SYSID'94: 10th IFAC Symposium on System Identification, Copenhagen, 4–6 July, 1994.
Preprints, Vol. 4, pp. 23-28.

FREQUENCY DOMAIN SYSTEM IDENTIFICATION TOOLBOX FOR MATLAB: A COMPLEX APPLICATION EXAMPLE

I. KOLLÁR*, R. PINTELON and J. SCHOUKENS****

**Stanford University, Dept. EE, ISL, Stanford, CA 94305-4055, USA.*
Fax: + 1 415 723-8473, e-mail: kollar@isl.stanford.edu
On leave from: Technical University of Budapest, Dept. Instrument & Measurement Eng, H-1521 Budapest, Hungary, e-mail: kollar@mmt.bme.hu

***Vrije Universiteit Brussel, Dept. ELEC, Pleinlaan 2, B-1050 Brussels, Belgium.*

Abstract. Frequency domain system identification has been rapidly developing during the past years. The new comprehensive Matlab toolbox is a successful implementation of the methods of the whole identification procedure. Its core is a maximum likelihood estimate of parametric transfer function models, maybe with a delay, but also offers help for nonparametric transfer function identification. It contains tools for excitation signal design, data preprocessing, parameter estimation, model and error presentation, model validation, simulation, archivation and documentation.
This paper illustrates the use of many functions of the toolbox on the example of the identification of a model of a flexible robot arm.

Key Words. System identification; frequency domain; linear systems; parameter estimation; maximum likelihood; transfer functions; excitation signal; multisine; computer software; software tool; Matlab toolbox.

1. INTRODUCTION

The Frequency Domain System Identification Toolbox (FDIDENT) was first presented at IFAC Sysid'91 (Kollár et al, 1991). Since then, the methods and the routines were further developed and tested, and numerical stability, speed and graphics have been significantly improved. By now, a reliable set of routines offers a wide spectrum of services for all the important steps of system identification, as:

- excitation signal design,
- data preprocessing,
- parameter estimation,
- model and error presentation,
- model validation,
- simulation,
- archivation and documentation.

Time domain and frequency domain methods are complementary to each other; while both of them can be equally applied to obtain a good approximate system model under many circumstances, there are situations where one of them is superior. This justifies the existence of two identification toolboxes for MATLAB. The discussion of the similarities and differences deserves more space that is allowed for this paper, so only the major issues will be briefly mentioned, and then an elaborated example of frequency domain identification will be presented. For more on the comparison of the two methods, see e. g. (Andersson et al, 1994; Ljung, 1993; Schoukens and Pintelon, 1991; Schoukens et al, 1994).

Time domain methods are usually preferable when

- the system is slightly time varying,
- the excitation signal is random, and cannot be made periodic,
- a discrete-time (z-domain) model is sought,
- on-line (maybe recursive) identification is desirable.

In order to properly apply them it is required that

- the excitation signal is not seriously distorted by the actuator,
- a zero-order hold excitation signal is applied,
- the signal samples are obtained by significant oversampling.

Frequency domain methods are usually preferable when

- the excitation signal is distorted by the actuator,
- a physical (s-domain) model of the system is

sought,

- the signal samples are obtained just above the Nyquist rate,
- wide-band modeling is required,
- frequency bands of interest are to be selected.

In order to properly apply them it is required that

- the system under test is essentially time invariant,
- periodic excitation signal can be applied,
- block update of the system model is acceptable.

2. A CASE STUDY

This section illustrates the frequency domain identification procedure on the example of identification of a flexible robot arm[1]. The procedures described below include a few typical steps, applicable under a variety of circumstances. However, every identification task has its own flavor. For systems, other than this robot arm, some other procedures may also be adequate. Therefore, while this session serves well as an example, and suggests a series of good solutions, making use of the services of the FDIDENT toolbox, it is certainly not the only way of doing identification.

2.1. Description of the Experiment

The behavior of a flexible robot arm was measured by applying controlled torque to the vertical axis at one end of the arm, and measuring the tangential acceleration of the other end.

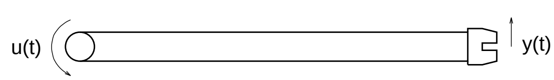

Fig. 1. Measurement of robot arm

The excitation signal was a multisine. Its components had equal amplitudes at frequencies [1,3,. . . , 199]**df*, with $df = fs/N = 500/4096 = 0.125$ Hz, that is, the frequency range of excitation was about 0.125 Hz to 25 Hz. The phases of the individual components were optimized to assure a low crest factor, in order to decrease the effect of nonlinearities. The odd harmonic frequencies provided that components produced by a squaring nonlinearity would not disturb the identification.

The input and output signals were sampled with sampling frequency $fs = 500$ Hz. Sampling was synchronized to the excitation signal so that exactly 4096 samples were taken from each period. The data records contain 40960 points, that is, 10 periods were measured.

2.2. Investigation of the Time Domain Data

First the time domain data and the autocorrelation function can be investigated. Let us have a look at the time domain data.

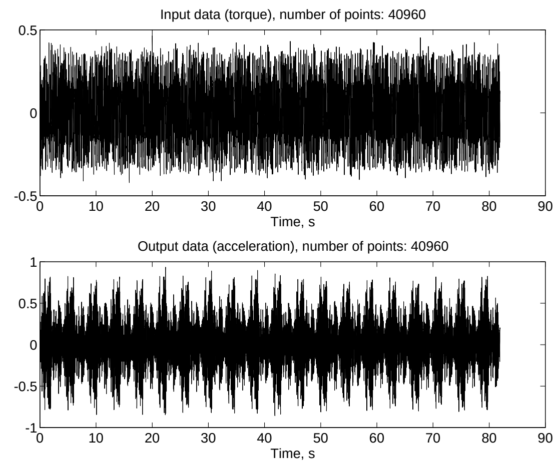

Fig. 2. Time domain data

Not much can be stated on basis of the time functions, not even the period length can be read off. The reason is that the 100 harmonic components make the signal shape very complex. What can be seen, however, is that the crest factor is somewhat worse than that of usual optimized signals: the actuator was not able to follow the desired waveform well enough. So, it was quite reasonable to measure the torque directly, and the input-output noise model of frequency domain identification will be of good use.

The input signal has apparently a smaller crest factor than the output one, which can be expected since the system dynamics strongly modify the amplitudes and phases of the multisine.

More can be determined from the autocovariance function. Figure 3 shows the so-called circular autocovariance. First of all, there is indeed a periodicity of $4096 \cdot dt = 8.192\,\mathrm{s}$. The autocovariance function corresponds to a bandlimited white spectrum, the negative peaks correspond to the use of odd harmonics. The signal was oversampled by a factor of $2048/199 \approx 10$, that is, from the $\sin(x)/x$ shaped main lobes of the autocovariance function about 20 points are sampled.

It is obvious from the enlarged peaks that synchronization is very good between the excitation signal and the sampling clock. This can be also verified in the frequency domain, by estimation

[1] This measurement was made at the Department of Mechanical Engineering, Catholic University of Leuven (KUL), in cooperation with Department ELEC, Vrije Universiteit Brussel (VUB), as a part of the Belgian program "Interuniversity Attraction Poles (IUAP50: Robotics and Industrial Automation)" initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. These data belong to the public domain and can be freely used by anyone.

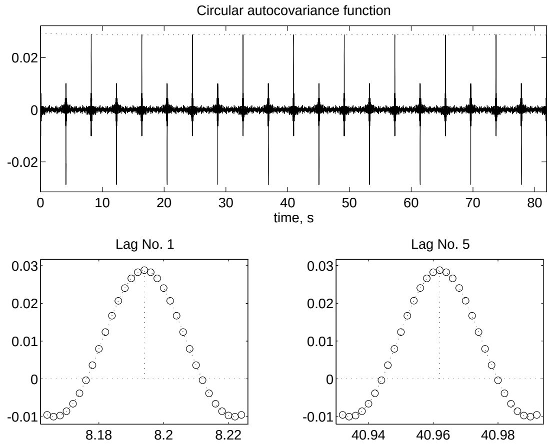

Fig. 3. The circular autocovariance function

the time delay between experiments. If there was a slip, it could not be more than about 0.003% ($dt/2$ in a time of $4{\cdot}4096{\cdot}dt$).

2.3. *Examination of the Signal-to-Noise Ratios*

The autocovariance function can also be used for approximate determination of the signal-to-noise ratio: the power of the periodic components is approximately given by one of the peaks at the nonzero lags, the total power is given by the covariance value at zero.

$C_x(0) = 0.0293,\ \ C_x(k * T_p) = 0.0288,$
$\mathrm{SNR_{inp}} = 17.3\,\mathrm{dB}.$

However, this not exactly what we need: the useful signal has power at the given frequencies only, the rest is spurious components, produced mostly by the nonlinearities. This can be seen from the plot of the Fourier amplitudes (Fig. 4).

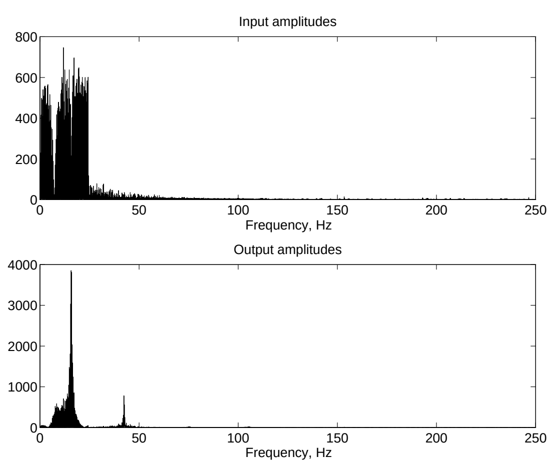

Fig. 4. Input and output Fourier amplitudes

The input amplitude spectrum is not really flat. The cause is most probably the non-flat transfer function of the system composed of the actuator and the device under test. The two dips at 7.2 Hz and 15.7 Hz in the input spectrum correspond probably two resonance points of the system, where the actuator (an electrical motor) is not capable to maintain the signal level. This is not a serious problem since the frequency range of interest is sufficiently covered by nonzero excitation amplitudes.

The powers of the deterministic components, of useful signal, and of the noise can be calculated for both the input and the output signals.

Input signal
total power: 0.0293
useful power: 0.0286, noise power: 0.000508
spurious periodic components: 0.000218
SNR: 17.5 dB, for useful components: 16.0 dB

Output signal
total power: 0.0828
useful power: 0.0797, noise power: 0.00139
spurious periodic components: 0.00171
SNR: 17.7 dB, for useful components: 14.1 dB

The signal-to noise ratios are quite good, although the nonlinearities produce significant components. Since the noise is larger than the nonlinearity products, these will hopefully not deteriorate the identification significantly. Let us notice moreover that we are going to select certain excitation lines only, so the SNR will be improved in the frequency domain by a factor of about 20 (13 dB).

2.4. *Preprocessing of Data*

Now variance analysis follows. 10 periods were measured. We will treat each period as a separate experiment, thus an approximate noise analysis can be performed.

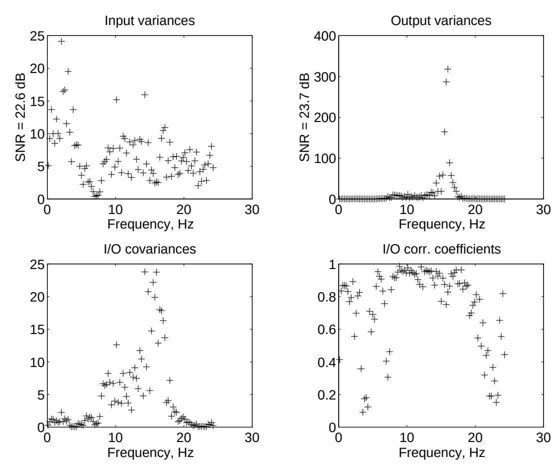

Fig. 5. The results of noise analysis

The total SNR is increased by the selection of the points of interest:

$\mathrm{SNR_{in}} = 22.6\,\mathrm{dB},\ \mathrm{SNR_{out}} = 23.7\,\mathrm{dB}.$

The increase of about 5-6 dB is due to the oversampling and selection of points of interest only. It is less than expected, probably because the noise has less power at higher frequencies than in the lower frequency band, so a large part of it contaminates the lines of interest.

The input-output covariances are quite large, so they may not be neglected. This may mean that a part of the noise goes through the system, that is, the estimation itself is corrupted by less noise than calculated above. This can be verified by plotting the frequency domain input-output covariances, divided by the input variances: if an important part of the noise goes through the system, this plot will have a shape similar to the transfer function. It can be seen that this is indeed the case (see Figs. 6 and 7). By taking input-output covariances into account, identification can be done properly.

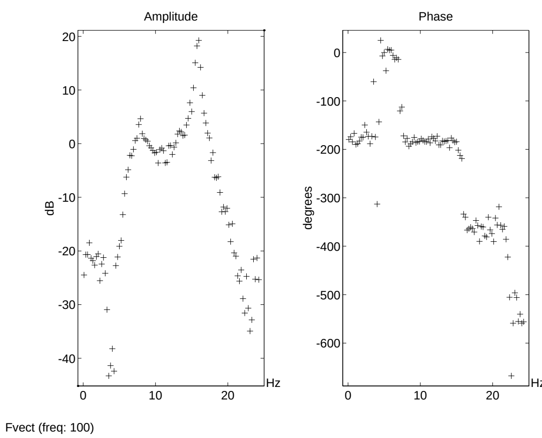

Fig. 6: The shape of cyx./vx, showing a pattern similar to the transfer function

A rough guess about the gain in SNR can be obtained by calculation of the SNR of the nonparametric estimate of the transfer function (see Fig. 7).

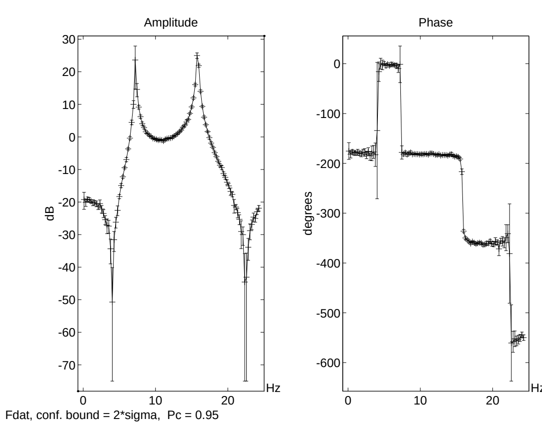

Fig. 7: Uncertainties of the nonparametric transfer function estimate

The SNR of this nonparametric estimate is somewhat smaller than those above, because the division Ym/Xm amplifies the noise significantly where the input amplitudes are small. In proper frequency domain parametric identification, done also by this toolbox, this division is therefore avoided. Instead, an appropriate cost function is minimized (Schoukens and Pintelon, 1991).

However, it is clear from Fig. 7 that the SNR is quite good indeed.

2.5. Identification

We can now proceed with identification. Since the experiments are very well synchronized, the average of the complex amplitudes, mx and my can be used. For the estimation algorithm, the numerator and denominator orders of the transfer function have to be given. From the nonparametric plot it is obvious that at least two complex pole pairs and two complex zero pairs will be necessary. So, let us start with a system 4/4.

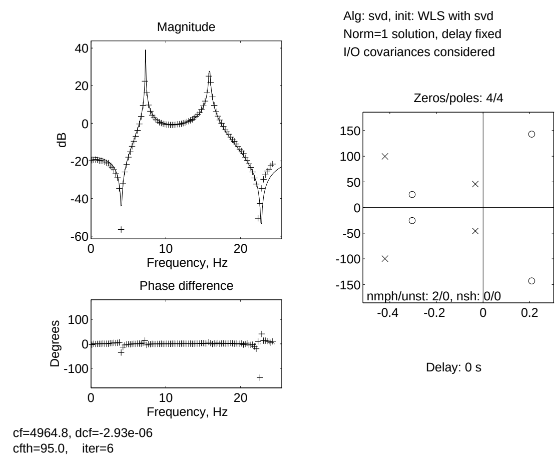

Fig. 8. Result of the 4/4 fit

The fit is quite good, but the cost function is still large (about 50 times the theoretical value), and there is an apparent mismatch at the higher frequency band. It seems to be reasonable to increase the orders. Let us try a 6/6 system.

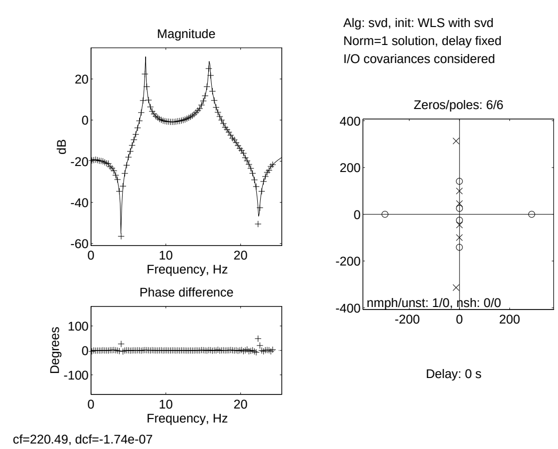

Fig. 9. Result of the 6/6 fit

The fit is much better. The cost function got quite close to the theoretical value, however, it is still larger than the theoretical value by about a factor of 2.5, so there are probably still small modeling errors. A model of order 8/8 and other order combinations can still be tried. However, these models are not much better than the 6/6 one, the cost function decreases by 5–10% only. The large number of necessary iterations is also an indicator of probable overmodeling. None of these trials is successful finding a better fitting stable model than the 6/6 one. The modeling error is probably due to nonlinearities. The order need not be further increased.

However, still there is a chance that a lower-order system can be as good as the 6/6 one. A plot of the confidence ellipses of the poles and zeros can be made, but the uncertainties are quite small. We can speculate that the real zero pair far from the imaginary axis (see Fig. 9) plays no important role, so it is reasonable to decrease the numerator order by 2. This will also allow that the transfer function decreases for higher frequencies, the usual behavior of physical systems. The two poles may correspond to the resonance around 42 Hz (260 rad/s), shown in the complex output amplitude plot. At this frequency there was no excitation applied, however, the nonlinearities produced enough overharmonics to show this resonance. For a proper identification of it, the complex input/output amplitudes around this frequency should also be used, and a broader excitations signal should have been applied. We are not going to specifically deal with this resonance, and will proceed with the above data.

Let us do now a 4/6 fit.

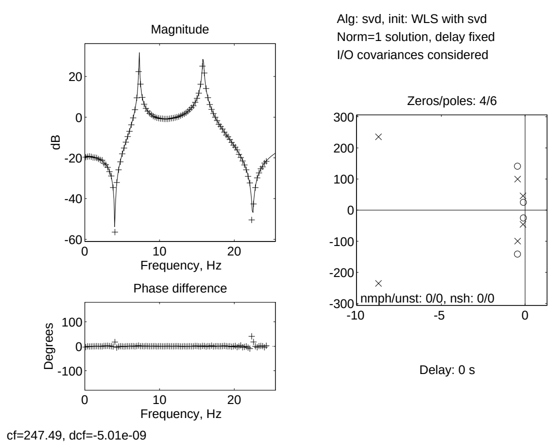

Fig. 10. Result of the 4/6 fit

2.6. Model Validation

The fit seems to be very reasonable. We have a good identified model.

For a thorough study of this fit, a report file can be generated, with estimated parameters, zeros and poles with uncertainties, diary of the estimation session, information about numerical stability, modeling errors, the value of the Akaike criterion and so on.

The model can be verified using the standard techniques of the toolbox. We will not do all the possible tests, but will do some of the typical ones. One of the most important indicators of the quality of the fit is the value of the cost function, already discussed above. It is also important to examine visually the quality of the fit on the plot of elis. In this case the error is too small to be easily detected on the plots. The phase errors at the zeros are not really important, since here the phase information of the measurements is small. The confidence interval plots using ploteltf could also be used, but the confidence intervals have to be magnified for visual checking.

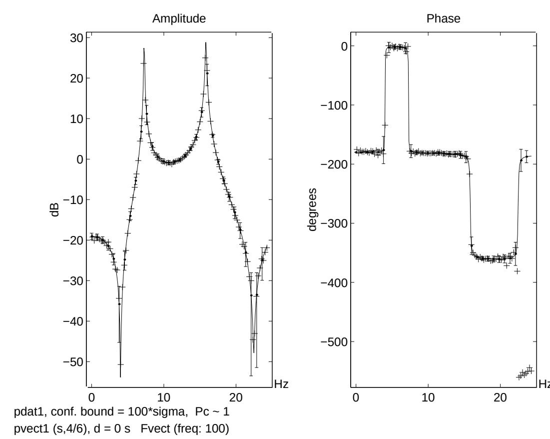

Fig. 11. Magnified confidence bounds of the 4/6 fit

Even with a bound of 100*sigma, not much can be seen. It is better to look for other tests. We think that there are modeling errors, so let us check the approximate mean model error (see Schoukens and Pintelon, 1991). This has about the same value for both the 4/6 and 6/6 models: 0.21, while the mean absolute value of the transfer function is about 1.30. These values illustrate that the modeling error is not completely negligible, and is in the same order of magnitude for all three fits. Because of the modeling error, the Akaike criterion cannot be used. For closer investigation of the quality of the fit, the residuals can be calculated.

The residuals can be studied for normal distribution and whiteness. We will do a few tests for the residuals of Ym/Xm, with respect to the parametric transfer function estimate. First, it has to be standardized, dividing the residuals at each frequency point by the standard deviation. If the fit is good, the standardized residuals have to exhibit circular standard normal distribution at each point, and they have to be independent. These properties will be checked by simple tests. First, let us draw the histograms of the real and of the imaginary parts. The dotted lines show the standard normal probability density function, scaled up to the histogram which is made of 1000 points, with a bin width of 0.2.

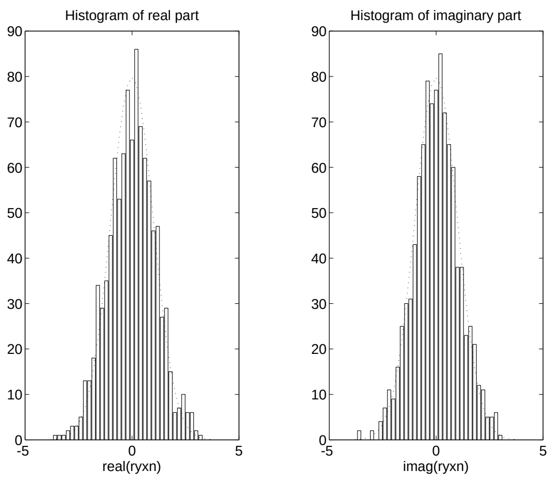

Fig. 12: Histograms of the real and imaginary parts of the complex residuals of Ym/Xm

The fit is good. The chi-square test can also be evaluated for both histograms, simply using the approximate probabilities. The theoretical value is 20.

$$\{\chi^2_{\mathrm{real}}\} = 22.5\,, \quad \{\chi^2_{\mathrm{imag}}\} = 10.4\,.$$

The test shows no significant deviation from the standard normal distribution.

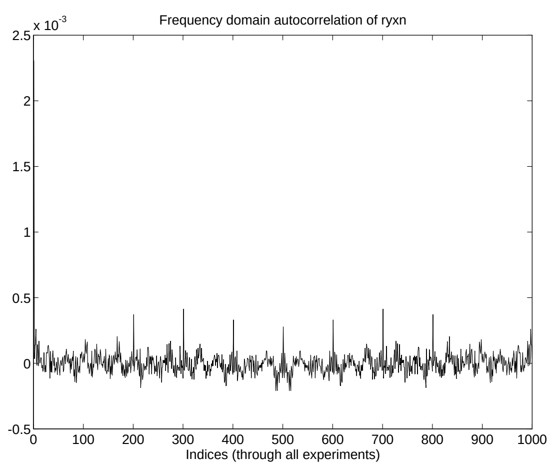

Fig. 13: Frequency domain autocorrelation of the residuals of Ym/Xm

As a last test, let us plot the autocorrelation function of the frequency domain residual series.

The autocorrelation function has a dominant peak at zero, an indication of the approximate uncorrelatedness of the residuals. The repeated smaller peaks are at lag distances of experiment lengths each, which indicates a small modeling error again, since it corresponds to a repetitive pattern in the residuals.

REFERENCES

Andersson, T., Pucar, P. and L. Ljung (1994). *Identification Aspects of Inter-Sample Input Behavior.* Technical Report LiTH-ISY-R-1606, Department of Electrical Engineering, Linköping University.

Guillaume, P. (1992). *Identification of Multi-Input Multi-Output Systems Using Frequency-Domain Models.* PhD Thesis, Vrije Universiteit Brussel, Dept. ELEC, Brussels, Belgium.

Kollár, I. (1994). *Frequency Domain System Identification Toolbox.* Natick, MA, The MathWorks.

Kollár, I., R. Pintelon, J. Schoukens (1991). Frequency Domain System Identification Toolbox for Matlab. *IFAC Symposium on Identification and System Parameter Estimation,* Budapest, 8-12 July, 1991. pp. 1243–1246.

Ljung, L. (1987). *System Identification. Theory for the User.* Englewood Cliffs, Prentice-Hall.

Ljung, L. (1992). *System Identification Toolbox.* (First version: 1986) Natick, MA, The MathWorks.

Ljung, L. (1993). Some Results on Identifying Linear Systems Using Frequency Domain Data. *IEEE International Conference on Decision and Control,* San Antonio, TX, Dec. 15–17, 1993. pp. 3534–3538.

Pintelon, R., P. Guillaume, Y. Rolain, J. Schoukens and H. Van hamme (1993). Parametric Identification of Transfer Functions in the Frequency Domain, a Survey. *IEEE International Conference on Decision and Control,* San Antonio, TX, Dec. 15–17, 1993. pp. 557–566.

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

Schoukens, J., R. Pintelon and H. Van hamme (1994a). Identification of Linear Dynamic Systems Using Piecewise Constant Excitations: Use, Misuse and Alternatives. *Automatica,* **30,** 7. To appear.

Schoukens, J., P. Guillaume and R. Pintelon (1994b). *Design of Broadband Excitation Signals.* In *Perturbation Signals for System Identification,* ed. K. Godfrey, London, Prentice Hall.