Automatic Model Selection for Linear Time-Invariant Systems – Practical Issues

Gyula Simon, Johan Schoukens and Yves Rolain

Abstract– A completely automatic identification system is described which uses the latest results of the frequency domain identification approach to provide an easy-to-use and reliable tool to the inexperienced users. The proposed system accepts periodic measurement data and performs the whole identification procedure automatically from the data-preprocessing step to model and parameter selection. The validation phase provides information on the results at the level of the non-expert user. The performance of the system is illustrated by results obtained from real measurement data.

***Keywords*: model selection, system identification**

I. INTRODUCTION

The theory of estimation of linear time-invariant (LTI) systems and systems with a dominant LTI behavior perturbed by nonlinear disturbances is well defined and available in the literature [1,2,3,4]. The system identification process, however, is so complicated that the proper usage of the theoretical results (or even the usage of the available tools [5,6]) requires a solid theoretical background, signal processing skills, and know-how on the identification field. While some parts of the identification process can be (and have been) automated, some parts still require decisions of human experts. To allow users that are experienced in their own field to use identification – and hence extract high quality validated models – without having to become an identification expert, the process should be fully automated. The final goal is that users "drop in the measurement results, have a coffee, and get the final validated model" without further intervention.

Many new contributions have been published lately in the identification field [1,2,3,4,8,9], which can serve as fundamental elements in the construction of a fully automated identification process.

Based upon these results in this paper an automatic model selection and validation system is described.

The flowchart of the system identification process can be seen in Fig. 1. After the measurement, that is set up to obtain repeated measurements (at least 4 repeated measurements or 4 periods of the signal) a non-parametric noise model is extracted out of the data by calculating the non-parametric variances [7]. This preprocessing step can easily and automatically be accomplished by using periodic excitation signals, as will be discussed in Section II.

The solution of the complete identification problem is then further split into two parts: a model order estimation part and a parameter estimation part.

The model selection phase estimates the order of the LTI model, which can adequately describe the linear dynamic behavior of the system. Note that many practical systems contain perturbations that avoid the system to be fully described by a LTI model. Hence, no exact model order will exist as the system falls out of the chosen model class. Other small perturbations, such as device non-linearities, unmodeled device dynamics, instrumentation errors or calibration residuals will further contribute to the lack of identifiability of "the exact model" even if it exists. The method proposed will hence deliver a model which describes the system dynamics adequately, and in this sense is "close" to this ideal model order. In any case, at some moment modeling errors may require a user decision to be made.

In some cases, there is a priori knowledge upon the system: the order determination may then be straightforward through the analysis of the underlying physical phenomena. In most cases, however, the knowledge of the system is too idealized, and hence enough only to an approximate estimation of the correct model order. In many situations, there may not be any *a priory* information about the system's order. In these cases, the model order must be determined either by guess-and-trial methods (which are very commonly used), or by systematic order selection methods, which to the best knowledge of the authors, do not exist in the literature.

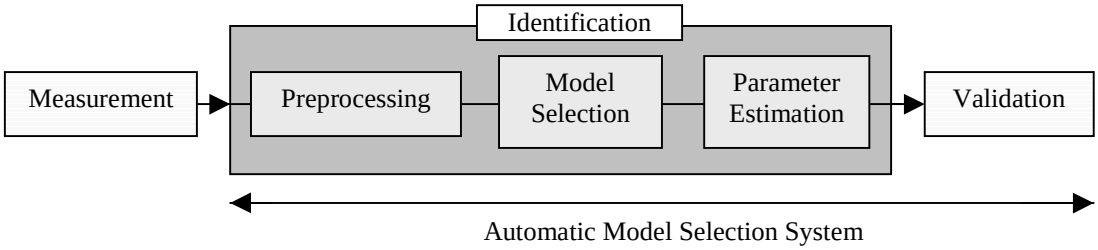

Fig.1. The identification process and the tasks of the Automatic Model Selection System

A systematic order selection method can use either bottom-up (from simple to complex) or top-down (from complex to simple) approach.

A bottom-up approach starts the search from a simple model and tries to improve the fit by adding additional degrees of freedom until the desired (user defined) precision is reached. The main advantage of this approach is that the complexity of the model during the estimation process remains always below the final complexity, which is a useful property from the viewpoint of the numerical stability of many widely used algorithms. The top-down approach starts from a high-order system and the complexity is reduced while the fit is still satisfactory. Practical experience shows that this approach can avoid local minima with higher probability, and thus the top-down models can provide better estimations than their bottom-up counterparts, especially in the cases when the model order is high. Since top-down methods must be able to cope with high orders and common pole-zero pairs, the numerical stability is a key issue for the considered application.

The parameter estimation phase is a very deeply examined field and thus many theoretical results are available [1,2,3,4]. The main problem of all the known algorithms is that they tend to find local minima of different quality, depending on the initial conditions. Therefore, these algorithms require a good initial guess in order to provide "optimal" results. The determination of the initial values of the algorithms and the parameterization of the sophisticated search methods to make the best of their capabilities requires a great deal of know-how from the users of these algorithms.

The heart of the described system is a top-down mixed order and model approximation algorithm. In an initialization phase, an appropriate starting value is determined, based upon [8]. This initial order is first determined by a coarse order estimation step, and then using this as an initial guess, the final order and model are determined in an iterative way. To provide the necessary numerical stability orthogonal polynomials are used in the representation of the LTI model, and during the iterative estimation loop sophisticated initial values are provided for the model estimation algorithm.

Since the model estimation should serve the requirements of the user, the proposed system is able to provide a set of models to fulfill strict requirements with higher model orders and loosen the requirements with lower complexity as well. The high order models extract all linear dynamics from the data, while a low order model is assumed to have errors below a user defined value. Note that although the system gives reasonable models, it is possible that personal experience leads to a better decision. The full process will be described in Section III.

The calculated models are useless unless the user knows their quality, thus the automatic validation and qualification of the models is as important for the user as the model itself. Since the model quality strongly depends on the goal of the user, the final decision is in his/her hands. The proposed system can give useful hints and tips in addition to the global qualification of the provided model, which is based on different numerical and statistical properties. The additional textual information helps the inexperienced user to understand and interpret the meaning of the validation tests, as described in Section IV.

The performance of the proposed system is illustrated by examples based on real measurement data in Section V.

II. AUTOMATIC DATA PREPROCESSING

To make automatic processing easy, the measurement must be made with periodic excitation signals. The described approach uses frequency domain methods, so each period of the measured input and output signals is transformed to the frequency domain by the DFT. Using these input and output spectra, the sample mean $\left(\overline{U}(\omega_k), \overline{Y}(\omega_k)\right)$, the sample variance $\left(\sigma_U^2(\omega_k), \sigma_Y^2(\omega_k)\right)$, and the sample covariance $\left(\sigma_{YU}^2(\omega_k)\right)$ of the input and output spectra are calculated [7].

III. MODEL SELECTION AND PARAMETER ESTIMATION

The inputs of the automatic model and order selection algorithm are the preprocessed data with the non-parametric noise model (i.e. the mean value, the sample variances and covariance of the input and output). The process can be divided into four major components:

A. coarse order estimation,
B. parameter estimation and model validation,
C. order reduction #1, and
D. order reduction #2.

The coarse order estimation step gives a rough (and usually conservative) estimation of the model order alone. Step B determines the model parameters with the estimated order and tries to validate it. If the validation fails, the model order is increased until the validation is successful. The first order reduction step produces a validated good quality model with the lowest possible order, while the second order reduction step tries to further decrease the order at the expense of a minor diminution of the model quality. The components of the automatic model and parameter estimation system are explained below.

Step A: The coarse order detection algorithm is based upon the method proposed in [8], which is able to reduce the initial model order using stochastic methods, without any parameter estimation.

Since this step does not require parameters to be estimated, it is faster than a full estimation run. Therefore, it is worthwhile to refine this step, to obtain maximal accuracy for the model order, as it will reduce the number of subsequent runs to get the

final model. In the proposed system, the original coarse estimation method is extended in two ways: to provide possibly lower order estimates, and to avoid the necessity of an initial order, which is "high enough".

Since the coarse order reduction step is fast, it is repeated iteratively until no further decrease is gained. If the initial order is detected to be too low, the order is automatically increased by a constant factor (in the present system by 20 percent) until a satisfactory (or the maximum allowed) order is reached. See Appendix 1. for details.

Step B: Because of the stochastic nature of this order estimation algorithm, it sometimes reduces the model order too much. Hence, the result must be verified before further processing. Using the estimated order, first the model parameters are estimated [4,7]. Then, the model is checked using correlation test on the residual error [9]. If the correlation test fails, the model order is increased and the parameter estimation and correlation test steps are repeated until the correlation test succeeds, or the maximum allowed order is reached.

Note that the reason of the failure of the correlation test may be not only the low model order. A suboptimal solution (local minimum) of the used parameter estimation algorithm may lead to analogous results. In both cases, however, the right solution to the validation failure is to increase the model order until the parameter estimation algorithm can find an acceptable – even high order – solution, from which Step C is able to produce a good quality lower order model.

Step C: The validated model is then processed by a peeling algorithm based upon [8]. The peeling process eliminates the roots (poles, zeros, or canceling pairs) which have no significant effect on the transfer function. Theoretically, the peeling algorithm is very simple:

Step 1 Choose some roots which can possibly be eliminated.
Step 2 Re-estimate the model and parameters with the decreased order.
Step 3 Check the result. If the reduced order model is validated then accept the decreased order and go to Step 1.
Step 4 Restore the previous order. If there are other roots to try, go to Step 1.

Apart from the very high computational requirements of this approach, several other questions arise concerning the practical realization of the peeling process. The most important practical issues are the following:

- how to decrease the computational complexity of the root elimination procedure (Step 2),
- how to decide which roots to eliminate (Step 1),
- is it possible to decrease the possibility of local minima in the parameter estimation phase using available extra information (Step 2),
- how to decide whether an estimation is acceptable or not (Step 3)?

The computational complexity can effectively be decreased by reducing the number of the parameter estimation steps. The greater the number of the eliminated roots in one parameter estimation step, the faster the algorithm. The used method is based upon the idea proposed in [8], which uses a simplified approach to give a conservative estimation on the cost function after eliminating some roots. (Note that in this paper the term "cost function" refers to the minimum description length cost function [1,11].) This non-parametric estimation is much faster than the parameter estimation step (which is from now referred to as slow estimation), but since the remaining roots are not rearranged, good candidates may be refused in this step. However, each hit saves precious seconds or even minutes.

The proposed algorithm combines the fast and slow estimation methods as follows:

Step 1' Choose a (real or complex) zero, pole, and a pole-zero pair, which can possibly be eliminated. Rank the three candidates.
Step 2' Make a non-parametric estimation on the cost function by eliminating the best candidate which was not yet tried.
Step 3' Check the result. If the reduced order model is acceptable then accept the decreased order and go to Step 1.
Step 4' Restore the previous order. If there are other candidates to try, go to Step 2'.
Step 1'' Choose a (real or complex) zero, pole, and a pole-zero pair, which can possibly be eliminated. Rank the three candidates.
Step 2'' Re-estimate the model and parameters by eliminating the best candidate, which was not yet tried.
Step 3'' Check the result. If the reduced order model is acceptable then accept the decreased order and go to Step 1'.
Step 4'' Restore the previous order. If there are other candidates to try, go to Step 2''.

In both the fast (') and the slow ('') estimation steps 3 possible candidates are tested: the less significant pole, zero, and pole-zero pair. If the fast estimation step can no longer decrease the order, then it is tried by the slow estimation method. If any roots are eliminated in the slow step, the fast method is tried again with the new root placement. The algorithm terminates if no more pole, zero, or pole-zero pair can be eliminated. More details on the algorithm can be found in Appendix 2.

The choice of the candidates in Step 1' is made by estimating the contribution of the roots to the transfer function. In the case of the poles and zeros, the distance from the measurement band is considered, while in the case of the canceling pairs the distance of roots from each other is considered [8]. The order of the trials is determined by ranking the results (see Appendix 2 for details).

In the present system the candidates in Step 1'' are the same as in Step 1', but their ranking is re-evaluated using the cost function estimations produced in the previous three (unsuccessful) fast estimation steps.
The parameter estimation in Step 2'' uses an iterative parameter estimation algorithm [4,7], for which the initial values must be set. The convergence speed and the quality of the supplied model strongly depend on the initialization. It is obvious, that the elimination of an unnecessary root does not change significantly the transfer function. It can also be assumed that the elimination of such a root does not change too much the numerator and the denominator of the transfer function, either. Based upon this assumption the following initialization scheme is used:
If the estimated fit in Step 2' is 'not too bad' (which means that the candidate root possibly can be eliminated with a small rearrangement of the other roots) then use the reduced-order (non-parametric) numerator and denominator, otherwise (the candidate root can not be eliminated, or major rearrangement is necessary) use the previous good non-reduced-order (non-parametric) numerator and denominator, as initial values.
The decision in Step 3' is made upon the estimated cost function: if it is smaller than a maximal allowed value [8], then the reduced order is accepted otherwise it is rejected. In Step 3'' the quality test is combined: in addition to the cost-function test a correlation test is also made on the residuals [9].
Since the automatic model selection algorithm is a top-down approach, high model orders may appear during the root elimination phase. The numerical stability of the algorithms was ensured by using orthogonal polynomials instead of power polynomials [10]. With this solution, orders of 100/100 were successfully handled.
Step D: The second peeling step produces a somewhat lower quality, lower order model starting from the verified model produced in Step D. (Note that a possible and straightforward further extension of the system is the dynamic order reduction, which enables the user to specify the requirements in terms of the allowed model errors.) In many cases, the loose model is also acceptable for the user, who may decide which model to choose.
The only difference between Step C and D is that the algorithm uses a different qualification method in Step 3''. While in the validated case the peeling algorithm uses the combined cost function and correlation test, in the loose model case only the cost function values are tested against a maximum value determined from the validated model's cost function.

IV. QUALIFICATION OF MODELS

The two models provided by the model selection algorithm must be validated and qualified before being applied. The user may not have the necessary theoretical background and experience to use and interpret the conventional validation tests, such as the theoretical and the observed cost function, and the correlation test. Instead, the proposed system provides some additional easy-to-understand information on the model quality, the presence of the unmodeled dynamics, and the effect of nonlinear distortion.
Based upon the result of the correlation test [9] it can be decided whether the linear dynamic behavior is correctly modeled or if there are some unmodeled linear dynamics present. The correlation test also gives estimation on the presence of (unmodeled) nonlinear distortions. The ratio of the mean error power and the mean of the variance of the transfer function taken over the frequency gives the basis of the qualification of the overall model fit. Upon the provided textual information, even the inexperienced user is able to decide which of the provided models to choose (see Fig. 3).

V. EXPERIMENTAL ILLUSTRATIONS

In this section, three examples are presented to illustrate the performance of the automatic model selection system. The examples cover a wide range of difficult practical problems. The first example is a Brüel&Kjaer passive bandpass filter, where the SNR of the measurement is very high, so the presence of small nonlinear distortions is disturbing. The second example is a second order mechanical system with poor SNR and high nonlinear distortion. The third example is the radial servo system of a CD player, which is a high complexity system. On the figures the following notation is used: measured transfer function (+), estimated transfer function (solid line), variance of the transfer function (dotted line), and the residual error (x with dotted line).

A. Brüel&Kjaer bandpass filter

The measured transfer function and the estimated variance can be seen in Fig. 2, along with the estimated model and the model error. No loose model is provided in this case, since no lower order model can be found to fulfill the requirements in Step D. The information on the validated model can be seen in Fig. 3.

B. Second order mechanical system

A mechanical resonating system (mass, viscous damping, nonlinear spring) was simulated with an electrical circuit. Although all the linear dynamics are modeled, the overall fit is very poor because of the unmodeled nonlinear behavior, as can be seen in Fig. 4. The automatic identification process produced the expected 0/2 order system, for which the qualification message can be seen in Fig. 5.

C. Radial servo system of a CD player

The radial servo system of a Philips CD320/00G was measured. Since open-loop measurement is impossible, the system was operated in closed loop while applying an external excitation signal and the input and output of the servo system was measured. The result of the automatic identification process is a

14/15 order verified model and a 7/8 order loose model, as can be seen in Figs. 6-9.
Note, that in this case a human expert was able to provide somewhat better model quality by increasing the model order (the decrease of the cost function by 30% required a 24/24 order model).

VI. CONCLUSION

The described automatic identification system is able to provide reasonable linear time invariant models for linear and almost linear systems from measurement data, without any human interaction. The input of the system is the measured data, and the output is one strictly validated, and one (somewhat lower quality) 'loose' model. The generated models are also qualified in a form, which is interpretable also by the inexperienced user.

The proposed system was build using advanced frequency domain identification algorithms, and a special care was taken to choose proper excitation signal to enable fully automatic processing, robust data representation to ensure numerical stability. The speed of the algorithm was increased by combining fast, but not precise techniques with slow and reliable methods.
The system after a rough order estimation uses a top-down approach to select the proper model order and parameters. The main advantage of this approach is that the possible local minima of the parameter estimation algorithms can be avoided with greater chance, using appropriate initialization methods.
The performance of the proposed system was also illustrated though examples with real measurement data.

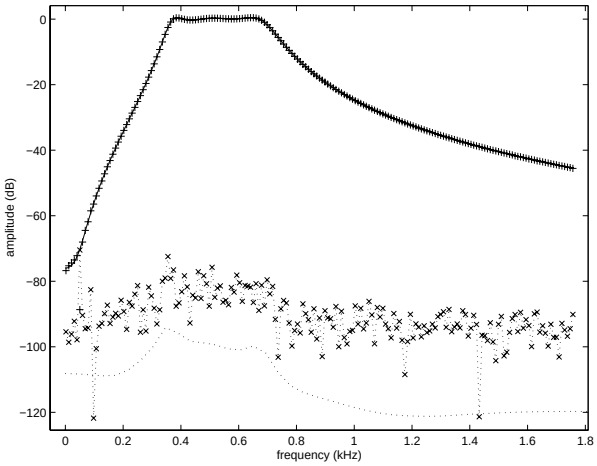

Fig. 2. Measurement data and identification results of the Brüel&Kjaer bandpass filter.

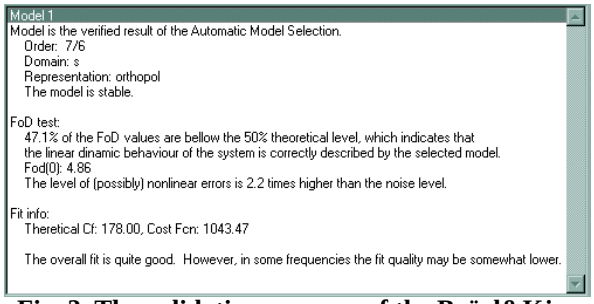

Fig. 3. The validation message of the Brüel&Kjaer bandpass filter.

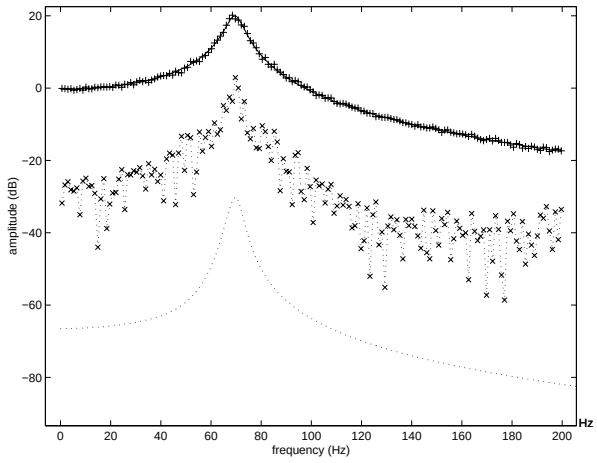

Fig. 4. Measurement data and identification results of the mechanical system.

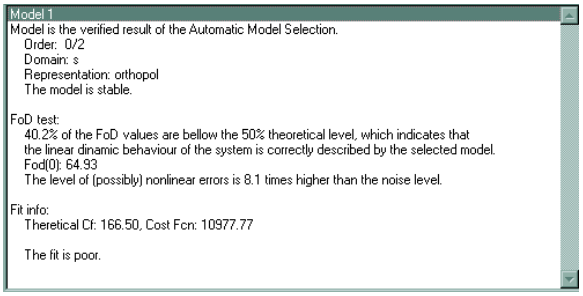

Fig. 5. The validation message of the mechanical system.

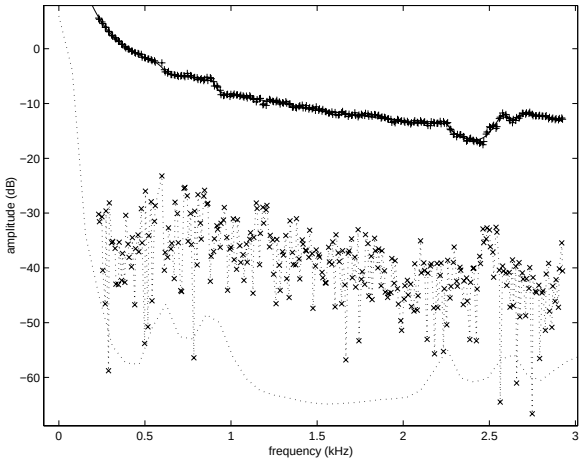

Fig. 6. Measurement data and the validated identification results of the CD player.

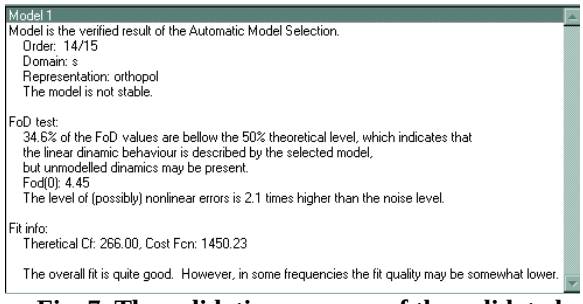

Fig. 7. The validation message of the validated identification results of the CD player.

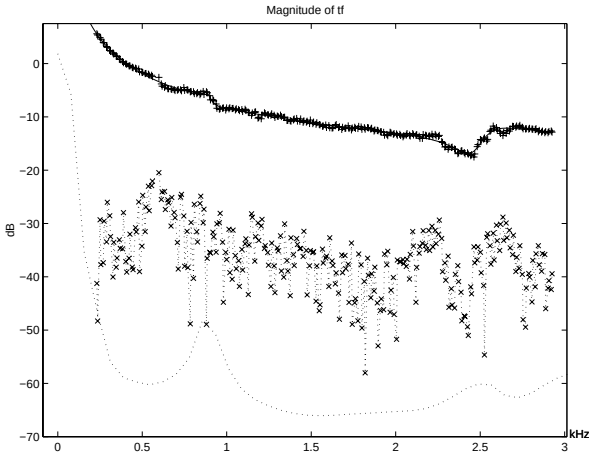

Fig. 8. Measurement data and the loose identification results of the CD player.

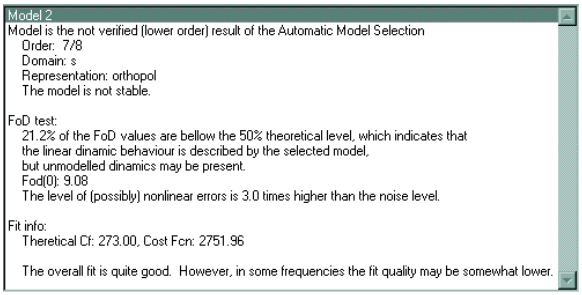

Fig. 9. The validation message of the loose identification results of the CD player.

Appendix 1: Flowchart of the Automatic Model Selection Algorithm

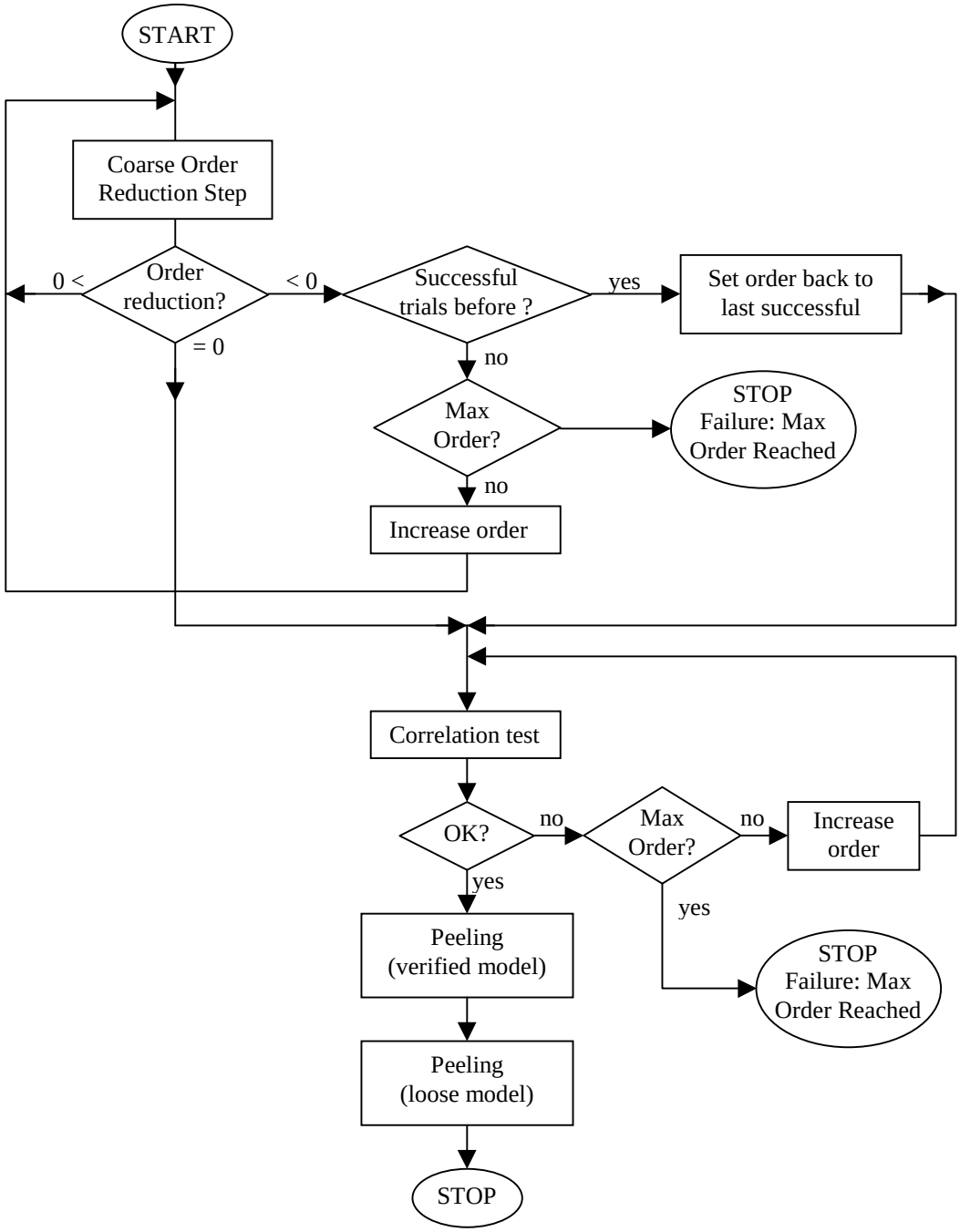

Fig.10. The flowchart of the automatic model selection algorithm.

Appendix 2: The peeling algorithm

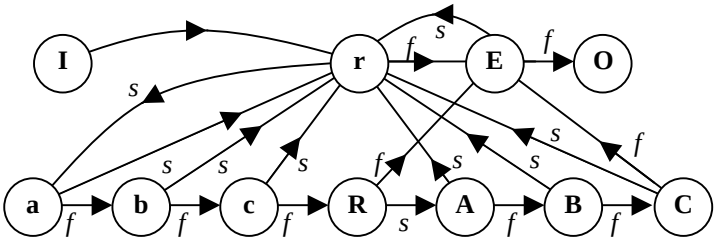

Fig. 11. The state-machine description of the peeling algorithm. s: success, f: failure.

The state-machine description of the peeling algorithm can be seen in Fig. 11. The detailed description of the states is as follows:

I:*Initial state.*
Calculate the maximum allowed cost function Cf_{MAX} based on the initial model's cost function (see Note 1. below).

r: *Ranking of poles, zeros, and canceling pole-zero pairs.*
The best pole, zero, and pole-zero pair candidates are chosen, and the three candidates are ranked. (Note that there may be less than three candidates in case there are no more poles or zeros in the root set. In such cases the corresponding elimination steps **b**, **c**, **B**, and **C** are skipped.)
The significance of poles and zeros is quantified by using the distance between the roots and the measurement band (see in [8], equations 23-24):

$$C_A = \frac{\min_{\omega\in\Omega_C} |j\omega - b|}{\max_{\omega\in\Omega_C} |j\omega - b|},$$

where b is the complex root, and ω is the angular frequency in the measurement band Ω_C.
The significance of canceling pole-zero pairs is calculated by using the relative distance of the roots (see in [8], equations 21-22):

$$C_E = \max_{\omega\in\Omega_E} \left| \frac{\Delta b}{j\omega - b} \right|,$$

where b is the location of the pole, and $b+\Delta b$ is the location of the zero.
In order C_A and C_E be comparable, C_E is transformed to C_E':

$$C_E' = \frac{1}{1+C_E}$$

The ranking of the candidates is based upon the two C_A-like quantities and C_E'. The largest the value C the better the candidate.
Note that the ranking of the candidates has no significant effect on the elimination process, only the speed can be increased if the correct root is chosen first.
Success if order is not 0/0, failure otherwise.

a,**b**,**c**: *Fast root elimination.*
Remove the first (**a**), the second (**b**), or the third (**c**) candidate (if there is any) from the root set. Calculate the new numerator and denominator of the transfer function, using the rest of the roots (there is no estimation to be done). Calculate the cost function for the decreased order system (and store it for later usage).
Failure if the estimated cost function is above Cf_{MAX} or there are no more candidates, success otherwise.

R: *Re-ranking of the three candidates based on the estimated cost functions in **a**, **b**, and **c**.*
Failure, if order is 0/0, success otherwise.

A,**B**,**C**: *Slow root elimination with parameter estimation.*
Remove the first (**A**), the second (**B**), or the third (**C**) candidate (if there is any) from the root set. Estimate the new parameters of the decreased order model.
Initialization of the parameter estimation routine [4,7]:
If the estimated cost function in the corresponding slow elimination step is 'not too bad' (in the present system: $Cf < 10*Cf_{MAX}$) then use the estimated (non-parametric) numerator and denominator for the initialization. Otherwise, use the latest validated model's numerator and denominator for that purpose. The initial weight vector is calculated as follows:

$$W^2(w_k) = \sigma_U^2(w_k) * |N(w_k)|^2 + \sigma_Y^2(w_k) * |D(w_k)|^2 - 2real\left(D(w_k)\overline{N}(w_k)\sigma_{YU}^2(w_k)\right),$$

where N and D are the non-parametric numerator and denominator values, and the upper bar is the conjugate operator.
The model parameters are calculated using different methods. In the current system the Total Least Squares (TLS) and Weighted Generalized Total Least Squares (WGTLS) methods are used [12, 13]. The initial model with the smaller cost function is used to initialize the parameter estimation routine.
Calculate the cost function of the estimated model, and also perform correlation test on the residuals.
Perform model quality test:
When peeling is used to find validated models, the quality check is the following:

1. Loose cost function test:
   Q1 = (Cf<$2Cf_{MAX}$) (see Note 1. below).
2. Correlation test on the residuals
   Q2 = (At least 25% of the correlation values are below the 50% theoretical level)
3. Loose test on the central 10% of the lags part:
   Q3 = (At least 20% of the central correlation values are below the 95% theoretical level)
   Q4 = (At least 60% of the central correlation values are below the 99.5% theoretical level)

Q = (*Q1* and *Q2* and *Q3* and *Q4*).
When peeling is used to find the loose model, the quality test is the following:
Q = (Cf<Cf_{MAX}) (see Note 1. below).
Failure if the result of the model quality test *Q* is false, or there are no more candidates. Success otherwise.

E: *Parameter estimation without root elimination*.
In this step the model parameters are re-estimated, if the last successful elimination step was not **A**, **B**, **C**, or **E** (in these cases end with failure). This step is necessary to compute the parameters of the reduced order model, if the last order reduction was achieved in a fast step.
The algorithm is the same as in **A**, **B**, **C**, except for no roots are eliminated before the estimation.

O: *Final state*.

Note 1. The Cf_{MAX} value is calculated the following ways depending on the purpose of the peeling:
During validated model estimation the maximum allowed cost function is [8]:

$$Cf_{MAX} = Cf_{high\,order} + 2\sigma_{Cf},$$

which is calculated in the initial phase (**I**) based on the initial (input) model, and after each successful slow estimation steps (**A**, **B**, **C**, or **E**) based on the estimated model.
During loose model estimation

$$Cf_{MAX} = 2Cf_{validated} - Cf_{noise},$$

which is calculated only in the initial phase (**I**), and where $Cf_{validated}$ is the cost function of the validated input model. See also Note 2.
Note 2. The total cost function is the following:

$$Cf_{TOTAL} = \sum_{w} \frac{|e(w)|^2}{W(w)},$$

where *e(w)* is the residual and *W(w)* is the applied weight [4,7].
The contribution of the noise to the cost function can be calculated by

$$Cf_{NOISE} = F - \frac{P}{2},$$

where *F* is the number of frequencies, and *P* is the number of free parameters. The Minimum Description Length (MDL) cost function is defined as follows:

$$Cf_{MDL} = Cf_{TOTAL} + \frac{P}{2}\ln(4F).$$

In the description of the algorithm, the term 'cost function' and the notation Cf always refer to Cf_{MDL}. [1,11]

References

[1] L. Ljung (1999). System Identification :Theory for the User. Prentice Hall, Englewood Cliffs, NJ.
[2] T. Söderström and P. Stoica (1989). System Identification. Englewood Cliffs, NJ : Prentice-Hall, pp. 256.
[3] J. P. Norton (1986). An Introduction to Identification, Academic, London, England.
[4] J. Schoukens and R. Pintelon (1991). Identification of Linear Systems : A Practical Guideling to Accurate Modeling. Pergamom, Oxford.
[5] L. Ljung (1986). System Identification Toolbox for Use with Matlab. The MathWorks, Inc., Sherborn, A.
[6] I. Kollár (1994). Frequency-Domain System Identification Toolbox for Use with Matlab, The Mathworks, Natick, M. A.
[7] J. Schoukens, R. Pintelon, G. Vandersteen and P. Guillaume (1997). Frequency-domain system identification using non-parametric noise models estimated from a small number of data sets. Automatica, Vol. 33, no.6, pp. 1073-1086.
[8] Y. Rolain, J. Schoukens, and R. Pintelon (1997). Order estimation for linear time-invariant systems using frequency domain identification methods. IEEE Trans. on Automatic Control, vol. 42, No. 10, pp. 1408-1417.
[9] J. Schoukens, T. Dobrowiecki, and R. Pintelon (1998). Identification of linear systems in the presence of nonlinear distortions. A frequency domain approach. IEEE Trans. on Automatic Control, Vol. 43, No.2, pp. 176-190.
[10] Y. Rolain, R. Pintelon, K. Q. Xu, and H. Vold (1995). Best conditioned parametric identification of transfer function models in the frequency domain. IEEE Trans. on Automatic Control, vol. AC-40, no. 11, pp. 1954-1960.
[11] J. Rissanen (1978). Modeling by shortest data description. Automatica, vol. 14, pp. 465-471.
[12] S. Van Huffel and J. Vandewalle (1991). The total least squares problem: computational aspects and analysis. SIAM, Frontiers in Applied Mathematics, Philadelphia.
[13] R. Pintelon, P. Guillaume, G. Vandersteen, and Y. Rolain (1998), Analysis, development, and applications of TLS algorithms in frequency domain system identification. SIAM J. Anal. Appl, vol. 19, No.4, pp 983-1004.