EXCITATION SIGNALS FOR STRUCTURE SELECTION OF NONLINEAR MODELS – RANDOM AND DETERMINISTIC CASES

Alexandro Garro Brito*, Elder Moreira Hemerly†, Waldemar de Castro Leite Filho*

*Instituto de Aeronáutica e Espaço.

†Instituto Tecnológico de Aeronáutica

‡Pça. Mal. Eduardo Gomes, 50. Vila das Acácias. São José dos Campos, SP. 12228-900. Brasil

Emails: alegbrito2@gmail.com, hemerly@ita.br, waldemarwclf@iae.cta.br

Abstract— Many studies have been conducted about the structure selection of autorregressive nonlinear models and some techniques are available for an efficient definition of black-box models that fit the experimental data set. In spite of the amount of work, few papers clearly consider questions involving the excitation signal. Random excitation is common, however this is not the general situation in practical applications. In this paper, some aspects involving the excitation signals are discussed, including both the random and stochastic cases. The aim is to aware the user about the results that should be expected according to the input signal, and some desirable excitation characteristics so that the model selection procedures is improved.

Keywords— Nonlinear systems, Identification, Structure selection.

Resumo— Muitos estudos têm sido realizados sobre a escolha de estrutura de modelos não-lineares autorregressivos e algumas técnicas estão disponíveis para a definição de modelos caixa-preta que ajustem os dados experimentais. Apesar da variedades de técnicas, poucos trabalhos tratam claramente questões envolvendo os sinais de excitação. Tais textos apresentam comumente o uso de excitações aleatórias, mesmo se sabendo que esse tipo de sinal não é usual em aplicações práticas. Neste artigo, alguns aspectos envolvendo tais sinais de excitação são discutidos, incluindo-se os casos determinístico e aleatório. O objetivo é alertar o usuário a respeito dos resultados esperados do processo de escolha da estrutura com base no sinal de excitação utilizado, bem como apresentar algumas características importantes para que o procedimento de seleção produza resultados adequados.

Palavras-chave— Sistemas não-lineares, Identificação, Seleção de estrutura

1 Introduction

The structure selection plays an important role in the nonlinear system identification and some methods have been developed for NARX/NARMAX models. The *Error Reduction Ratio (ERR)* (Billings et al., 1989) aims to evaluate the influence of each candidate regressor over the output signal formation in a Least Squares scenario. This criterion relates the regressor importance in terms of its capacity of explaining the output variance.

The most of paper uses the ERR strategy in a scenario where random excitation is assumed (Leontaritis and Billings, 1987; Aguirre et al., 2002). In spite of the relevant characteristics of the ERR criterion and the important properties of random signals, it lacks a more profound study on how the ERR behaves in the statistical sense. Moreover, the random excitation is not always proper to be used in practical applications since many mechanical systems cannot be excited by such signals. Then, alternative multisinusoidal signals have been considered for structure selection (Brito, 2013) but their influence over ERR should be previously established.

This paper aims to discuss some questions involving the excitation signal for model structure selection. The data length and sampling effects on the selection results are discussed. Finally, the broadband class of excitation is presented and some desirable characteristics for structure selection are studied. This enables a proper input design.

The paper is organized as follows: section I presents an overview about NARMAX models and the ERR criterion. Section III shows some experiments involving the structure selection by using white noise excitations so that conclusions about the ERR statistical characteristics can be drawn. Section IV establishes the equivalences between the multisine and random excitation in terms of the structure selection efficiency. Finally, some conclusions are drawn in section V.

2 NARMAX models and structure selection

There are several mathematical representations for nonlinear systems available in the literature, each of them with its own characteristics, assumptions and range of applicability. One of the most successful is the NARX/NARMAX model (Leontaritis and Billings, 1985; Chen and Billings, 1989). These authors showed that, based on mild assumptions, a wide class of nonlinear systems can be expressed through a nonlinear recursive discrete formulation written as

$$y_k = F(y_{k-1}, \ldots, y_{k-n_y}, u_{k-1}, \ldots, u_{k-n_u}), \quad (1)$$

where u and y are respectively the input and the output and n_u and n_y are the maximum lag for each signal in the model.

A small subset of this representation is constructed through polynomial functions relating the past states of the discrete signal u and y. These are called *NARX polynomial models* and they can be written as (Aguirre et al., 2002)

$$y_k = \sum_{m=0}^{d} \sum_{p=0}^{m} \sum_{n_1,n_m}^{n_y,n_u} c_{p,m-p}(n_1, \cdots, n_m) \times \times \prod_{i=1}^{p} y_{k-n_i} \prod_{i=p+1}^{m} u_{k-n_i} \tag{2}$$

where d is the polynomial degree and

$$\sum_{n_1,n_m}^{n_y,n_u} \equiv \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_y} \cdots \sum_{n_m=1}^{n_u} . \tag{3}$$

The NARX polynomial models are linear in the parameters $\theta_{(i,j)}$. Thus, a *Least Squares* identification procedure can be used. This is given by

$$\begin{aligned} y_k &= \Psi_{k-1}\hat{\theta} + \xi_k \\ &= \sum_{i=1}^{n_\theta} \hat{\theta}_i \psi_{i,k-1} + \xi_k \end{aligned} \tag{4}$$

where the regressor matrix Ψ is constructed with all the terms of (2). ξ is the estimation residue.

The first problem to be faced during the identification of a NARX polynomial model is: which are the necessary regressors in (4) (or the monomial terms in (2)) so that the nonlinear behavior is correctly captured by the model? The answer is not easy to obtain, because both underparametrized and overparametrized models can conduct to very poor output response (Aguirre and Billings, 1995). Hence, a prior structure selection step plays key role in the identification process.

In this paper, a *relevant regressor* is denoted as one that is able to represent the output in a better fashion than another, based on some criterion. Several criteria can be used for this purpose (Piroddi and Spinelli, 2003; Kukreja et al., 1999) and the most obvious is the *identification error* provided by each regressor individually. Thus, one can rank a determined regressor according to its ability to "explain" the output data when compared to another one. This approach is the basis of the *Error Reduction Ratio (ERR)* (Billings et al., 1989). This ratio is based on the full orthogonalization of the regression matrix, so that the individual importance of each regressor can be evaluated.

Consider again (4) and an auxiliary model

$$y_k = \sum_{i=1}^{n_\theta} \hat{g}_i w_{i,k-1} + \xi_k \tag{5}$$

where the regressors w_i are orthogonal over the data, or, $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = 0$, with $\langle \cdot \rangle$ denoting the inner product. It can be shown that

$$\langle \mathbf{y}, \mathbf{y} \rangle = \sum_{i=1}^{n_\theta} \hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle + \langle \xi_i, \xi_i \rangle. \tag{6}$$

Eq. (6) permits to quantify the contribution of each regressor w_i introduced into the model. The *Error Reduction Ratio* due to the ith regressor is defined as

$$[ERR]_i = \frac{\hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle}. \tag{7}$$

In this way, a criterion to be used for the regressor selection is to choose those regressors that have the highest ERR. This ratio relates the importance of the included term to its capacity to explain the output variance (Piroddi and Spinelli, 2003).

3 ERR and the white noise excitation

Many studies in structure selection of NARX/NARMAX models are based on white noise excitation, but few realizations are performed and richer information about the ERR behavior is not possible (Swain and Billings, 1998; Mao and Billings, 1999; Piroddi and Spinelli, 2003; Leontaritis and Billings, 1987). Then, this paper proposes some statistical experiments to evaluate and discuss the characteristics of such method.

3.1 Test conditions

To perform the mentioned tests, two different NARX systems where chosen,

$$\begin{aligned} \mathcal{S}_1 : \; y_k &= 0.5y_{k-1} + 0.8u_{k-2} + u_{k-1}^2 - \\ &- 0.05y_{k-1}^2; \end{aligned} \tag{8}$$

$$\begin{aligned} \mathcal{S}_2 : \; y_k &= 0.8y_{k-1} + 0.4u_{k-1} + 0.4u_{k-1}^2 + \\ &+ 0.4u_{k-1}^3; \end{aligned} \tag{9}$$

The first system was considered in (Piroddi and Spinelli, 2003), while $\mathcal{S}_2$ was studied in (Bonin et al., 2010). In both papers, these models are used to test the efficacy of model structure selection approaches, always based on the white noise excitation.

For this paper, an initial NARX candidate regressor set is obtained by considering $d = 3$ and the linear regressors $\{u_k, u_{k-1}, \cdots, u_{k-3}, y_{k-1}, \cdots, y_{k-3}\}$ in (2). This produces a regression matrix with 119 linear and nonlinear regressors. Along the discussions, these regressors are coded by their initial position in the candidate regression matrix. This was done to render the following tables more compact. Obviously, only a small set of the initial regressors will appear along the text and Table 1 presents the numeric codification of them.

Table 1: Numeric codification of the regressors.

1.	y_k	7.	y_{k-3}	38.	$u_k^2 y_{k-1}$
2.	u_{k-1}	8.	u_k^2	43.	$u_k u_{k-1}^2$
3.	y_{k-1}	9.	$u_k u_{k-1}$	44.	$u_k u_{k-1} y_{k-1}$
4.	u_{k-2}	15.	u_{k-1}^2	64.	u_{k-1}^3
5.	y_{k-2}	21.	y_{k-1}^2	65.	$u_{k-1}^2 y_{k-1}$
6.	u_{k-3}	37.	$u_k^2 u_{k-1}$	85.	y_{k-1}^3

In all the proposed experiments, 100 realizations of the excitation signals are performed and the ERR values are obtained for each regressor and realization. Then, the ERR mean values of the most important regressors are presented, considering the 99.73% confidence limits ($\pm 3\sigma$).

3.2 Data length

The first question to be answered is if the data length of a white noise excitation has any effect over the ERR method. In spite of one can use its insight, more detailed analysis is necessary. To test this, the regressor ERR was evaluated for Gaussian signals with unitary variance and three amounts of points: 1000, 500, 250 and 100. The results are summarized in Table 2.

The mean values for ERR along the realizations were not considerably modified by the amount of points N in the excitation. However, the variance seems to be increased for all the regressors when N decreases, as it could be already expected. Such high variance levels can make certain regressors statistically insignificant and the user should be aware about this, since that useless regressors can be selected with basis on false information provided by only one realization.

For example, one can notice that the confidence limits for the regressor 7 ($\mathcal{S}_1$) are bigger than the mean ERR. This results in a standard deviation higher than the mean level, an inconsistent conclusion. Although this can be removed by considering a larger amount of realizations, this behavior can lead the user to choose an insignificant regressor if only one experiment is taken to perform the structure selection.

3.3 Sampling time

Perhaps the most important factor to be considered in the signal design is the sampling time. This topic has been discussed in several papers and many authors presented important results that show how a wrongly chosen sampling time can damage the ERR accuracy (Piroddi and Spinelli, 2003; Aguirre et al., 2002).

A very simple way to see this is to analyze the following example. Consider a continuous-time nonlinear system expressed by

$$\frac{dy(t)}{dt} = u(t) + bu^2(t) + cu^3(t) - ay(t). \tag{10}$$

This equation can be seen as a nonlinear system modeled by a Hammerstein formulation. Eq. (10) can be straightforwardly discretized through Euler's derivative approximation, and the same system can be expressed by the difference equation

$$\begin{aligned} y_k = {} & (1 - a\Delta T)y_{k-1} + \Delta T u_{k-1} + b\Delta T u_{k-1}^2 \\ & + c\Delta T u_{k-1}^3, \end{aligned} \tag{11}$$

where ΔT is the sampling time. Without loss of generality, consider $a = 0.5$ and $b = c = 2$. Notice that if $\Delta T = 0.4$, (11) collapses to (9). However, if $\Delta T = 0.04$ and $\Delta T = 0.004$ the difference equations become, respectively,

$$\begin{aligned} y_k = {} & 0.98y_{k-1} + 0.04u_{k-1} + 0.08u_{k-1}^2 \\ & + 0.08\Delta T u_{k-1}^3, \end{aligned} \tag{12}$$

$$\begin{aligned} y_k = {} & 0.998y_{k-1} + 0.004u_{k-1} + 0.008u_{k-1}^2 \\ & + 0.008\Delta T u_{k-1}^3. \end{aligned} \tag{13}$$

Is easy to see that the smaller is the sampling time the larger is the importance of the regressor y_{k-1} over the others. In other words, when the sampling time is very low, the model behaves as a one-step predictor and their data looses almost all the nonlinear information contained in the real system. If the ERR is used to rank and select the most important regressors, it will attribute a level of more than 0.99 to y_{k-1} and the other regressors are almost surely lost.

An empirical, but very effective, result to choose a proper sampling time for the structure selection (and identification) is to use ΔT not less a tenth of the dominant constant time of the system to be modeled (Aguirre et al., 2002).

4 ERR and the multisinusoidal excitation

All the results presented in the last section refer to white noise excitation. This is very common in papers concerning the identification and structure selection of NARX and NARMAX models due to the nice properties that such signal has. Unfortunately, this type of signal is not proper for many practical identification problems because several real systems (mechanical, chemical, etc.) cannot be excited by a pure random signal. Some pseudo-random signals have been proposed to circumvent this difficulty, but even in this case some experimental problems can be faced.

The use of broadband signals, where the multisine signal is included, is an option that has been successfully used in many applications (Pintelon and Schoukens, 2001). Due to their strong practical appeal and the flexibility in choosing a proper spectral content, these signals can be useful in many problems, including the structure selection (Chen and Billings, 1989). The question to be answered is about the necessary characteristics such that these signals are really applicable. In other

DOI: 10.5540/03.2013.001.01.0072

Table 2: Data length effect on ERR.

	N=1000		N=500		N=250		N=100	
	Reg.	ERR	Reg.	ERR	Reg.	ERR	Reg.	ERR
$\mathcal{S}_1$	15	0.459±0.165	15	0.453±0.112	15	0.435±0.180	15	0.368±0.228
	3	0.304±0.141	3	0.314±0.154	3	0.308±0.221	3	0.308±0.316
	4	0.127±0.070	4	0.123±0.084	4	0.125±0.126	4	0.125±0.172
	5	0.036±0.030	5	0.036±0.034	5	0.038±0.049	2	0.038±0.148
	7	0.021±0.020	7	0.021±0.022	7	0.022±0.039	5	0.022±0.059
$\mathcal{S}_2$	3	0.729±0.044	3	0.725±0.071	3	0.721±0.091	3	0.703±0.155
	2	0.175±0.044	2	0.181±0.061	2	0.185±0.085	2	0.200±0.136
	64	0.056±0.038	64	0.047±0.046	64	0.035±0.044	15	0.027±0.059
	15	0.027±0.035	15	0.026±0.039	15	0.028±0.047	64	0.013±0.025
	8	0.002±0.002	1	0.002±0.010	1	0.004±0.019	1	0.010±0.036

words, the user should be able to design the signals so that the structure selection is properly performed. This paper discusses some details such that the parameters of a multisine signal can be set for the structure selection purpose.

A multisine signal is defined as

$$u_k = \sum_{l=l_{min}}^{l_{max}} A_l \sin\left(\frac{2\pi k}{\omega_s} l\omega_0 + \phi_l\right), \qquad (14)$$

where A_l is the amplitude of the l^{th} harmonic ($l \in \mathbb{N}$), ω_s is the sampling frequency, ω_0 is the multisine fundamental frequency and $\phi_l \sim U(-\pi, \pi)$. l_{min} and l_{max} are the signal bandwidth cutoffs, set by the user according to the application. In spite of the phase is a random number for each harmonic, one can note that y_k is a deterministic signal since this random phase is maintained constant along the signal realization. The data length of one multisine signal period is given by the relation between ω_s and ω_0, or, $N = {}^{\omega_s}/_{\omega_0}$.

The first point to be discussed is if a multisine signal can provide ERR behavior similar to that provided by a white noise excitation. In principle, there is no major difficulty according to what follows. Consider a Gaussian white noise with N points. In the frequency domain, this signal contains useful power in the entire frequency spectrum. This means that if a Fast Fourier Transform is applied and by assuming a sampling frequency of ω_s, the result – an N-point complex data – will contain useful power spectra for all the frequency lines up to ${}^{\omega_s}/_2$ (according to the Shannon theorem). Moreover, it is easy to conclude that the FFT frequency resolution is given by ${}^{\omega_s}/_N$. Based on this information, one can construct a N-point multisine with similar spectral content by setting $\omega_0 = {}^{\omega_s}/_N$, $l_{min} = 1$, $l_{max} = {}^N/_2$ and A_l such that both the white noise and the multisine present the same RMS value. Hence, both the signals are expected to have the same persistence and "aperiodic" behavior for one multisine period. In conclusion, they should provide similar ERR and identification results if the same problem conditions are assumed.

To exemplify this, consider the ERR values from Table 2 obtained through Gaussian white noise signals with 1000 points. Assume, for instance, a sampling time of 0.1 sec. By applying FFT to these signals, one will obtain a spectrum with information between $\pm 10\pi$ ${}^{rad}/_s$. Moreover, each the frequency lines are separated by intervals of $\omega_0 = {}^{\omega_s}/_N$. Now, take (14) with these characteristics and set A_l such that multisine has RMS value of 1 (equals to the white noise). Such multisine signals were applied to the test systems and the results are summarized in Table 3. As it was expected, the ERR mean values are quite similar to Table 2, however the confidence limits are considerably low. The reason for this is that the Gaussian white noise is actually produced by a pseudorandom generator. Besides its good results for the noise mean and variance, the pseudorandom generator only approximates the signal whiteness, working better for a bigger amount of data points.

Table 3: Multisine signal equivalent to a Gaussian white noise.

$\mathcal{S}_1$		$\mathcal{S}_2$	
Reg.	ERR	Reg.	ERR
15	0.462±0.083	3	0.731±0.048
3	0.315±0.108	2	0.176±0.043
4	0.127±0.056	64	0.054±0.032
5	0.036±0.026	15	0.028±0.029
7	0.022±0.019	8	0.002±0.002

The results indicates that it is possible to design a multisine signal proper for structure selection through ERR. However, a more detailed strategy is necessary to design a multisine signal properly. How should the multisine parameters be chosen, considering the application practical constraints? The answer is explored in what follows.

4.1 Spectral content

The last section pointed out that an N-point multisine with ${}^N/_2$ excited frequency lines (or harmonics) will behaves similarly to an N-point Gaussian white noise with the same RMS value. In other words, a multisine with N_{exc} excited frequency lines has the same persistence that a Gaussian white noise with $2N_{exc}$ points. This information is the key to set the spectral content of a specific multisine signal. For example, suppose a multisine signal defined as (14) and consider, initially, $l = l_{min}, \cdots, l_{max} \in \mathbb{N}$. Consider a multisine with $\omega_s = 20\pi\ {}^{rad}/_{s}$ and $\omega_0 = 0.01\pi\ {}^{rad}/_{s}$, what results in a signal of 2000 points. By assuming $l_{min} = 1$ and $l_{max} = 1000$, the signal has 1000 excited harmonics. Hence, such multisine is equivalent to a Gaussian white noise of 2000 points. 100 realizations of the multisine signals with unitary RMS value and the characteristics above were applied to the test systems $\mathcal{S}_1$ and $\mathcal{S}_2$ and the results are in Table 4 (Condition 1). Note that the ERR values are very similar to those in Table 3, but the variance is smaller as already expected.

More flexible multisine bandwidth can be chosen, but this will result in increasing of the signal data points when compared to a white noise signal. To see this, consider now an odd-multisine whose l in (14) are only odd numbers between l_{min} and l_{max} (l_{min} and l_{max} are also odd numbers). Obviously, $N_{exc} = {}^{(l_{max}-l_{min})}/_{2}$ and the equivalent white noise has $2N_{exc}$ points. However, to produce an odd-multisine with the same spectral content than a "full" multisine, the double of points are necessary since the even frequency harmonics are not excited. For instance, consider the same parameter settings used in the Condition 1 above, but for an odd-multisine. This will give the same 2000 data points but with only 500 excited frequency lines. Then, this signal will be approximately equivalent to a gaussian white noise of 1000 points, half than the designed multisine. 100 realizations of such multisine were performed and applied to the test systems (Condition 2 in Table 4). Once again, the ERR values are very similar to those in Table 3, but with smaller confidence limits and, in principal, these limits should be similar. The reason is that the A_l amplitude is higher for this case, then the power spectra applied to each frequency line is higher. This seems to have effect on the ERR variance.

As a final experiment, consider an odd-multisine whose $\omega_s = 20\pi\ {}^{rad}/_{s}$ and $\omega_0 = 0.005\pi\ {}^{rad}/_{s}$, what results in a signal of 4000 points. By assuming $l_{min} = 1$ and $l_{max} = 1000$, it will again contain 500 excited frequency lines. Note that this is a narrower bandwidth signal since $l_{max} < {}^{N}/_{2}$. 100 realizations of such multisine were performed and applied in the test systems (Condition 3 in Table 4). The ERR values are very similar to those in Condition 2, but the confidence limits are still smaller. This shows that the relation between the white noise data length and the multisinusoidal amount of frequency lines should be considered only as a preliminary design parameter. In all the cases, however, the multisine input presented better selection behavior.

4.2 Harmonic's amplitude selection

In many moments along this paper, it is considered that the Gaussian white noise and the multisine have the same "excitation level". This was related to the RMS value but no relation to any harmonic amplitude A_l was provided. This is done in what follows. The RMS value of a signal x_k is calculated by

$$x_{RMS} = \sqrt{\frac{1}{N}\sum_{k=1}^{N} x_k^2}. \tag{15}$$

For a perfect zero-mean Gaussian white noise with variance σ^2, this results

$$\xi_{RMS} = \sqrt{\frac{1}{N}\sum_{k=1}^{N} \xi_k^2} = \sqrt{\sigma^2} = \sigma. \tag{16}$$

Then, the RMS value of a white noise in these conditions equals its standard deviation.

The RMS value of the multisine expressed as (14) is given by

$$u_{RMS} = \sqrt{\sum_{l} \left(\frac{A_l}{\sqrt{2}}\right)^2}. \tag{17}$$

By considering that all the harmonics have the same amplitude $A_l = A$ and that there are N_{exc} excited harmonics in the signal, the RMS value simplifies to

$$u_{RMS} = A\sqrt{\frac{N_{exc}}{2}}. \tag{18}$$

By combining (16) and (18), one can obtain

$$A = \sigma\sqrt{\frac{2}{N_{exc}}}. \tag{19}$$

In conclusion, a multisine signal with N_{exc} harmonics of amplitude A has the same excitation level of a zero-mean Gaussian white noise with variance σ^2 if (19) holds. This result was used to adjust the multisine harmonic amplitudes such that the RMS value was equal to the respective random signal in the experiments presented in Tables 3 and 4.

5 Conclusions

This paper presented several considerations about the excitation signals used for the structure selection and the identification of NARX/NARMAX models. This study was performed by using the Error Reduction Ratio (ERR), one of the most successful methods for the task.

Table 4: ERR for different multisine spectral contents.

	Condition 1		Condition 2		Condition 3	
	Reg.	ERR	Reg.	ERR	Reg.	ERR
$\mathcal{S}_1$	15	0.469±0.050	15	0.469±0.067	15	0.475±0.054
	3	0.307±0.072	3	0.309±0.081	3	0.304±0.066
	4	0.129±0.042	4	0.129±0.011	4	0.126±0.009
	5	0.037±0.016	5	0.036±0.019	5	0.037±0.014
	7	0.021±0.012	7	0.021±0.016	7	0.022±0.013
$\mathcal{S}_2$	3	0.731±0.030	3	0.728±0.034	3	0.729±0.025
	2	0.175±0.025	2	0.176±0.022	2	0.176±0.018
	64	0.060±0.025	64	0.060±0.027	64	0.063±0.022
	15	0.026±0.021	15	0.027±0.004	15	0.027±0.003
	8	0.002±0.001	43	0.002±0.013	8	0.016±0.001

Random signals have been used in a lot of contexts involving the structure selection and identification of nonlinear systems due to their important properties. However, this study shows that the selection efficiency is strongly related to the signal amount of points N. As it was observed, the ERR variance depends on N and regressors with irrelevant statistical ERR values can be inadvertently chosen if only one realization is considered.

In addition, these signals are not always applicable to many real systems due to their strong pulsating aspect. Then, other signals should be studied and an alternative is to use a broadband excitation such as the multisine signals. In this case, the user is able to select the frequency content and this paper presented a study of the equivalence between this class of signals and the random excitation. The guidelines provided herein help to design a multisinusoidal signal with equivalent characteristics to a random one. As discussed, the results are qualitatively valid for the multisine input design. Hence, the user can apply it to perform the model structure selection in a practical application, where a random excitation is not possible.

References

Aguirre, L. A. and Billings, S. (1995). Dynamical effects of overparametrization in nonlinear models, *Physica D* **80**: 26–40. DOI: 10.1016/0167-2789(95)90053-5

Aguirre, L. A., Corrêa, M. and Cassini, C. (2002). Nonlinearities in NARX polynomial models: representation and estimation, *IEE Cont. Theory App.* **149(4)**: 343–348. DOI: 10.1049/ip-cta:20020398

Billings, S., Chen, S. and Korenberg, M. (1989). Identification of MIMO nonlinear systems using a forward-regression orthogonal estimador, *Intern. Journ. Control* **49(6)**: 2157–2189. DOI: 10.1080/00207178908961377 DOI: 10.1080/00207178908559767

Bonin, M., Seghezza, V. and Piroddi, L. (2010). NARX model selection based on simulation error minimisation and LASSO, *IEE Cont. Theory App.* **4(7)**: 1157–1168. DOI: 10.1049/iet-cta.2009.0217

Brito, A. (2013). *Seleção de estrutura de modelos autorregressivos não-lineares no domínio da frequência*, PhD thesis, ITA.

Chen, S. and Billings, S. (1989). Representations of nonlinear systems: the NARMAX model, *Intern. Journ. Control* **49(3)**: 1013–1032. DOI: 10.1080/00207178908559683

Kukreja, S., Galiana, H. and Kearney, R. (1999). Structure detection of narmax models using bootstrap methods, *Proc.* 38^{th} *Conf. Decision and Control*.

Leontaritis, I. and Billings, S. (1985). Input-output parametric models for nonlinear systems. Part I: deterministic non-linear systems, *Intern. Journ. Control* **41(2)**: 303–328. DOI: 10.1080/0020718508961130 DOI: 10.1080/0020718508961129

Leontaritis, I. and Billings, S. (1987). Model selection and validation methods for non-linear systems, *Intern. Journ. Control* **45(1)**: 311–341. DOI: 10.1080/00207178708933730

Mao, K. and Billings, S. (1999). Variable selection in non-linear systems modelling, *Mechanical Syst. & Sign. Processing* **13(2)**: 351–366. DOI: 10.1006/mssp.1998.0180

Pintelon, R. and Schoukens, J. (2001). *System identification - a frequency domain approach* IEEE Press. DOI: 10.1002/0471723134

Piroddi, L. and Spinelli, W. (2003). An identification algorithm for polynomial NARX models based on simulation error minimization, *Intern. Journ. Control* **76(17)**: 1767–1781. DOI: 10.1080/00207170310001635419

Swain, A. and Billings, S. (1998). Weighted complex orthogonal estimator for identifying linear and nonlinear continuous time models from generalized frequency response functions, *Mechanical Syst. & Sign. Processing* **12(2)**: 269–292. DOI: 10.1006/mssp.1997.0146

