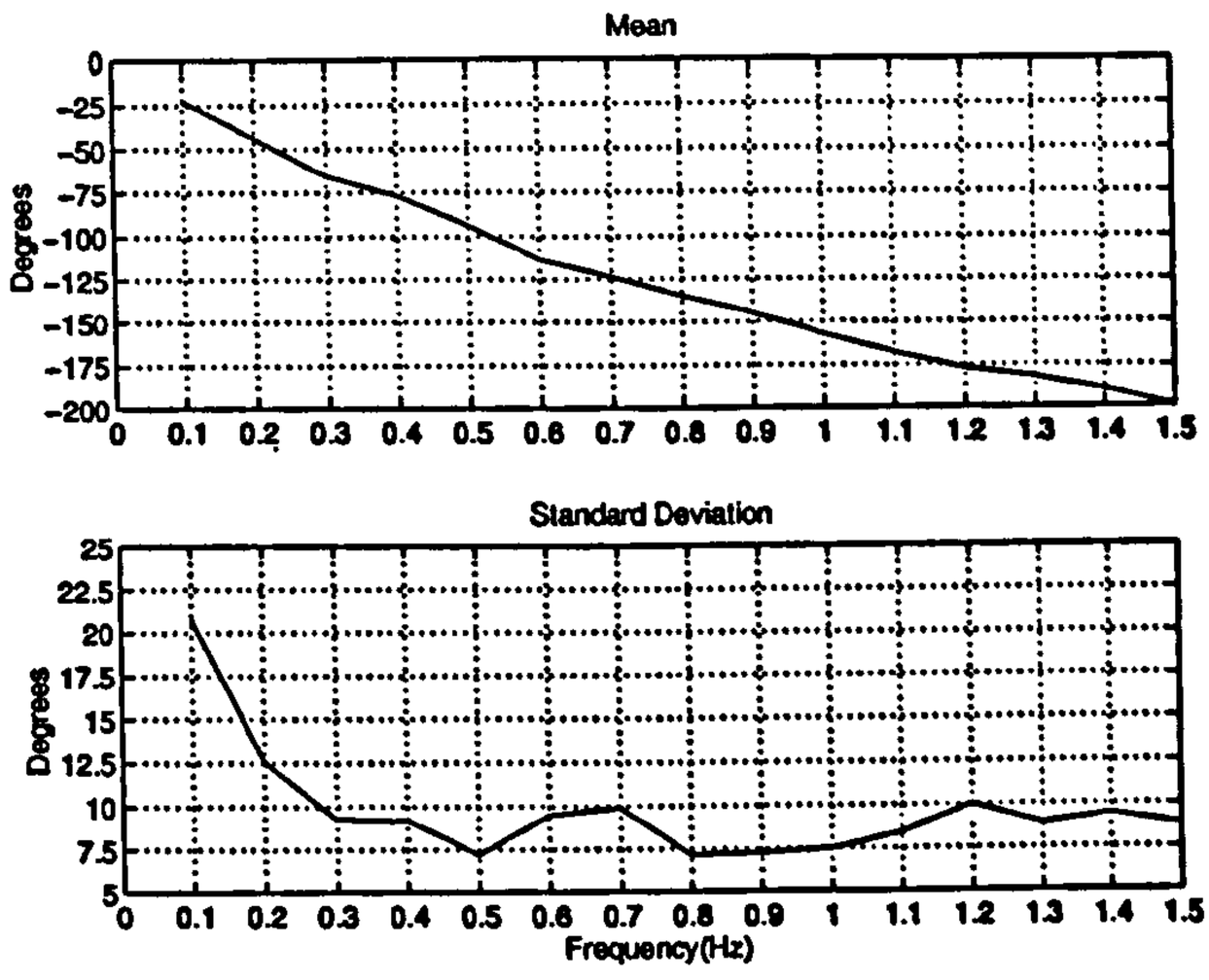

Figure 7.8: Mean and standard deviation of the empirical frequency response function phase, estimated for the hot-air flow device. Open-loop configuration.

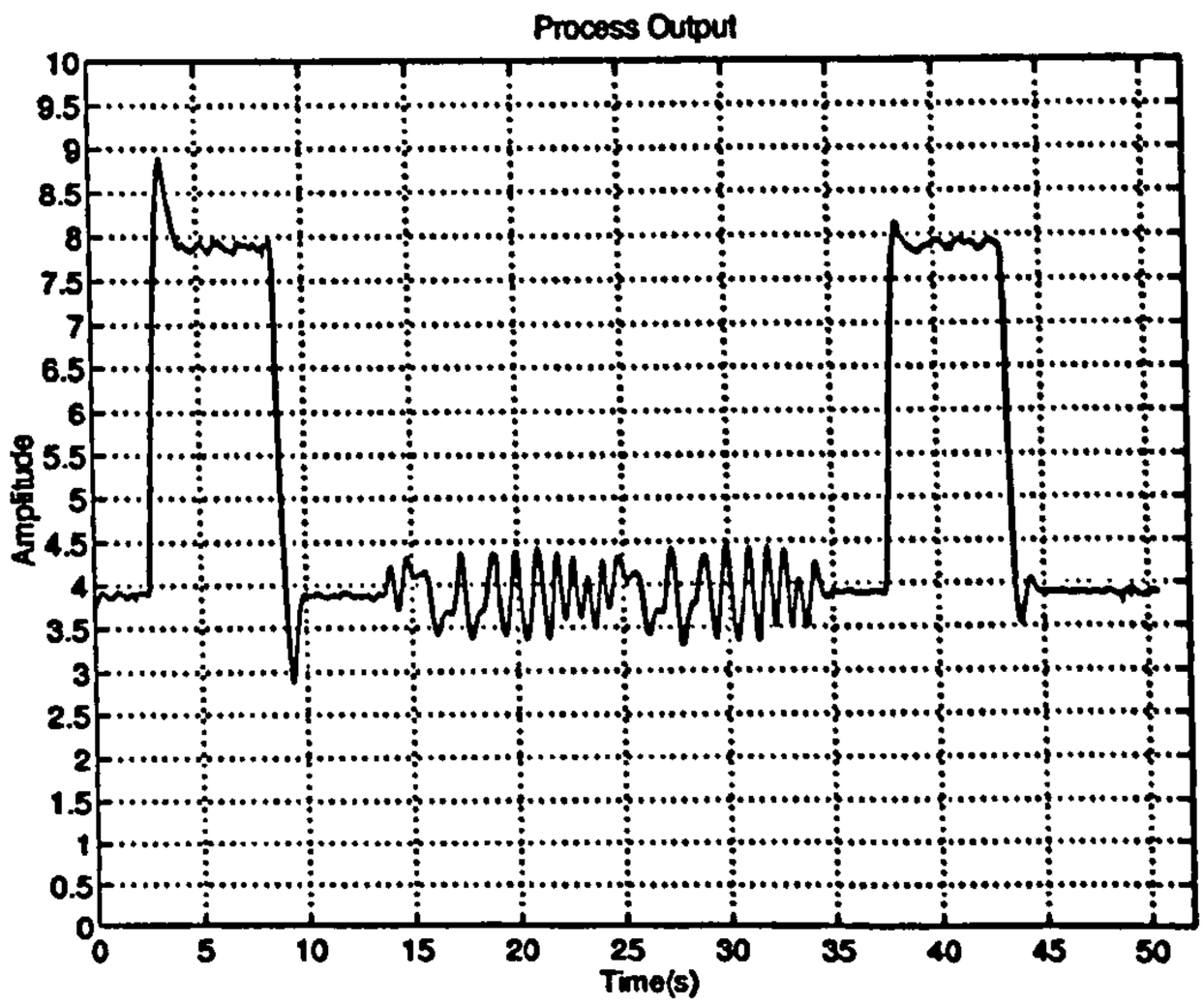

Figure 7.9: Output of hot-air flow device showing step response, tuning phase, and re-tuned step response

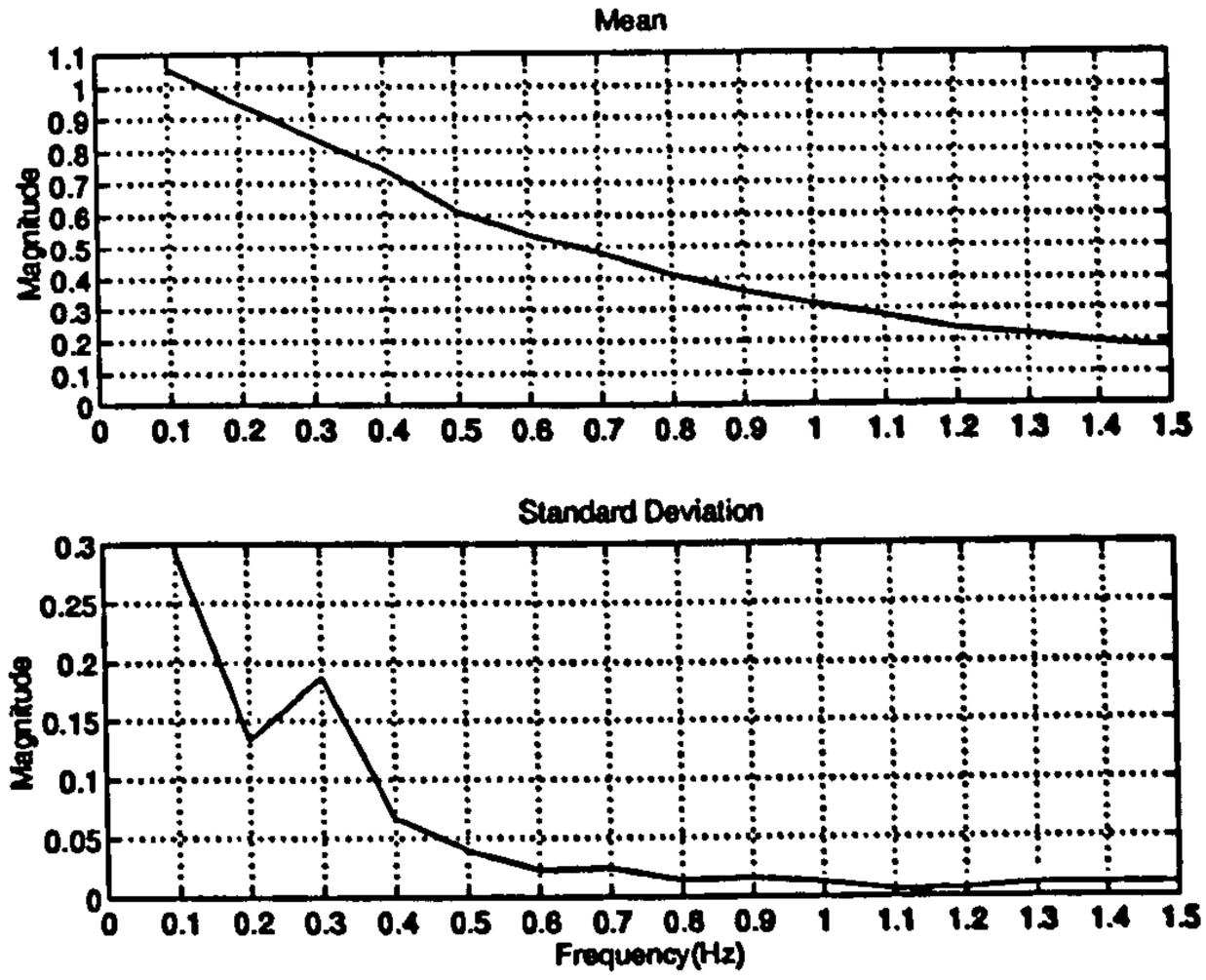

Figure 7.10: Mean and standard deviation of the empirical frequency response function magnitude, estimated for the hot-air flow device. Closed-loop configuration.

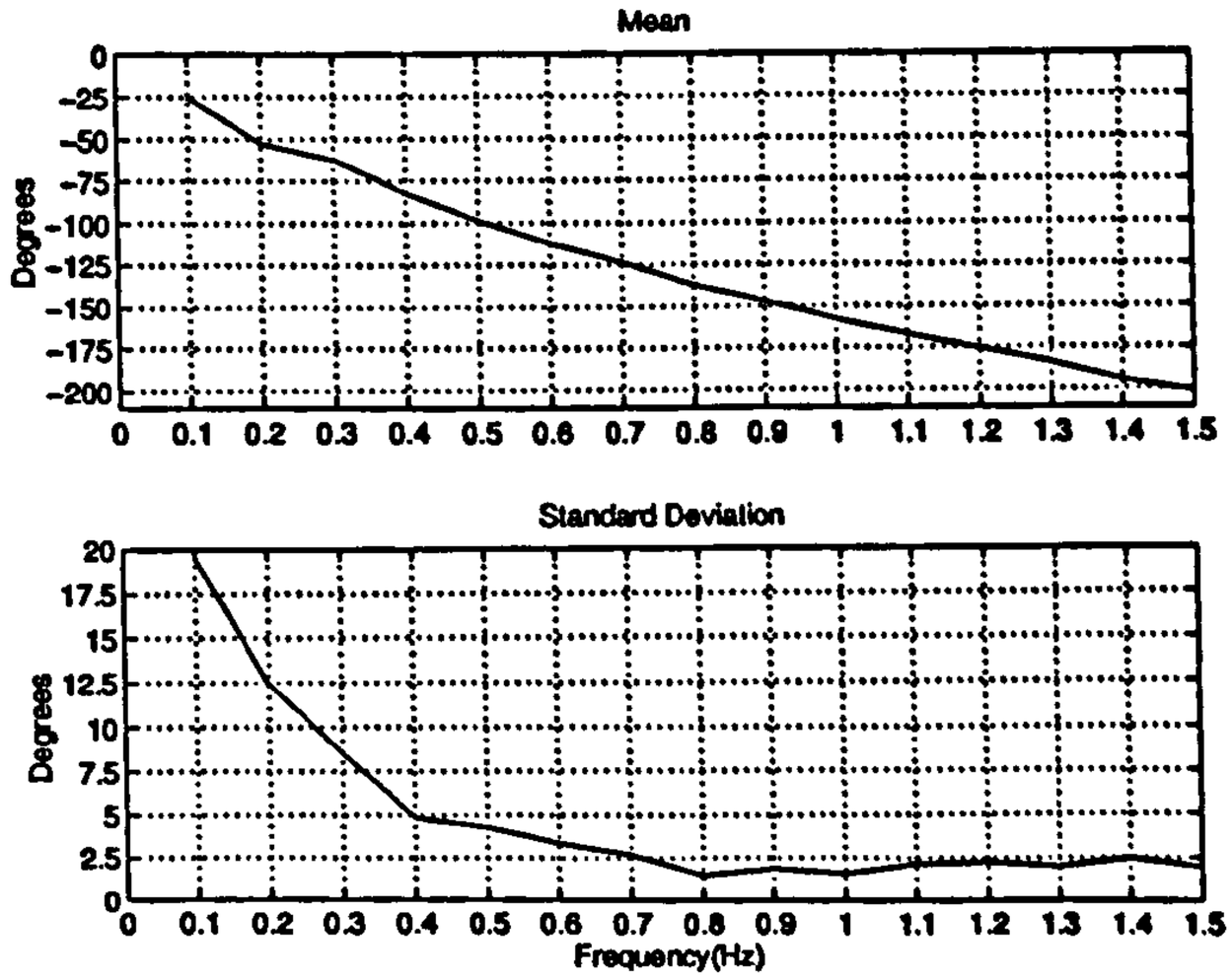

Figure 7.11: Mean and standard deviation of the empirical frequency response function phase, estimated for the hot-air flow device. Closed-loop configuration.

low-frequency measurements still means that a parametric model is required if this information is to be used confidently for the design of the controller.

During normal operation of the controller, a Laplace transfer function with first-order numerator and second-order denominator was estimated from the spectra of $u(t)$ and $y(t)$. Note that this represents the system dynamics as having a zero-order numerator, first-order denominator. The process dead-time is being encompassed within a non-minimum phase first-order system. This procedure was repeated twenty times, the same as for the non-parametric identification. Figures 7.12 and 7.13 show the mean and standard deviation of the frequency response functions calculated with the parametric models. It can be seen that these estimates show a

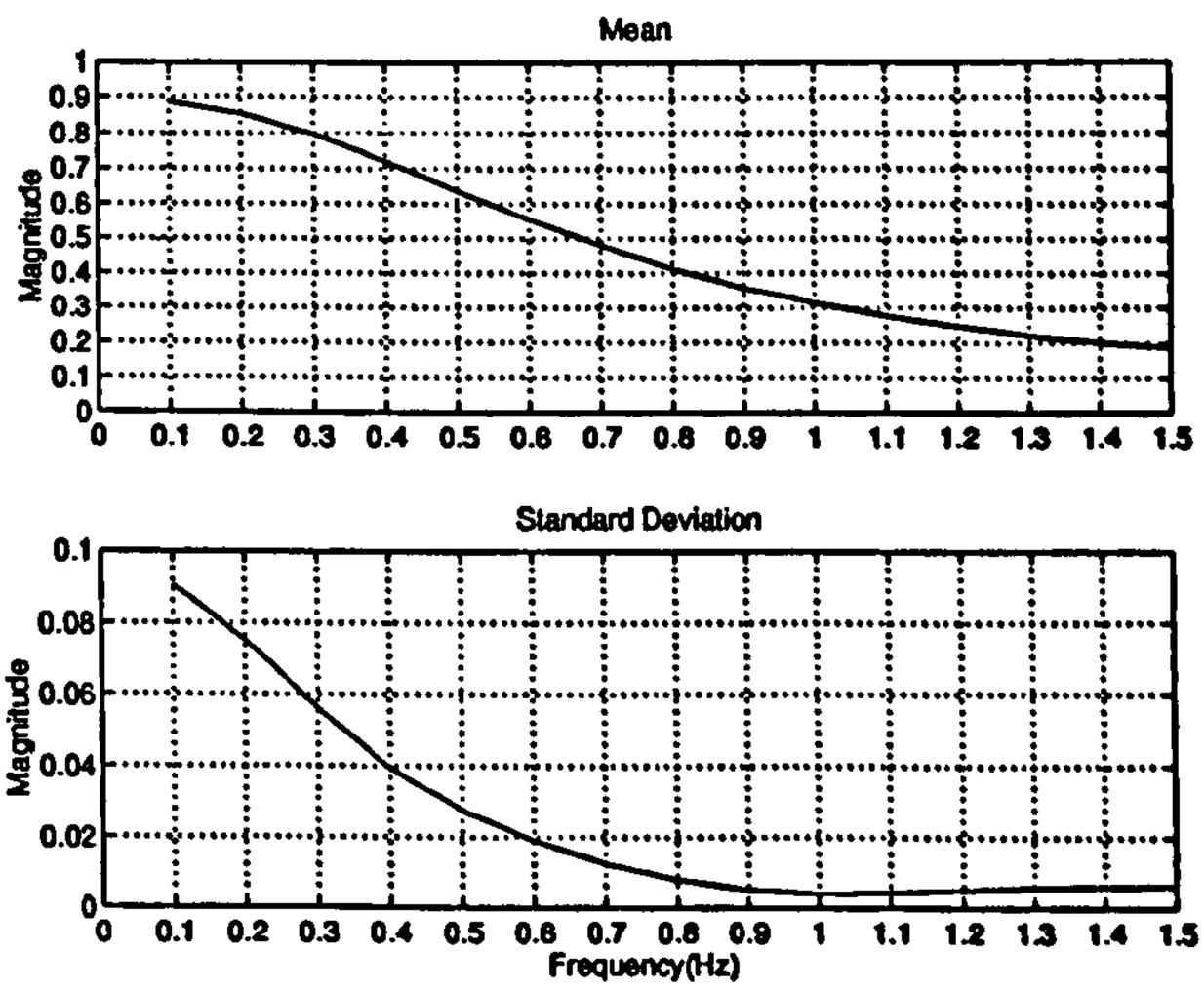

Figure 7.12: Mean and standard deviation of the second-order model frequency response function magnitude, estimated for the hot-air flow device. Closed-loop configuration.

greatly increased accuracy, especially at the low-harmonics. However the second-order structure (non-minimum phase) is unable to match the phase response of the system, resulting in the phase response being systematically under-estimated. A first-order numerator, third-order denominator model was then specified, and the

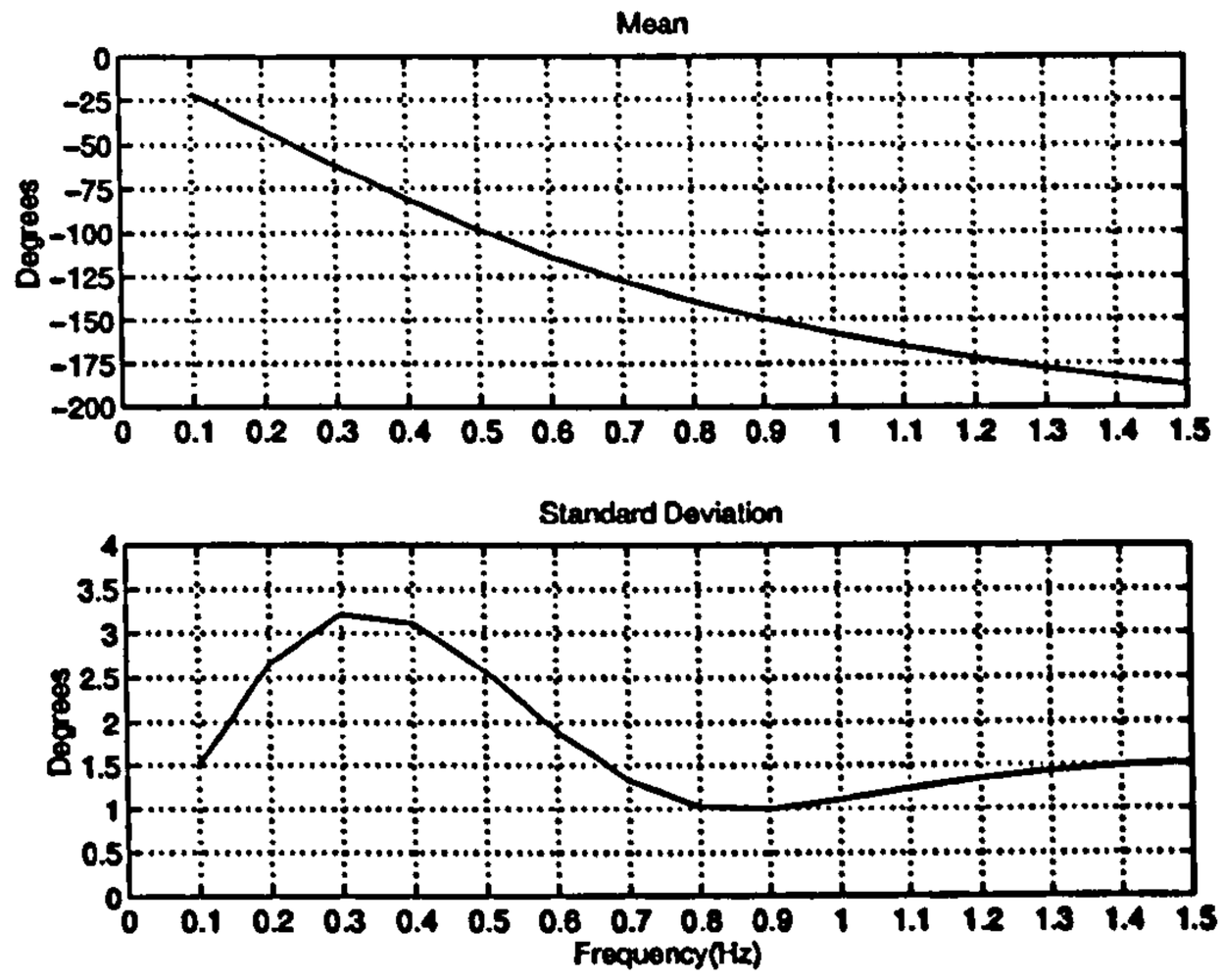

Figure 7.13: Mean and standard deviation of the second-order model frequency response function phase, estimated for the hot-air flow device. Closed-loop configuration.

procedure repeated. The mean and standard deviations of the resulting frequency response functions are shown in Figures 7.14 and 7.15. This time, the model characterises the system well, also with good accuracy. However, numerical problems were encountered when fitting this model (parameters strongly coupled), which meant that this structure could not be used confidently. Any attempts to alleviate the numerical problems inevitably increases the duration of the tuning phase, which is assumed not to be an option here. Since the non-parametric estimate is sufficient at the higher harmonics, and the second-order model is sufficient for the characterisation of the low-frequency behaviour, it is possible to envisage the joint use of this information.

In this feedback system, the accuracy of the non-parametric and parametric open-loop identification strategies can be seen to be very good. The accuracy of the non-parametric estimate is satisfactory at the higher harmonics, and so the Ziegler-Nichols and KLV designs could take advantage of the empirical frequency response

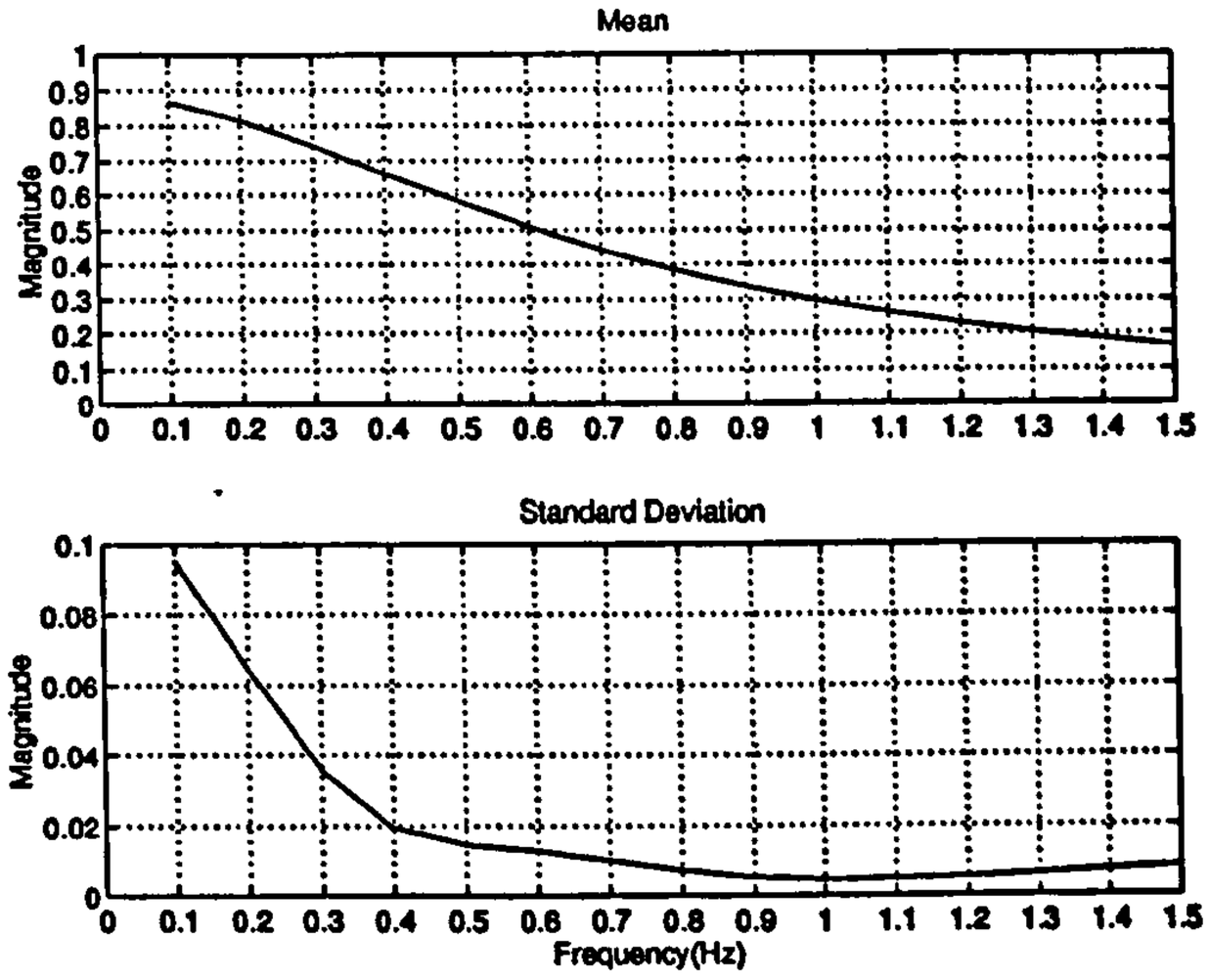

Figure 7.14: Mean and standard deviation of the third-order model frequency response function magnitude, estimated for the hot-air flow device. Closed-loop configuration.

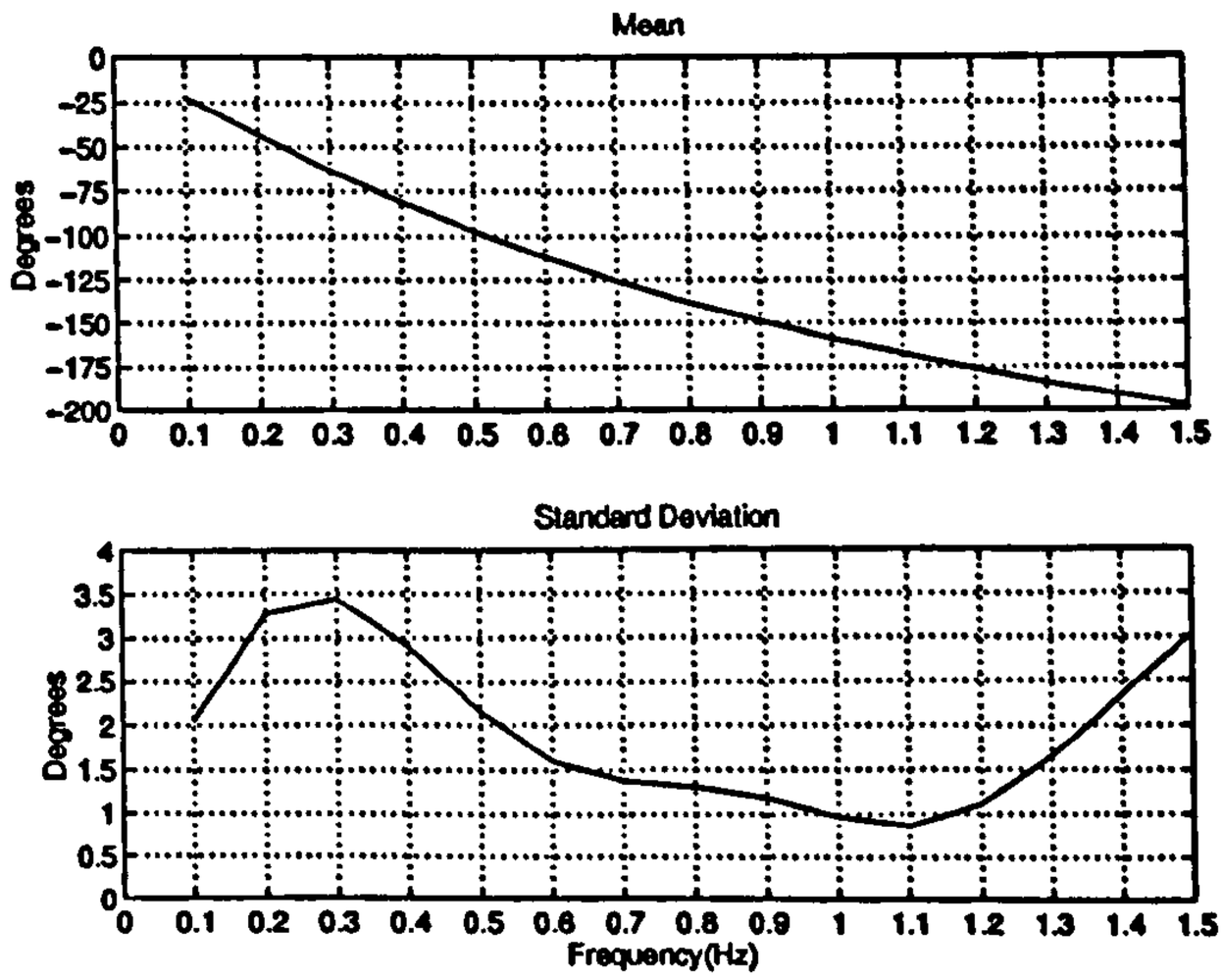

Figure 7.15: Mean and standard deviation of the third-order model frequency response function phase, estimated for the hot-air flow device. Closed-loop configuration.

function estimate. It is only when knowledge of the low-frequency behaviour of the system is required, that the parametric model is required.

For each parametric estimate, controller parameters were calculated using the model frequency response functions. The mean and standard deviations for these are shown in Table 7.2. Although the -180^o frequency of the second-order model

Model	*Rule*	K_p	T_i	T_d	β
1/2	RZ-N	$2.79 \pm 3.48\%$	$0.37 \pm 2.18\%$	$0.094 \pm 2.46\%$	$0.56 \pm 7.76\%$
	ISTE	$2.37 \pm 3.49\%$	$0.56 \pm 10.23\%$	$0.094 \pm 2.46\%$	-
	KLV	$1.67 \pm 2.41\%$	$0.76 \pm 1.18\%$	$0.057 \pm 1.23\%$	-
1/3	RZ-N	$2.74 \pm 2.4\%$	$0.40 \pm 1.65\%$	$0.10 \pm 1.70\%$	$0.58 \pm 6.56\%$
	ISTE	$2.33 \pm 2.46\%$	$0.59 \pm 11.43\%$	$0.1 \pm 1.70\%$	-
	KLV	$1.83 \pm 2.35\%$	$0.75 \pm 1.44\%$	$0.056 \pm 1.43\%$	-

Table 7.2: Controller parameter estimates calculated from the parametric frequency response functions.

has been systematically over-estimated (resulting in the magnitude being under-estimated), the effect on the parameter values is not great. It is clear that the parameters calculated solely from the higher harmonics are well-behaved. The parameters utilising the low-frequency characteristics of the process show a greatly increased variability.

Figures 7.16 and 7.17 show the step response of the process, obtained with each of the rule based controller designs. Firstly, the discrepancy in the controller parameters for the second- and third-order models has not caused any significant change in the performance of the system. The controller design which gives the best (tracking) performance for this system is the KLV method, which is advantageous since this method only requires knowledge of a single frequency response estimate. With this system, the non-parametric identification method would have sufficed.

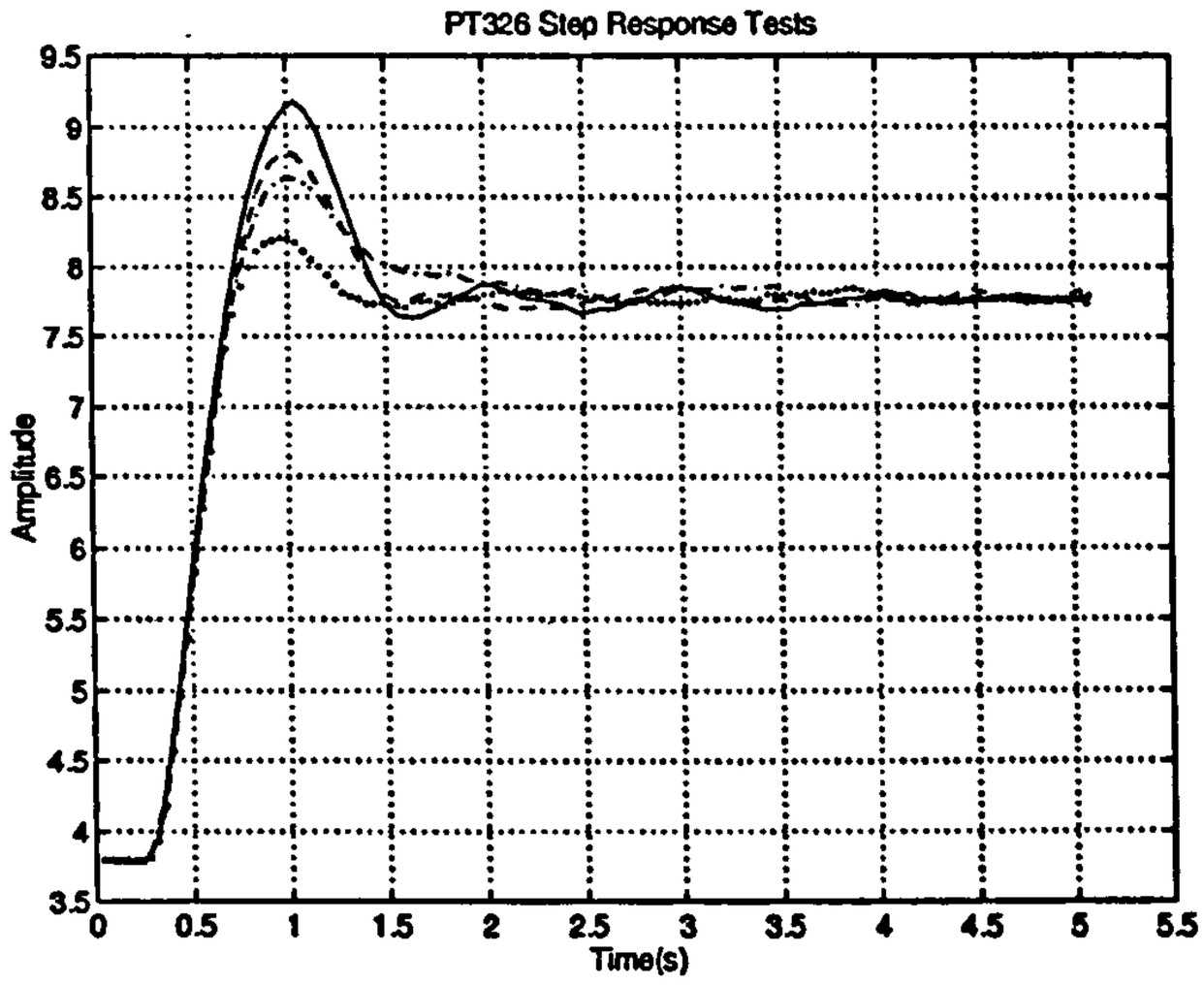

Figure 7.16: Step response tests obtained for the hot-air flow device with rule based controllers designed from second-order model. Ziegler-Nichols (solid), refined Ziegler-Nichols (dashed), ISTE (dashed-dot), KLV (dotted).

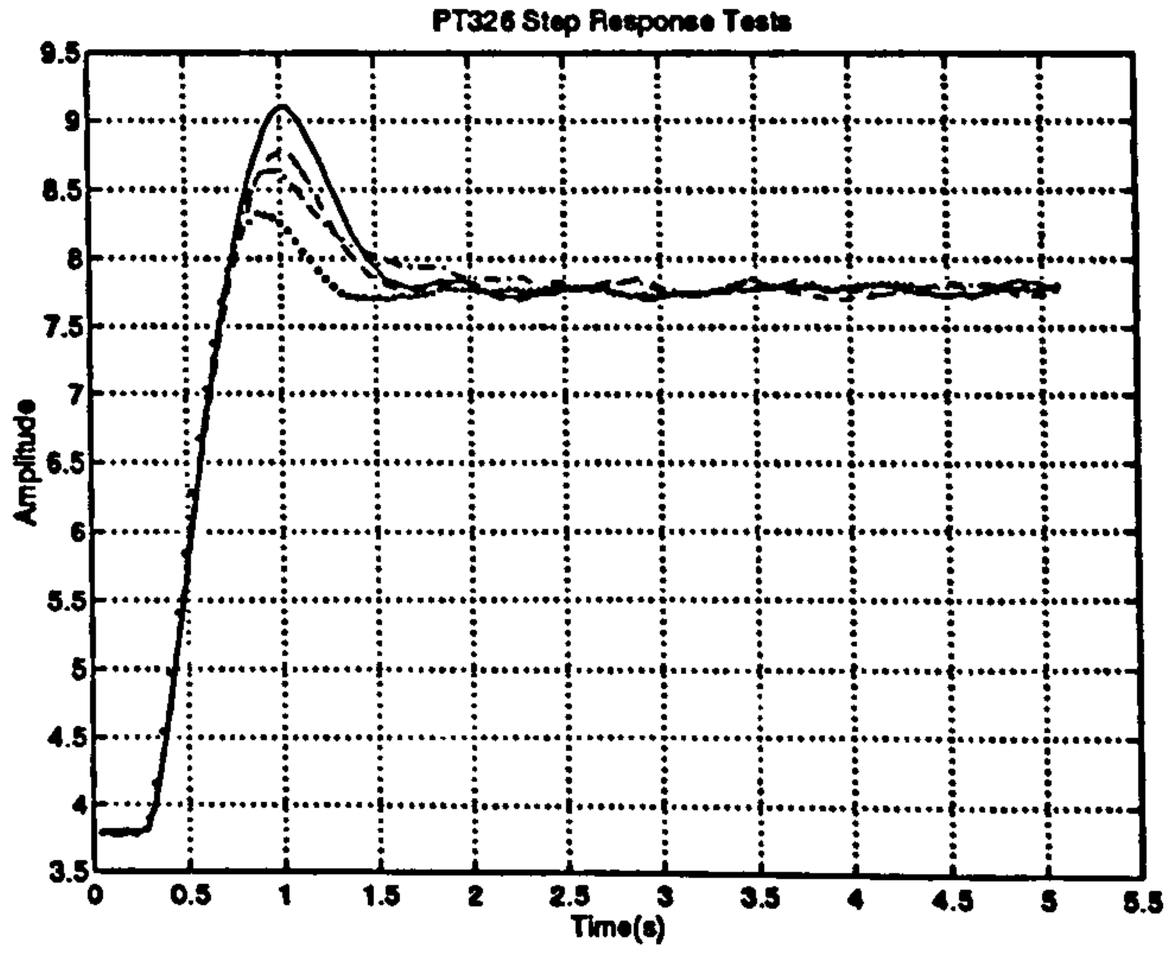

Figure 7.17: Step response tests obtained for the hot-air flow device with rule based controllers designed from third-order model. Ziegler-Nichols (solid), refined Ziegler-Nichols (dashed), ISTE (dashed-dot), KLV (dotted).

7.6 Conclusions

An identification approach to the problem of automatically tuning the controller parameters in a PID structure was discussed. The identification method centred on using a multiharmonic excitation to simultaneously measure several points of the process frequency response, without the need to break the feedback loop or halt the operation of the controller. Conceptually, the method is similar to relay excitation, but it is felt that a small disturbance to the setpoint during tuning is more favourable in practice compared to replacing the controller with a relay element. If one wishes to simply mimic the relay excitation method, a narrow-band signal can be designed to estimate several points of the frequency response, all in close proximity to a previously determined value of the critical frequency. However, a more conventional identification approach has been adopted here.

Non-parametric and parametric identification methods have been discussed. The main advantages of the identification approach can be gained if the former can be used. If this is the case, the properties of the excitation can be directly utilised, and the computations required are negligible. A similar identification approach is elaborated by Hang and Sin (1991), with a PRBS signal used as the setpoint disturbance. The ability to compress the signal power into a low number of harmonics with the multiharmonic signals is the only main difference between the two experimental situations. If there is a significant disturbance within the loop, the non-parametric estimate is expected to perform badly. From the results given in Section 7.5, the non-parametric estimate would have been unusable if the level of disturbance present at the low frequencies were found throughout the frequency range.

Parametric identification provides a way of improving the accuracy of the estimates, but it also increases the complexity of the problem many-fold. First of all comes the problem of structure selection. It must be emphasised that systems

typically encountered in the process industries are usually well characterised by low-order parametric models (invariably non-minimum phase). Therefore structure selection is simplified to some extent. With regard to the particular identification method implemented on the PT326 system, the necessity to have the tuning time as short as possible, meant that numerical problems were encountered when identifying the parameters of a very natural representation of the system. Although no use is made of the actual parameter values (except for frequency response estimation), this is not an ideal situation.

Recursive time-domain identification methods are to be investigated in the future. The main novelty of the approach is unfortunately lost at this point, as the computational complexity involved in structure determination and approximate identification, for the purposes of estimating the frequency response, make the system requirements far from simple. The only way to assess whether an over-complicated identification strategy is being used, is to assess the results obtained from real systems. This will be carried out in the future as part of a project entitled "Rule based and model based tuning of process controllers using frequency domain identification", which is the subject of funding by the EPSRC under Grant GR/K 09373. With the process considered in Section 7.5, the accuracy of the majority of the non-parametric frequency response estimates was sufficient for controller tuning.

The use of the estimated frequency response function played only a small role in this chapter. This said, all the tuning rules known to the author have been tested on a real system. On the PT326 process, the KLV tuning method gave superior tracking performance, which is heartening since the method does not require knowledge of the low-frequency characteristics of the system. However, it would be dangerous to generalise this result.

Further work needs to be carried out on all the topics considered in this chapter, with considerable work necessary on the utilisation of the frequency response

data. It may turn out that the present tuning rules already give the level of control performance which is possible with rule based designs, but this needs to be assessed. A comparison is also needed between the straightforward PID structure and the pseudo-derivative feedback (PDF) structure. The PDF controller has both the proportional and derivative terms of the controller acting on the process output, instead of the error signal. This simple modification has been shown to give modest improvements in the performance of marine systems (Vahedipour, 1990), and so deserves consideration.

Chapter 8

Conclusions and Discussion

8.1 Summary

The main results of this research will now be assessed, and where necessary, indications given to where further work is necessary. Consideration must also be given to current issues in the field of system identification, and to how the work described in this thesis fits into the general framework. The contents of the individual chapters show that the emphasis of the work has been on the practical utilisation of periodic excitations in system identification, with attention given to three main areas:- perturbation signal design; pre-treatment of data; and identification methods for linear and nonlinear systems. Since these areas are arguably the most important areas of system identification practice, the actual significance of the work reported here needs to be carefully evaluated. This will be carried out, in turn, for the areas indicated above.

8.2 Periodic Signal Design

System identification methods generally require that an input signal be *persistently exciting*, while if *a priori* knowledge of the system dynamics exists, it is possible to design an optimal excitation, using well-established techniques. For frequency-domain identification, a further requirement is that the measurements be transformed to the frequency-domain without systematic errors, hence enforcing the use of periodic excitations. What more is there to say about the design of signals for system identification ? Fortunately, quite a lot.

Immediately, one can usually dispense with any notion of designing an optimal excitation since, generally, this requires *a priori* knowledge of the system, of a quality that makes the identification experiment obsolete. Also, for practical identification, the use of a signal that is persistently exciting, and no more, makes the implicit assumption of a known model structure. Hence perturbation signals which

guarantee persistent excitation, and allow the user some flexibility when it comes to discriminating between different model structures, seem to have some merit. However, these problems are relatively well known, and indeed can be circumvented using the pseudo-random or random excitations. The type of excitation considered in this thesis, and typified by the multilevel signals of Chapter 1, are suitable for maximising the input power, in an arbitrary spectrum, for a given time-domain constraint, but is this really necessary, when a pseudo-random signal (established for over thirty years), can fulfil nearly all of these tasks quite well ? Well, for the parametric, linear, single-input identification of a linear system, it would be a mistake to detract from the pseudo-random excitations, and try to justify the use of these "new" periodic signals, on an accuracy basis. However, the systems encountered in practice are often multi-input and rarely linear. Furthermore, the control community is questioning the accuracy of models obtained using conventional system identification methods. The first two points will be considered here, with the last point entering the discussion in Section 8.4.

A good place to start is the non-parametric identification of single-input and multiple-input systems. For linear system modelling, the accuracy of each frequency response estimate can be directly influenced by the selection of the input power spectrum. For multiple-input systems, the design of uncorrelated periodic signals corresponds to specifying orthogonal frequency sets, and so can be accomplished using the signal designs considered in this work. For the parametric identification of multiple-input systems, uncorrelated input signals need not be specified, but the identification is certainly aided by the availability of the frequency response function of each linear sub-system. This is, in essence, similar to the multivariable identification and model selection procedures proposed by Zhu and Backx (1993) and Zhu (1994). The availability of the non-parametric frequency response function when a periodic excitation is used, alleviates the need to initially estimate a high-order

model of the system. Hence for linear system identification, periodic excitations that are capable of realizing an arbitrary power spectrum can be seen to be very desirable. Of course, the availability of powerful time-domain optimisation procedures to carry out the signal design, means that the signals always exhibit a very high power content.

When the system under study is nonlinear, the possible applications of periodic excitations increases. Firstly, if a linear model is required, low-order nonlinear distortion errors can be controlled by proper selection of the properties of the input signal, and the quantitative presence of nonlinear effects can be assessed using signal processing procedures which are non-parametric in nature. If it is the actuator, alone, which is nonlinear, a large selection of input signals is available to guarantee the accurate application of the signal to the dynamics of interest. The multilevel signals discussed in Chapter 1 have significant application in this area. When a nonlinear model is required, it was shown in Chapters 5 and 6 that the flexibility of periodic signals, together with the ability to utilise frequency-domain measurements, can aid nonlinear system modelling. Ignoring the identification procedures for the moment, the need in parametric nonlinear system modelling for input signals with a rich harmonic content, means time-domain optimisation procedures are essential, since the potential to create a waveform with a high crest factor is considerable.

To conclude this section, it can be said that for practical system identification, the consideration given to the properties of periodic input signals in Chapters 1, 5 and 6, shows that periodic input signals have a wide-enough area of applicability to justify their use in most system identification experiments. The multilevel signals prove useful when it is difficult to apply a multisine to the system, and when the properties of binary signals are unsatisfactory. The most important conclusion, though, is that a well-designed periodic signal will give accurate identification results, with excellent facilities for model validation, including the assessment of

nonlinear distortions. The flexibility of being able to choose between different time-domain realizations is an added advantage.

8.3 Pre-Treatment of Data

Leaving aside the discussion of the merits of periodic excitations, attention is now given to signal processing procedures which usually precede the estimation of the system parameters. It is shown in Chapters 3 and 4 that drift and transient effects can be diminished, relatively easily, when a periodic excitation is used for the identification experiment. Looking firstly at the treatment of drift effects, three procedures were proposed to estimate a polynomial approximation of the drift disturbance, that are applicable in different experimental situations. The first method is appropriate in situations where the drift disturbance can be represented by a low-order polynomial, and when it is only possible to excite the system with a low number of periods of the input signal. The second method used a conceptually similar procedure, but assumed that an appropriate number of signal periods can be applied to the system. This leaves a number of low-frequency harmonics, in the output signal, free for approximating the drift. Finally, the third method used a coupled method to simultaneously estimate the drift and spectra of the output signal, and is applicable in all experiments that utilise a periodic input signal. The second and third methods were shown to give approximately the same results in the presence of additional stochastic disturbances. The first method was shown to give acceptable results, when low-frequency harmonics can be left free in the input signal for the purposes of drift removal.

Turning now to the suppression of transient effects, it was shown that signal processing procedures can be used to alleviate the requirement of allowing transient effects to decay before acquiring the periodic mode of response. This is particularly

necessary when frequency-domain identification is used. However, the work was carried out initially, to explore alternatives to using aperiodic excitations and identification methods, which are also appropriate in situations where transient effects are a problem. It was shown in Chapter 4 that there is no need to specify alternative excitations or identification procedures (e.g. aperiodic cross-correlation), to side-step the transient problem. The methods proposed were shown to give accurate results with realistic experimental data. Both methods of transient suppression are compatible with the drift suppression procedures studied in Chapter 3

8.4 Linear and Nonlinear System Identification

Identification procedures were discussed in Chapters 2 and 6 for linear and nonlinear system identification. A control-relevant system identification procedure was also discussed in Chapter 7. The material in Chapter 2 was included to illustrate the use of linear system identification procedures, and show that time-domain and frequency-domain system identification procedures are complementary to each other. The nonlinear system identification procedures, developed in Chapter 6, showed that both non-parametric and parametric nonlinear models may be obtained from frequency-domain measurements. The identification methods employed included techniques which are completely dependent upon specific forms of periodic input signals, as well as parametric methods which use the frequency-domain data in much the same way as linear system identification procedures. In the area separating these two strategies, identification procedures were given for Hammerstein and Weiner models, which impose weak constraints on the type of excitation that may be used for the experiment.

The application of linear system identification to PID controller tuning, considered in Chapter 7, encompasses many of the current issues in control-relevant system

identification. Although the application can be thought of as being simplistic in the light of modern control theory, the coupling of identification and control methods, as is the case here, is the focus of much research by the identification community. Several approaches were given in Chapter 7 for estimating the frequency response of the system under control. Signal design issues were also dealt with.

A great deal of further research is required in this area, to create both a robust identification method, and tuning rules which exploit the identification results. It is currently believed that the necessary identification procedures will be parametric, and so the recent work on model error quantification (Goodwin et al., 1992; Schoukens and Pintelon, 1994b), is of much interest. The identification method is, however, only half of the picture. With the availability of a large number of frequency response estimates, research is required to assess whether the tuning rules, available at the present time, can be improved upon, and of equal importance, whether different controller structures (albeit still of the PID type), can offer improvements.

8.5 Concluding Remarks

Signal design procedures were developed for the identification of linear and nonlinear systems. Signal processing procedures were given for the suppression of drift and transient effects from the identification results, and for the assessment of nonlinear contributions to the measurements. Frequency-domain identification techniques were given for a class of nonlinear systems. Finally, a new autotune method was described, which improves upon many of the features of current methods.

Bibliography

Advantest Corporation (1989). R9211B/C FFT servo analyzer instruction manual, Tokyo.

Andria, G., Savino, M. and Trotto, A. (1989). Windows and interpolation algorithms to improve electrical measurement accuracy, *IEEE Transactions on Intrumentation and Measurement* **38**(4): 856–863.

Åström, K. and Hägglund, T. (1988). *Automatic Tuning of PID Controllers*, Instrument Society of America, Research Triangle Park, NC 27709.

Åström, K. J. (1980). Maximum likelihood and prediction error methods, *Automatica* **16**: 551–574.

Åström, K. J. and Hägglund, T. (1984). Automatic tuning of simple regulators with specifications on phase and amplitude margins, *Automatica* **20**(5): 645–651.

Åström, K. J., Hägglund, T., Hang, C. C. and Ho, W. K. (1993). Automatic tuning and adaptation for PID controllers - A survey, *Control Engineering Practice* **1**(4): 699–714.

Barker, H. A. (1967). Choice of pseudorandom binary signals for system identification, *Electronics Letters* **3**(11): 524–526.

Barker, H. A. (1969). Reference phase of pseudorandom signals, *Proceedings IEE* **116**(3): 429–435.

Barker, H. A. (1993). Design of multi-level pseudo-random signals for system identification, *in* K. Godfrey (ed.), *Perturbation Signals for System Identification*, Prentice Hall, U.K., chapter 11.

Barker, H. A. and Al Hilal, M. H. (1985). Nonlinear system identification using pseudorandom signals with partially orthogonal transforms, *7th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, York, pp. 415–420.

Barker, H. A. and Davy, R. W. (1975). System identification using pseudorandom signals and the discrete Fourier transform, *Proceedings IEE* **122**(3): 305–311.

Bayard, D. S. (1994). High-order multivariable transfer function curve fitting: Algorithms, sparse matrix methods and experimental results, *Automatica* **30**(9): 1439-1444.

Bendat, J. and Piersol, A. (1980). *Engineering Applications of Correlation and Spectral Analysis*, John Wiley & Sons, New York.

Bendat, J. S. (1976). Solutions for the multiple input/output problem, *Journal of Sound and Vibration* **44**(3): 311–325.

Bendat, J. S. (1990). *Nonlinear System Analysis and Identification from Random Data*, John Wiley & Sons, New York.

Billings, S. A. (1985). Introduction to nonlinear systems analysis and identification, *in* K. Godfrey and P. Jones (eds), *Signal Processing for Control*, Springer-Verlag, Berlin, chapter L10.

Billings, S. A. and Fakhouri, S. Y. (1977). Identification of nonlinear systems using the Wiener model, *Electronics Letters* **13**(17): 502–504.

Billings, S. A. and Fakhouri, S. Y. (1979). Nonlinear system identification using the Hammerstein model, *International Journal of Systems Science* **10**(5): 567–578.

Billings, S. A. and Voon, W. S. F. (1986). A prediction-error and stepwise-regression estimation algorithm for non-linear systems, *International Journal of Control* **44**(3): 803–822.

Billings, S. A. and Zhu, Q. M. (1991). Rational model identification using an extended least squares algorithm, *International Journal of Control* **54**: 529–546.

Billings, S. A. and Zhu, Q. M. (1994). A structure detection algorithm for nonlinear dynamic rational models, *International Journal of Control* **59**(6): 1439–1463.

Boyd, S., Tang, Y. S. and Chua, L. (1983). Measuring Volterra kernels, *IEEE Transactions on Circuits and Systems* **30**(8): 571–777.

Briggs, P. A. N. and Godfrey, K. R. (1966). Pseudorandom signals for the dynamic analysis of multivariable systems, *Proceedings IEE* **113**(7): 1259–1267.

Brown, R. F. (1968). Drift correction in periodic crosscorrelation schemes, *Electronics Letters* 4(22): 478–479.

Buckner, M. R. and Kerlin, T. W. (1972). Optimum binary signals for reactor frequency response measurments, *Nuclear Science and Engineering* **49**: 255–262.

Chen, C. H., Kerlin, T. W. and Fry, D. N. (1972). Experiences with binary periodic signals for dynamic testing at the HFIR, *IEEE Transactions on Nuclear Science* **19**: 828–836.

Chen, C. T. (1993). The use of non-periodic signals in system identification, *B(Eeng) in Engineering Electronics Final Year Project Report*, Department of Engineering, University of Warwick.

Chen, S., Billings, S. A. and Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification, *International Journal of Control* 50(5): 1873–1896.

Darnell, M. (1993). Periodic and non-periodic, binary and multi-level pseudo-random signals, *in* K. Godfrey (ed.), *Perturbation Signals for System Identification*, Prentice Hall, U.K., chapter 5.

Darnell, M. and Kemp, A. H. (1988). Synthesis of multilevel complementary sequences, *Electronics Letters* 24(19): 1251–1252.

Davies, W. D. T. and Douce, J. L. (1967). On-line system identification in the presence of drift, *IFAC Symposium on Identification in Automatic Control Systems*, Prague. Preprints part 1, paper 3.12.

Ditmar, W. and Pettitt, R. (1993). Multi-frequency signals for plant identification, *in* K. Godfrey (ed.), *Perturbation Signals for System Identification*, Prentice Hall, U.K., chapter 13.

Evans, C., Rees, D., Jones, L. and Hill, D. (1994a). Time and frequency domain identification of jet engine dynamics: Problems and solutions, *10th IFAC Symposium on System Identification*, Copenhagen, pp. 2.243–2.248.

Evans, D. C., Rees, D. and Jones, D. L. (1992). Design of test signals for identification of linear systems with nonlinear distortions, *IEEE Transactions on Intrumentation and Measurement* 41(6): 768–774.

Evans, D. C., Rees, D. and Jones, D. L. (1994b). Identifying linear models of systems suffering nonlinear distortions, *IEE International Conference "Control 94"*, University of Warwick, pp. 288–296.

Eykhoff, P. (1974). *System Identification: Parameter and State Estimation*, John Wiley & Sons, New York.

Flower, J. O., Forge, S. C., Radcliffe, N. G. and Roust, C. B. (1978). Dynamic measurement of a nuclear reactor using low-peak-factor excitation signals, *Nuclear Science and Engineering* 68: 110–115.

Franck, G. and Rake, H. (1985). Identification of a large water heated crossflow heat exchanger with binary multifrequency signals, *7th IFAC Symposium on Identification and System Parameter Estimation*, York, pp. 1859–1864.

Gevers, M. (1991). Connecting identification and robust control: A new challenge, *9th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Budapest, pp. 1.1–1.10.

Godfrey, K. (1993a). Introduction to perturbation signals for frequency-domain system identification, *in* K. Godfrey (ed.), *Perturbation Signals for System Identification*, Prentice Hall, U.K., chapter 2.

Godfrey, K. (1993b). Introduction to perturbation signals for time-domain system identification, *in* K. Godfrey (ed.), *Perturbation Signals for System Identification*, Prentice Hall, U.K., chapter 1.

Godfrey, K. (ed.) (1993c). *Perturbation Signals for System Identification*, Prentice Hall, UK.

Godfrey, K. R. and Briggs, P. A. N. (1972). Identification of processes with direction-dependent dynamic responses, *Proceedings IEE* **119**(12): 1733–1739.

Godfrey, K. R. and Moore, D. J. (1974). Identification of processes having direction-dependent responses, with gas-turbine engine applications, *Automatica* **10**: 469–481.

Godfrey, K. R., McCormack, A. S. and Flower, J. O. (1995). Applying system identification using commercially available software and hardware, *Control Engineering Practice*. Accepted for Publication.

Goodwin, G. and Payne, R. (1977). *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press, New York.

Goodwin, G. C., Gevers, M. and Ninness, B. (1992). Quantifying the error in identified transfer functions with applications to model order selection, *IEEE Transactions on Automatic Control* **37**(7): 913–928.

Grace, A. (1992). *Optimization Toolbox for use with MATLAB*, The MathWorks Inc., South Natick, Mass.

Grandke, T. (1983). Interpolation algorithms for discrete Fourier transforms of weighted signals, *IEEE Transactions on Intrumentation and Measurement* **32**(2): 350–355.

Greblicki, W. and Pawlak, M. (1986). Identification of discrete Hammerstein systems using kernel regression estimates, *IEEE Transactions on Automatic Control* **31**(1): 74–77.

Guillaume, P., Pintelon, R. and Schoukens, J. (1992). Parametric identification of two-port models in the frequency domain, *IEEE Transactions on Intrumentation and Measurement* **41**(2): 233–239.

Guillaume, P., Schoukens, J., Pintelon, R. and Kollár, I. (1991). Crest-factor minimization using nonlinear Chebyshev approximation methods, *IEEE Transac-*

tions on Intrumentation and Measurement 40(6): 982–989.

Hägglund, T. and Åström, K. J. (1991). Identification of systems using periodic excitations, *9th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Budapest, pp. 2.1157–2.1162.

Hang, C. C. and Sin, K. K. (1991). On-line auto tuning of PID controllers based on the cross-correlation technique, *IEEE Transactions on Industrial Electronics* 38(6): 428–437.

Hang, C. C., Åström, K. J. and Ho, W. (1991). Refinements of the Ziegler-Nichols tuning formula, *IEE Proceedings Part D* 138(2): 111–118.

Harris, S. L. (1993). Generation and applications of binary multi-frequency signals, *in* K. Godfrey (ed.), *Perturbation Signals for System Identification*, Prentice Hall, U.K., chapter 6.

Hewlett Packard Co. (1989). HP 3563A operating manual, U.S.A.

Hunter, I. W. and Korenberg, M. J. (1986). The identification of nonlinear biological systems: Wiener and Hammerstein cascade models, *Biological Cybernetics* **55**: 135–144.

Institution of Electrical Engineers (1994). Improvements in furnace control: Current developments and new technologies, *IEE Colloquium, Digest 1994/018.*

Jain, V. K., Collins, W. L. and Davis, D. C. (1979). High-accuracy analog measurements via interpolated FFT, *IEEE Transactions on Intrumentation and Measurement* 28(2): 113–122.

Kollár, I. (1994). *Frequency Domain System Identification Toolbox for use with MATLAB*, The MathWorks Inc., South Natick, Mass.

Kollár, I., Pintelon, R. and Schoukens, J. (1994). Frequency domain system identification toolbox for MATLAB: A complex application example, *10th IFAC Symposium on System Identification*, Copenhagen, pp. 4.23–4.28.

Kollár, I., Pintelon, R., Rolain, Y. and Schoukens, J. (1991). Another step towards an ideal data acquisition channel, *IEEE Transactions on Intrumentation and Measurement* **40**(3): 659–660.

Korenberg, M., Billings, S. A., Liu, Y. P. and McIlroy, P. J. (1988). Orthogonal parameter estimation algorithm for non-linear stochastic systems, *International Journal of Control* **48**(1): 193–210.

Kvashnin, Y. F. (1969). Measurement of the correlation function and spectrum of a nonperiodic m-sequence, *Telecommunications* **23**(2): 56–58.

Lawrence, P. J. (1981). Estimation of Volterra functional series of a nonlinear system using frequency response data, *IEE Proceedings Part D* **128**: 206–210.

Lawrence, P. J. and Rogers, G. J. (1979). Sequential transfer-function synthesis from measured data, *Proceedings IEE* **126**(1): 104–106.

Leontaritis, I. J. and Billings, S. A. (1985). Input-output parametric models for non-linear systems part II: Stochastic non-linear systems, *International Journal of Control* **41**(2): 329–344.

Leontaritis, I. J. and Billings, S. A. (1987). Experimental design and identifiability for non-linear systems, *International Journal of Systems Science* **18**(1): 189–202.

Levi, E. C. (1959). Complex curve fitting, *IRE Transactions on Automatic Control* **4**: 37–43.

Ljung, L. (1985). On the estimation of transfer functions, *Automatica* **21**(6): 677–696.

Ljung, L. (1986). *System Identification Toolbox for use with MATLAB*, The MathWorks Inc., South Natick, Mass.

Ljung, L. (1987). *System Identification: Theory for the User*, Prentice Hall Inc., Englewood Cliffs, NJ 07632.

Ljung, L. and Söderström, T. (1983). *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, Mass.

McCormack, A., Flower, J. and Godfrey, K. (1994a). The suppression of drift and transient effects for frequency-domain identification, *IEEE Transactions on Intrumentation and Measurement* **43**(2): 232–237.

McCormack, A. S., Godfrey, K. R. and Flower, J. O. (1994b). The detection of and compensation for nonlinear effects using periodic input signals, *IEE International Conference "Control 94"*, University of Warwick, pp. 297–302.

McCormack, A. S., Godfrey, K. R. and Flower, J. O. (1995). The design of multilevel multiharmonic signals for system identification, *IEE Proceedings Part D.* Accepted for Publication.

Norton, J. (1986). *An Introduction to Identification*, Academic Press Inc., London.

Nowak, R. D. and Van Veen, B. D. (1994). Random and pseudorandom inputs for Volterra filter identification, *IEEE Transactions on Signal Processing* **42**(8): 2124–2135.

Offelli, C. and Petri, D. (1990). Interpolation techniques for real-time multifrequency waveform analysis, *IEEE Transactions on Intrumentation and Measurement* **39**(1): 106–111.

Patra, A. (1993). Identification of a class of nonlinear continuous-time systems using Hartley modulating functions, *Technical report*, Ruhr-Universität, Bochum, Germany.

Pintelon, R., Guillaume, P., Rolain, Y. and Verbeyst, F. (1992). Identification of linear systems captured in a feedback loop, *IEEE Transactions on Intrumentation and Measurement* **41**(6): 747–754.

Pintelon, R., Guillaume, P., Rolain, Y., Schoukens, J. and Van hamme, H. (1994a). Parametric identification of transfer functions in the frequency domain – a survey, *IEEE Transactions on Automatic Control* **39**(11): 2245–2260.

Pintelon, R., Rolain, Y., Vanden Bossche, M. and Schoukens, J. (1990). Towards an ideal data acquisition channel, *IEEE Transactions on Intrumentation and Measurement* **39**(1): 116–120.

Pintelon, R., Schoukens, J. and Chen, H. (1994b). On the basic assumptions in the identification of continuous time systems, *10th IFAC Symposium on System Identification*, Copenhagen, pp. 3.143–3.152.

Rabiner, L. R., Schafer, R. W. and Rader, C. M. (1969). The chirp z-transform algorithm and its application, *The Bell System Technical Journal* **48**: 1249–1293.

Ream, N. (1968). Proof of the drift-resistant property of binary m-sequences, *Electronics Letters* **4**(18): 380–381.

Renders, H., Schoukens, J. and Vilain, G. (1984). High-accuracy spectrum analysis of sampled discrete frequency signals by analytical leakage compensation, *IEEE Transactions on Intrumentation and Measurement* **33**(4): 287–292.

Sanathan, C. K. and Koerner, J. (1963). Transfer function synthesis as a ratio of two complex polynomials, *IEEE Transactions on Automatic Control* 8: 56–58.

Schei, T. S. (1994). Automatic tuning of PID controllers based on transfer function estimation, *Automatica* **30**(12): 1983–1989.

Schoukens, J. and Pintelon, R. (1991). *Identification of Linear Systems: A Practical Guideline to Accurate Modeling*, Pergamon Press, London.

Schoukens, J. and Pintelon, R. (1994a). Identification of linear systems in the presence of model errors, *IEEE Instrumentation and Measurement Technology Conference IMTC/94*, Hamamatsu, pp. 1372–1376.

Schoukens, J. and Pintelon, R. (1994b). Quantifying model errors of identified transfer functions, *IEEE Transactions on Automatic Control* **39**(8): 1733–1737.

Schoukens, J. and Renneboog, J. (1986). Modeling the noise influence on the Fourier coefficients after a discrete Fourier transform, *IEEE Transactions on Intrumentation and Measurement* **35**(3): 278–286.

Schoukens, J., Guillaume, P. and Pintelon, R. (1993a). Design of broadband excitation signals, *in* K. Godfrey (ed.), *Perturbation Signals for System Identification*, Prentice Hall, U.K., chapter 3.

Schoukens, J., Pintelon, R. and Renneboog, J. (1988a). A maximum likelihood estimator for linear and nonlinear systems – a practical application of estimation techniques in measurement problems, *IEEE Transactions on Intrumentation and Measurement* **37**(1): 10–17.

Schoukens, J., Pintelon, R. and Van hamme, H. (1992). The interpolated fast Fourier transform: A comparitive study, *IEEE Transactions on Intrumentation and Measurement* **41**(2): 226–232.

Schoukens, J., Pintelon, R. and Van hamme, H. (1994). Identification of linear dynamic systems using piecewise constant excitations: Use, misuse and alternatives, *Automatica* **30**(7): 1153–1169.

Schoukens, J., Pintelon, R., Van der Ouderaa, E. and Renneboog, J. (1988b). Survey of excitation signals for FFT based signal analyzers, *IEEE Transactions on Intrumentation and Measurement* **37**(3): 342–352.

Schoukens, J., Rolain, Y., Montecelli, L. and De Locht, C. (1993b). Identification of linear systems in the presence of non-linear distortions, *Proc. 11th International Modal Analysis Conference*, Florida, pp. 479–485.

Schrama, R. J. P. (1991). An open-loop solution to the approximate closed-loop identification problem, *9th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, Budapest, pp. 2.1602–2.1607.

Schroeder, M. R. (1970). Synthesis of low-peak-factor signals and binary sequences with low autocorrelation, *IEEE Transactions on Information Theory* **16**: 85–89.

Seifer, W. W. and Steeg, C. W. (1960). *Control Systems Engineering*, McGraw-Hill.

Sinha, N. and Rao, G. (eds) (1991). *Identification of Continuous-Time Systems: Methodology and Computer Implementation*, Kluwer Academic Publishers, Dordrecht.

Söderström, T. and Stoica, P. (1989). *System Identification*, Prentice Hall International Ltd, UK.

Suki, B. and Lutchen, K. R. (1992). Pseudorandom signals to estimate apparent transfer and coherence functions of nonlinear systems: Applications to respiratory mechanics, *IEEE Transactions on Biomedical Engineering* **39**(11): 1142–1151.

Tsang, K. M. and Billings, S. A. (1994). Identification of continuous time nonlinear systems using delayed state variable filters, *International Journal of Control* **60**(2): 159–180.

Vahedipour, A. (1990). *Application of Pseudo-Derivative Feedback (PDF) Algorithm in Ship Control*, PhD thesis, University of Exeter.

Van den Bos, A. (1970). Estimation of linear system coefficients from noisy responses to binary multifrequency test signals, *2nd IFAC Symposium on Identification and Process Parameter Estimation*, Prague. paper 7.2.

Van den Bos, A. and Krol, R. G. (1979). Synthesis of discrete-interval binary signals with specified Fourier amplitude spectra, *International Journal of Control* **30**(5): 871–884.

Van den Eijnde. E. and Schoukens, J. (1991). On the design of optimal excitation signals, *9th IFAC Symposium on Identification and System Parameter Estimation*, Budapest, pp. 1.827–1.832.

Van den Hof, P. M. J. and Scrama, R. J. P. (1994). Identification and control: Closed loop issues, *10th IFAC Symposium on System Identification*, Copenhagen, pp. 2.1–2.13.

Van der Ouderaa, E., Schoukens, J. and Renneboog, J. (1988a). Peak factor minimization using a time-frequency domain swapping algorithm, *IEEE Transactions on Intrumentation and Measurement* **37**(1): 145–147.

Van der Ouderaa, E., Schoukens, J. and Renneboog, J. (1988b). Peak factor minimization of input and output signals of linear systems, *IEEE Transactions on Intrumentation and Measurement* **37**(2): 207–212.

Van hamme, H. (1992). *Identification of Linear Systems from Time- or Frequency Domain Measurements*, PhD thesis, Vrije Universiteit Brussel.

Victor, J. and Shapely, R. (1980). A method of nonlinear analysis in the frequency domain, *Biophysical Journal* **29**: 459–484.

Voda, A. and Landau, I. D. (1995). A method for the auto-calibration of PID controllers, *Automatica* **31**(1): 41–53.

Wellstead, P. E. (1981). Non-parametric methods of system identification, *Automatica* **17**(1): 55–69.

Whitfield, A. H. (1986). Transfer function synthesis using frequency response data, *International Journal of Control* **43**(5): 1413–1426.

Young, P. (1981). Parameter estimation for continuous-time models – a survey, *Automatica* **17**(1): 23–39.

Zhu, Y. (1994). A frequency domain criterion for MIMO model order/structure selection, *10th IFAC Symposium on System Identification*, Copenhagen, pp. 2.516–2.521.

Zhu, Y. and Backx, A. (1993). *Identification of Multivariable Industrial Processes: for Simulation, Diagnosis and Control*, Springer-Verlag, London.

Zhuang, M. and Atherton, D. P. (1993). Automatic tuning of optimum PID controllers, *IEE Proceedings Part D* **140**(3): 216–224.

Ziegler, J. G. and Nichols, N. B. (1943). Optimum settings for automatic controllers, *Trans. ASME* **65**: 433–444.

Van hamme, H. (1992). *Identification of Linear Systems from Time- or Frequency Domain Measurements*, PhD thesis, Vrije Universiteit Brussel.

Victor, J. and Shapely, R. (1980). A method of nonlinear analysis in the frequency domain, *Biophysical Journal* **29**: 459–484.

Voda, A. and Landau, I. D. (1995). A method for the auto-calibration of PID controllers, *Automatica* **31**(1): 41–53.

Wellstead, P. E. (1981). Non-parametric methods of system identification, *Automatica* **17**(1): 55–69.

Whitfield, A. H. (1986). Transfer function synthesis using frequency response data, *International Journal of Control* **43**(5): 1413–1426.

Young, P. (1981). Parameter estimation for continuous-time models – a survey, *Automatica* **17**(1): 23–39.

Zhu, Y. (1994). A frequency domain criterion for MIMO model order/structure selection, *10th IFAC Symposium on System Identification*, Copenhagen, pp. 2.516–2.521.

Zhu, Y. and Backx, A. (1993). *Identification of Multivariable Industrial Processes: for Simulation, Diagnosis and Control*, Springer-Verlag, London.

Zhuang, M. and Atherton, D. P. (1993). Automatic tuning of optimum PID controllers, *IEE Proceedings Part D* **140**(3): 216–224.

Ziegler, J. G. and Nichols, N. B. (1943). Optimum settings for automatic controllers, *Trans. ASME* **65**: 433–444.