Independent Scaling of a Delay in Frequency Domain System Identification

Jeroen Strompf, István Kollár and Yves Rolain

Abstract—***Numerical conditioning and identifiability are key factors in frequency domain system identification. Independent frequency scaling of a delay in transfer function models is studied in view of these two topics, both for ordinary s-domain descriptions as well as for descriptions with orthogonal polynomials in the Laplace-plane. Algorithms are proposed to device independent scaling during estimation sessions.***

Keywords—**System identification, parameter estimation, frequency domain, numerical conditioning, delay, scaling, transfer function.**

I. Introduction

This paper discusses some numerical and identifiability problems in frequency domain system identification of linear time-invariant (LTI) single input single output (SISO) systems with delay. The identification uses a maximum likelihood estimator [1]. The systems under estimation are represented as rational transfer functions, extended with a delay [2], [3], using either powers of s, or descriptions with orthogonal polynomials in the Laplace-plane [4]. The algorithm implemented in the estimator solves iteratively a nonlinear least squares problem.

Numerical problems occur when the numerical conditioning of the equations is worse than what the finite machine precision can handle. Numerical problems are exhibited by significant roundoff errors, discontinuities in functions, and mathematical inconsistencies in equations. Conditioning is expressed here by the 2-norm condition number of the equations ([5], pp. 79-80).

Identifiability problems occur when the algorithm is unable to estimate an accurate model. Identifiability problems may be exhibited by poor results during verification and simulation of the model, as well as being the cause of numerical instabilities. Identifiability problems need a careful reconsideration of the whole identification procedure.

These two topics are discussed in relation to independent scaling of a delay. First the properties of independent scaling are considered. After this, practical applications are discussed. All experiments are done by use of the *Frequency Domain System Identification Toolbox* for MATLAB with some extensions. This paper is a revised version of [6].

J. Strompf is with the Systems and Control Laboratory, Delft University of Technology, The Netherlands. PO Box 4, NL-2600 GA Delft, The Netherlands. Phone: +31 23 563-6902, fax: +31 23 564-1005, email: j.g.strompf@et.tudelft.nl

I. Kollár is with the Department of Measurement and Instrument Engineering, Technical University of Budapest, Hungary. H-1521, Phone: +36 1 463-1774, fax: +36 1 463-4112, email: kollar@mmt.bme.hu

Y. Rolain is with Department ELEC, Vrije Universiteit Brussel, Belgium. Pleinlaan 2, B-1050. Phone: +32 2 629-2944, fax: +32 2 629-2850, email: yrolain@vnet3.vub.ac.be

II. Preliminaries

We briefly discuss the cost function and related quantities which will be used in the sequel of this paper.

A. Cost Function

We assume here that the input and output observation noises are all independent and uncorrelated at different frequencies, and they have circularly symmetric Gaussian distributions. The cost function can be written for uncorrelated inputs and outputs as [1], [7]:

$$C(\mathbf{P}) = \frac{1}{2}\sum_{k=1}^{F} \frac{|e^{-j\omega_k T_d} N(j\omega_k,\mathbf{P})X_{mk} - D(j\omega_k,\mathbf{P})Y_{mk}|^2}{\sigma_{yk}^2|D(j\omega_k,\mathbf{P})|^2 + \sigma_{xk}^2|N(j\omega_k,\mathbf{P})|^2}, \tag{1}$$

where $N(j\omega,\mathbf{P})$ and $D(j\omega,\mathbf{P})$ are the numerator and the denominator of the transfer function, respectively. Measurements are made at angular frequencies ω_k, $k = 1, 2 \ldots F$. X_{mk} and Y_{mk} are corresponding input and output spectra, contaminated by noise. σ_{xk}^2 and σ_{yk}^2 are corresponding variances — halves of the total variances —, defined by:

$$\mathrm{var}\{N_{xk}\} = 2\sigma_{xk}^2 \qquad \text{and} \qquad \mathrm{var}\{N_{yk}\} = 2\sigma_{yk}^2, \tag{2}$$

where N_{xk} and N_{yk} are the complex input and output noises, respectively (errors-in-variables model). The unknown parameters are the coefficients of the transfer function, given by the vector $\mathbf{P}$. T_d is a delay and can be included in $\mathbf{P}$ or not.

B. Error Vector

The error vector is linked to the cost function by:

$$C(\mathbf{P}) = \frac{1}{2}\sum_{k=1}^{F} |\mathrm{ESm}(k)|^2 = \frac{1}{2}\mathbf{ESm}^{\mathrm{T}}\mathbf{ESm} \tag{3}$$

with

$$\mathrm{ESm}(k) = \frac{e^{-j\omega_k T_d/2} N(j\omega_k,\mathbf{P})X_{mk} - e^{j\omega_k T_d/2} D(j\omega_k,\mathbf{P})Y_{mk}}{\sqrt{\sigma_{yk}^2|D(j\omega_k,\mathbf{P})|^2 + \sigma_{xk}^2|N(j\omega_k,\mathbf{P})|^2}}. \tag{4}$$

The delay term is split into two parts for sake of numerical stability. This does not effect the value of $C(\mathbf{P})$. We will treat the Jacobian of this error vector,

$$\mathbf{J} = \frac{\partial \mathbf{ESm}}{\partial \mathbf{P}}. \tag{5}$$

Most iteration schemes directly operate on **J** to minimize the cost function, and numerical conditioning of the overall system depends on **J**.

C. Singular Values of **J**

The system equations determine the transfer function coefficients up to a common multiplying factor only. An additional equation has to be introduced to arrive at a full set of equations. Two commonly used possibilities are: fixing a nonzero coefficient, or fixing the norm of the parameter vector to 1. According to experience, fixing the norm of the parameter vector often results in superior results. Therefore, in this paper we only deal with this.

The number of singular values (SV's) of **J** equals the number of parameters under estimation. The close-to-zero SV which corresponds to the initially missing normalizing equation has no importance for us. Its value is only determined by roundoff errors. Therefore, the smallest SV will be disregarded in the sequel of the discussion.

The magnitude of a SV corresponds to the reciprocal of the precision with which an underlying parameter combination is estimated. Whenever the magnitude of a SV is not significant above its variance, the correspondence between this SV and its assumed physical background is questionable. Small SV's with doubtful physical significance may also occur when there are large SV's present: A large SV causes a large roundoff level. SV's smaller than this roundoff level will be lost. These phenomena are taken into account by the algorithm during estimation sessions: reciprocals of too small SV's are set to zero when inverses are computed. It may indicate a numerical problem as well as an identifiability problem. When the algorithm is forced to set SV's to zero, the experimenter is warned about this.

Small SV's may also be set to zero when **J** is bad conditioned, according to:

$$SV_{min} < freepar * SV_{max} * eps, \tag{6}$$

where SV_{min} and SV_{max} are the smallest and the largest SV of **J**, respectively. $freepar$ is the number of parameters under estimation, and eps the floating point relative accuracy, roughly $2.22 \cdot 10^{-16}$ on machines with IEEE floating point arithmetic.

D. Global Scaling

One way to improve numerical conditioning and thereby, enlarging the range of identifiability, is by scaling of every frequency related quantity by the scaling frequency ω_s, as discussed in [1] and [7]. This means that ordinary transfer function models are represented internally as:

$$H(j\omega) = e^{-j(\frac{\omega}{\omega_s})(T_d \omega_s)} \cdot \frac{(b_0\omega_s^0)(j\frac{\omega}{\omega_s})^0 + \ldots + (b_{nn}\omega_s^{nn})(j\frac{\omega}{\omega_s})^{nn}}{(a_0\omega_s^0)(j\frac{\omega}{\omega_s})^0 + \ldots + (a_{nd}\omega_s^{nd})(j\frac{\omega}{\omega_s})^{nd}}, \tag{7}$$

with a and b being the coefficients of the numerator and denominator, respectively. nn and nd are the orders of the numerator and denominator, respectively.

However, when starting identification, no usable information is at hand about poles and zeros. This is why usually a simple rule of thumb is used, based on the excitation band which is available from the Fourier data: the center angular frequency of the excitation band, which usually coincide with the band of interest:

$$\omega_s = \frac{\omega_{min} + \omega_{max}}{2}, \tag{8}$$

where ω_{min} and ω_{max} are the lowest and highest angular frequencies present in the excitation band, respectively. Global scaling is applied in all systems studied here.

III. Independent Scaling of the Delay

Independent scaling of the delay transforms the internal representation of the transfer function to

$$H(j\omega) = e^{-j(T_{ds}\frac{\omega}{\omega_s})(\frac{T_d}{T_{ds}}\omega_s)} \cdot \frac{(b_0\omega_s^0)(j\frac{\omega}{\omega_s})^0 + \ldots + (b_{nn}\omega_s^{nn})(j\frac{\omega}{\omega_s})^{nn}}{(a_0\omega_s^0)(j\frac{\omega}{\omega_s})^0 + \ldots + (a_{nd}\omega_s^{nd})(j\frac{\omega}{\omega_s})^{nd}}, \tag{9}$$

with T_{ds} the delay scaling factor.

The trajectories of the singular values of **J** are studied as function of the scaling of the delay.

A. General Behavior

The general behavior of the SV's is illustrated in Fig. 1. This is an artificially generated first order lowpass filter with a corner frequency at 0.8 Hz. The model is excitated in the range from 0 to 1000 Hz. The left plot corresponds to ordinary s-domain descriptions. The right plot corresponds to descriptions with orthogonal polynomials [4]. Note that with orthogonal polynomials, SV's are closer to each other. Therefore, these descriptions yield inherently better conditioned systems. The dominating feature of these figures is a number of horizontal lines, crossed by a slanted one. Only one SV changes as function of the delay scaling factor. The slope of the slanted trajectory is 20dB/decade (proportionality).

Formally it cannot be stated that there exists something as a SV associated to the delay. SV's are associated to

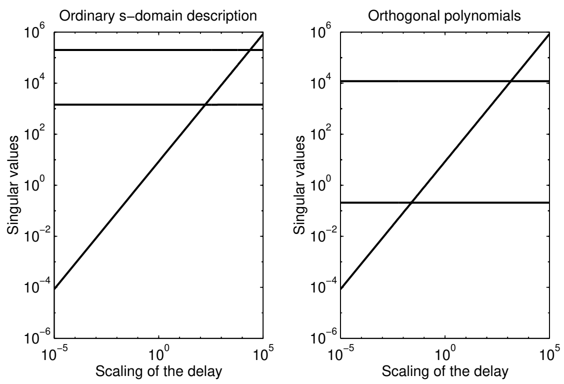

Fig. 1. Singular values of **J** for a first order lowpass filter

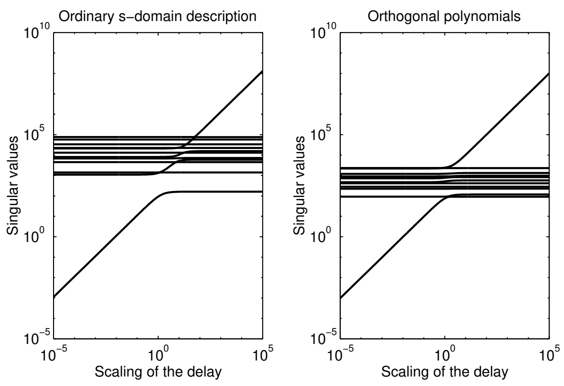

Fig. 2. Singular values of **J** for a 4/6 robot arm model

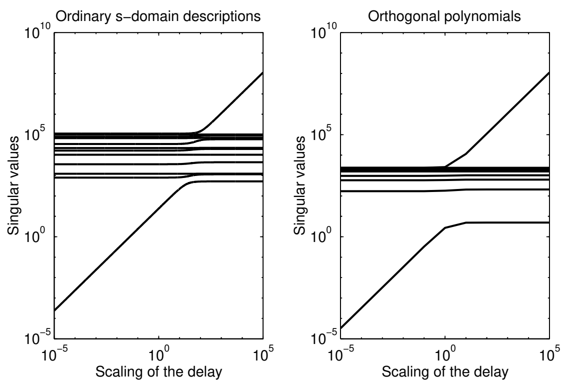

Fig. 3. Singular values of **J** for a 4/6 bandpass filter

combinations of parameters, not to individual ones. Nevertheless, the dominating feature of the figures is that the delay only interacts with one SV at a time. Therefore, we will loosely state these two entities to be usually associated to each other.

B. Distortions in the Trajectories

The behavior of the SV's as stated before is highly idealized and may seldom occur in practice. Usually, distortions occur due to reduced accuracy in the estimation of one or more parameters. This can even lead to a failure in identification.

Reduced accuracy is exhibited by 'bandgaps' in the figures. The slanted line does not cross such bands, but branches off into a horizontal line. This is illustrated in Figs. 2 and 3. The former is a 4/6 model of a robot arm. The latter a 4/6 model of an octave bandpass filter. Both models are based on measurements.

IV. Application of Independent Scaling

Independent scaling of the delay can be applied as an analytical tool and as a tool for dealing with certain numerical and identifiability problems. Its purpose in the latter case is, to bring the SV corresponding to the delay to among the other SV's.

A. Numerical Conditioning

Experiments show that scaling of the delay is often capable of improving the numerical conditioning of arbitrary systems, as it does for the models in the left plots in Figs. 2 and 3. By means of this, independent scaling can enlarge the region in which estimates are numerically acceptable, and reduces the number of cases in which the experimenter should consider to change something more fundamental in his experimental setup.

B. Identifiability

When there is a wide spread in precision among the parameter combinations under estimation, the algorithm may fail to identify all parameters. A wide spread in precision can occur between the delay and the transfer function parameters. Identifiability of the delay seems not to depend on the actual magnitude of the delay.

Fig. 4 gives an example of models where the delay can only be identified when independent scaling is applied. The left plot corresponds to the discussed lowpass filter. The right plot corresponds to a second order highpass filter with two poles close to 1. Both models are excitated in a very broad band from 0 to 10.000 Hz, at 1000 equally spaced frequencies. The delays are estimated with high precision. The transfer function parameters can not be estimated correctly without independent scaling of the delay.

Two cases in which the estimator fails, with or without scaling of the delay, are depicted in Fig. 5. Both models are based on the discussed lowpass filter. The left plot is generated by excitating the system in a too narrow frequency band from 0 to 0.01 Hz. A delay introduces an accumulating phase shift in response to excitation. Therefore, proper estimation of a delay requires excitation in a very broad band, as it is done for the models in Fig. 1. The transfer function parameters corresponding to the left plot in Fig. 5 are estimated accurately. The delay not. Identifiability problems with the delay seem to coincide with the phenomenon that the slanted line branches off into the *lowest* horizontal line.

The right plot in Fig. 5 corresponds to identification of the lowpass filter in a proper frequency band, but as an overmodeled 4/3 system. The SV's are distributed in two distinct bands. Due to bad numerical conditioning, the algorithm is forced to disregard the lowest two SV's. It may give a hint why scaling of the delay does not improve in this situation: if parameters cannot be identified at all, scaling makes no difference.

C. As an Analytical Tool

When independent scaling is included in identification procedures, it can provide some insight about the system under estimation and serve as a check for consistency. The experimenter can compare the suggested scaling factor with his expectation from the physical background of the experiment.

V. How to Apply Independent Scaling

We propose algorithms to determine an optimal delay scaling factor, on-line during estimation sessions, as well as off-line.

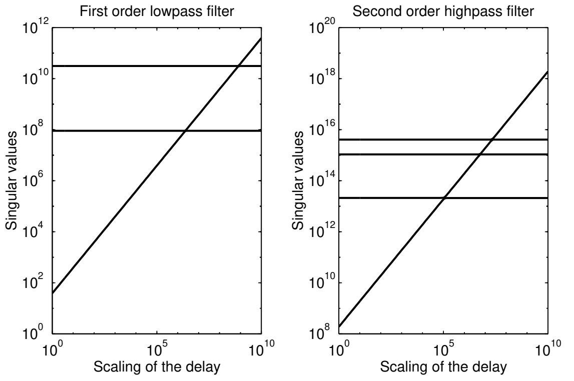

Fig. 4. Two models where independent scaling is necessary

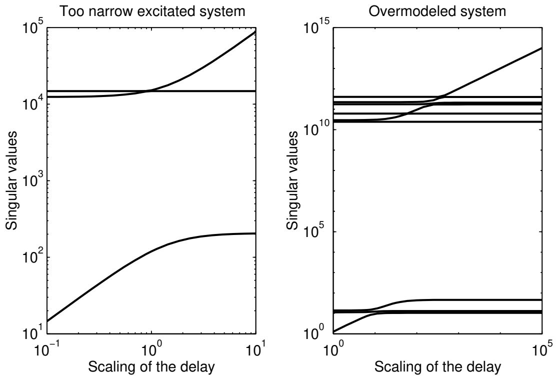

Fig. 5. Two unidentifiable lowpass filter models

A. On-line Estimation

The purpose of on-line estimation of a proper delay scaling factor is to device an algorithm that can be incorporated in estimation session procedures, as it is for the rule of thumb deviced for global scaling.

Best results are obtained by analysis of the matrix $\mathbf{J}$. $\mathbf{J}_{fix}$ is this matrix with its last column removed. This column corresponds to the delay. The SV's of $\mathbf{J}$ and $\mathbf{J}_{fix}$ are paired and only one SV of $\mathbf{J}$ remains alone. This one corresponds to the delay. Based on this value, the algorithm can propose a scaling factor which brings this SV to among the others. Matching of $\mathbf{J}$ and $\mathbf{J_{fix}}$ can be done with the so-called Hungarian method. It is based on looking for alternating paths in bipartite graphs ([7], pp. 2-69 … 2-70).

B. Off-line Estimation

In cases where the above proposed algorithm fails, offline estimation of the delay can be applied. A proper scaling factor corresponds to a minimal ratio between the largest and smallest SV, present in $\mathbf{J}$. There is a global minimum in all systems studied.The ratio decreases monitonically at both sides of this minimum, towards it. Based on these observations, robust algorithms can be used which make use of routines to search an ordered table, like *hunt* and *bisectioning* algorithms ([**?**], pp. 89-92). At least three estimation sessions under different scaling factors have to be performed, in order to collect $\mathbf{J}$ matrices. An initial value can be supplied by the previous discussed on-line algorithm.

VI. Conclusions

This paper presents some theoretical and practical results in the field of frequency domain system identification.

Experiments show that a delay in frequency domain models can be scaled independently of the transfer function to which it is associated. This is investigated for ordinary s-domain descriptions, as well as for descriptions with orthogonal polynomials in the Laplace plane. Only one SV associated to the system equations is significantly affected by scaling of the delay. The relation is predominantly proportional.

The proportional relation between the scaling factor and the apparently uniquely associated SV makes independent scaling of the delay a transparent tool. Algorithms are proposed to device independent scaling. Independent scaling can serve for a number of purposes:

1. To improve numerical conditioning during estimations of arbitrary systems.
2. To enhance identifiability in some systems with an extreme spread in accuracy among the estimated parameters.
3. As an analytical tool to provide the experimenter with some insight about the system under estimation.

References

[1] J. Schoukens and R. Pintelon, *Identification of Linear Systems: a Practical Guideline for Accurate Modeling*, Pergamon Press, London, 1991.

[2] R. Pintelon, "Phase correction of linear time invariant systems with digital allpass filters", *IEEE Transactions on Instrumentation and Measurement*, vol. IM-39, no. 2, pp. 324–330, 1990.

[3] R. Pintelon and L. van Biesen, "Identification of transfer functions with time delay and its application to cable fault location", *IEEE Transactions on Instrumentation and Measurement*, vol. IM-39, no. 3, pp. 479–484, 1990.

[4] Y. Rolain, R. Pintelon, K.Q. Xu, and H. Vold, "Best conditioned parametric identification of transfer function models in the frequency domain", *IEEE Transactions on Automatic Control*, vol. 40, no. 11, pp. 1954–60, November 1995.

[5] G. H. Golub and C. F. van Loan, *Matrix Computations*, The John Hopkins University Press, London, 1989.

[6] J. Strompf, I. Kollár, and Y. Rolain, "Independent scaling of a delay in frequency domain system identification", in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference*, Ottawa, Canada, May 19-21 1997, pp. 759–764.

[7] I. Kollár, *Frequency Domain System Identification Toolbox for use with Matlab*, The Mathworks Inc., Natick, MA, 1994.

[8] W. H. Preuss, B. P. Flamery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes – The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1986.