Independent Scaling of a Delay in Frequency-Domain System Identification

Jeroen Strompf, *Member, IEEE*, István Kollár, *Fellow, IEEE*, and Yves Rolain

***Abstract*— In frequency-domain identification of systems with delay, the delay can be scaled independently of the transfer function to which it is associated. Practical examples show that in most situations, the delay is properly scaled with the commonly used scaling. Examples have been provided where independent scaling can be useful to improve numerical conditioning and to enhance identifiability. Algorithms are proposed for these purposes and for system analysis.**

***Index Terms*— Delay estimation, delay systems, frequency-domain analysis, identification, numerical stability, parameter estimation, transfer functions.**

I. Introduction

THIS paper proposes *independent scaling of a delay* as a new method to avoid 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 [1]–[4], using either powers of *s*, or descriptions with orthogonal polynomials in the Laplace plane [5]. The algorithm implemented in the estimator solves these system equations iteratively as 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. They occur as significant rounding errors, discontinuities in functions, and mathematical inconsistencies in equations. Conditioning is expressed here by the 2-norm condition number of the system equations [6, pp. 79–80]. In principle, numerical problems can be avoided with the use of proper algorithms and data.

Identifiability problems occur when the algorithm is unable to estimate an accurate model. They give rise to 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.

After establishing the properties of independent scaling, we will discuss practical applications and proposed algorithms. All experiments have been done using the *Frequency-Domain System Identification Toolbox* for MATLAB [4] with some extensions.

Manuscript received June 1, 1997; revised February 1, 1998.

J. Strompf is with the Systems and Control Laboratory, Delft University of Technology, Delft, The Netherlands (e-mail: j.g.strompf@et.tudelft.nl).

I. Kollár is with the Department of Measurement and Instrument Engineering, Technical University of Budapest, Budapest, Hungary, H-1521 (e-mail: kollar@mmt.bme.hu).

Y. Rolain is with Department ELEC, Vrije Universiteit Brussel, Brussels, B-1050 Belgium (e-mail: yrolain@vnet3.vub.ac.be).

Publisher Item Identifier S 0018-9456(98)06138-5.

II. Preliminaries

We briefly discuss the cost function and related quantities.

A. Cost Function

In frequency-domain system identification, we usually assume that the input and output observation noises are all independent at different frequencies, and input noises are all independent of output noises. These noises have circularly symmetric Gaussian distributions. The cost function can be written for uncorrelated inputs and outputs as [1], [4]

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

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

$$\text{var}\{N_{xk}\} = 2\sigma_{xk}^2 \quad \text{and} \quad \text{var}\{N_{yk}\} = 2\sigma_{yk}^2 \tag{2}$$

where N_{xk} and N_{yk} are the complex input and output noises (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 $\mathbf{P}$ for estimation, or not.

B. Error Vector

The error vector is linked to the cost function by

$$C(\mathbf{P}) = \tfrac{1}{2}\sum_{k=1}^{F} |\text{ESm}(k)|^2 = \tfrac{1}{2}\mathbf{ESm}^{\text{H}}\mathbf{ESm} \tag{3}$$

with

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

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

the Jacobian of this error vector

$$\mathbf{J} = \begin{bmatrix} \mathrm{Re}\dfrac{\partial \mathbf{ESm}}{\partial \mathbf{P}} \\ \mathrm{Im}\dfrac{\partial \mathbf{ESm}}{\partial \mathbf{P}} \end{bmatrix} \tag{5}$$

since 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 set of equations of full rank. Two commonly used possibilities are: fixing a nonzero coefficient, or fixing the norm of the parameter vector, excluding the delay, to 1. Fixing a nonzero coefficient is only appropriate in the few situations when an *a priori* coefficient is known, e.g., from the physical background of the system. However, in the general situation, such *a priori* information is not at hand, and therefore, we only deal with systems where the norm is fixed to 1.

The number of singular values (SV's) of **J** equals the number of parameters under estimation. The smallest SV, associated to the normalizing equation has no physical meaning and its value is only determined by rounding errors. We will disregard this SV in the remainder of this discussion.

The magnitude of a SV approximately corresponds to the reciprocal of the precision with which an underlying parameter combination is estimated. However, there are situations in which the correspondence between SV's and their physical meaning becomes questionable:

- when the magnitude of a SV is not significantly above its variance;
- when small SV's have the same order of magnitude as the rounding error of large SV's.

Algorithmically, reciprocals of small SV's are set to zero and thus disregarded during estimation, when **J** is conditioned badly according to

$$SV_{\min} < \mathrm{freepar} \cdot SV_{\max} \cdot \epsilon \tag{6}$$

where $SV_{\min}$ and $SV_{\max}$ are the smallest and the largest SV of **J**, respectively, $\mathrm{freepar}$ is the number of parameters under estimation, and ϵ is the computer precision, roughly $2.22 \cdot 10^{-16}$ on machines with IEEE double floating point arithmetic. When SV's are disregarded, this may indicate a numerical or identifiability problem, and the experimenter is warned about this.

D. Global Scaling

One way to improve numerical conditioning and thereby to enlarge the range of identifiability, is by scaling every frequency-related quantity by the scaling frequency ω_s [1], [4]. This means that models are represented internally as

$$H(j\omega) = e^{-j(\omega/\omega_s)(T_d\omega_s)} \cdot \frac{(b_0\omega_s^0)\left(j\dfrac{\omega}{\omega_s}\right)^0 + \cdots + (b_{nn}\omega_s^{nn})\left(j\dfrac{\omega}{\omega_s}\right)^{nn}}{(a_0\omega_s^0)\left(j\dfrac{\omega}{\omega_s}\right)^0 + \cdots + (a_{nd}\omega_s^{nd})\left(j\dfrac{\omega}{\omega_s}\right)^{nd}} \tag{7}$$

with a_i and b_i the coefficients of the numerator and the denominator, and nn and nd the orders of the numerator and the denominator.

However, at the start of the identification process, no usable information is at hand about poles and zeros. This is why usually a simple rule of thumb is used: the center angular frequency of the excitation band, which usually coincides with the band of interest

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

where $\omega_{\min}$ and $\omega_{\max}$ are the lowest and highest angular frequencies present in the excitation band. Global scaling is applied in all models studied here.

III. Independent Scaling of a Delay

Independent scaling of a delay transforms the internal representation of a model to

$$H(j\omega) = e^{-j[T_{ds}(\omega/\omega_s)][(T_d/T_{ds})\omega_s]} \cdot \frac{(b_0\omega_s^0)\left(j\dfrac{\omega}{\omega_s}\right)^0 + \cdots + (b_{nn}\omega_s^{nn})\left(j\dfrac{\omega}{\omega_s}\right)^{nn}}{(a_0\omega_s^0)\left(j\dfrac{\omega}{\omega_s}\right)^0 + \cdots + (a_{nd}\omega_s^{nd})\left(j\dfrac{\omega}{\omega_s}\right)^{nd}} \tag{9}$$

with T_{ds} the delay scaling factor.

The trajectories of the SV's have been studied as functions of the scaling of the delay. These SV's are obtained from **J** at the solution; at the last iteration of an identification.

A. Idealized Behavior

The idealized behavior of the SV's is illustrated in Fig. 1 for an artificially generated first order lowpass filter with a corner frequency at 0.8 Hz. This model is excited in the range from 0 Hz to 1000 Hz at 1001 equally spaced frequencies. Variances of the input and output noises are 10^{-7}. The left plot corresponds to ordinary *s*-domain descriptions and the right plot to descriptions with orthogonal polynomials [5]. Note that in the latter case, SV's are closer to each other. Therefore, these descriptions usually yield inherently better conditioned systems. The dominating feature of these plots is a number of horizontal lines, crossed by a slanted one. Only one SV changes as function of the delay scaling factor, with a slope of 20 dB/decade (proportionality).

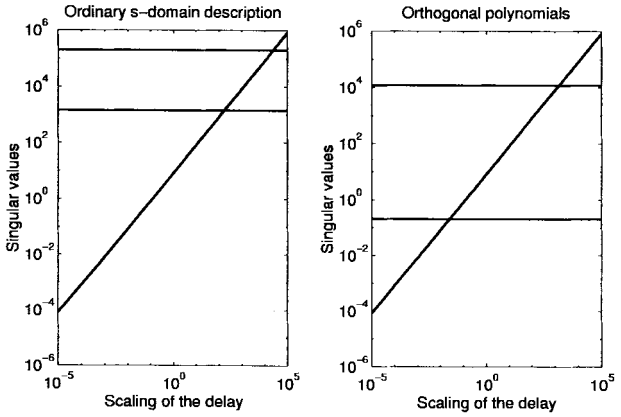

Fig. 1. Singular values of **J** for a first order lowpass filter.

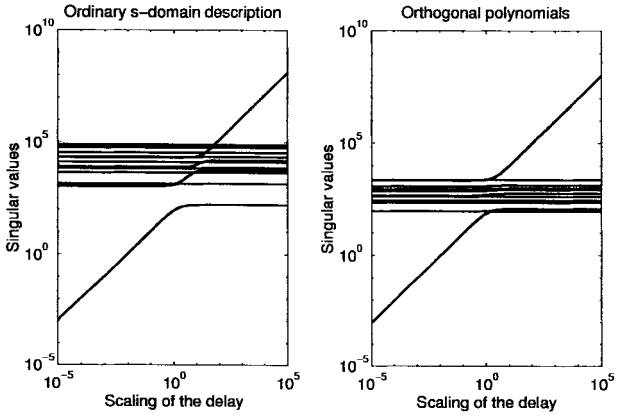

Fig. 2. Singular values of **J** for a 4/6 robot arm model.

Formally it cannot be stated that there exists something as a SV corresponding to the delay. SV's correspond to 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 corresponding to each other.

B. Distortions in the Trajectories

The behavior of the SV's as stated before occurs inexactly in practice. Usually, distortions appear due to reduced accuracy in the estimation of one or more parameters. 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 for two models, based on measurements: a 4/6 model of a robot arm, and a 4/6 model of an octave bandpass filter.

This reduced accuracy mostly happens to identification of the delay. A delay introduces an accumulating phase shift in response to excitation and therefore, proper estimation requires a very broad band. The idealized behavior of the model in Fig. 1 could only be obtained by using such a broad band. Identifiability problems with the delay seem to coincide with the phenomenon that the slanted line branches off into the *lowest* horizontal line.

We cannot fully explain why reduced accuracy is exhibited by bandgaps. Our guess is that probably the structure of **J**

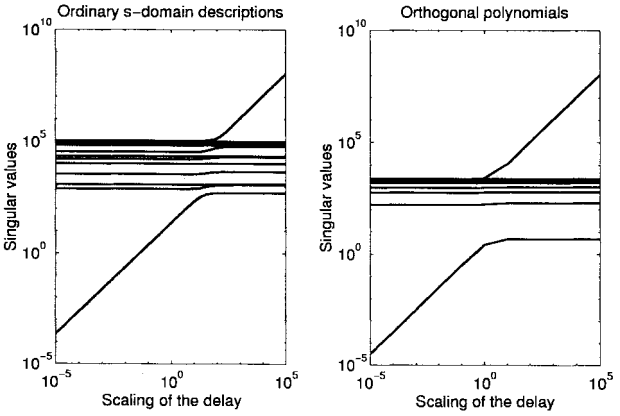

Fig. 3. Singular values of **J** for a 4/6 bandpass filter.

changes at these places sometimes, due to slight variations of the scaled elements, but we don't have sufficient proof as yet.

IV. Application of Independent Scaling

Independent scaling of a 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, thereby reducing conditioning problems.

A. Numerical Conditioning

Experiments show that scaling of a delay is often capable of improving numerical conditioning for arbitrary models, 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 his experimental setup more fundamentally.

An application may be in calibration of measurement systems. Modern calibration procedures need the best available accuracy from the measured data, and sophisticated identification schemes may break down when numerical stability is in stake. This can be especially important when digital signal processors (DSP's) are used, e.g., with single precision floating point, or fixed point arithmetic.

B. Identifiability

The algorithm may fail to identify all parameters when there is a wide spread in precision among the parameter combinations under estimation, as it can occur between a delay and the transfer function parameters.

Fig. 4 gives examples of models which can only be properly identified when independent scaling is applied. Both models are artificial in the sense that they have been excited in very broad bands with low frequency resolutions, in combination with very low variances of the noise levels. More realistic examples can probably be found when models of higher order are examined.

The left plot corresponds to the discussed lowpass filter. It is excited in the band from 0 Hz to 10^6 Hz at 11 equally spaced frequencies. Variances of the input and output noises are

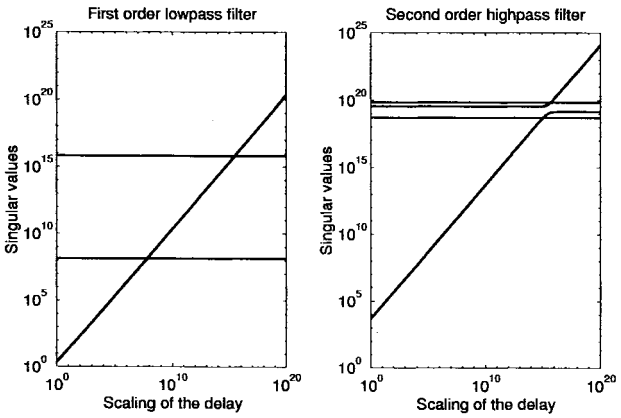

Fig. 4. Two models for which independent scaling is necessary.

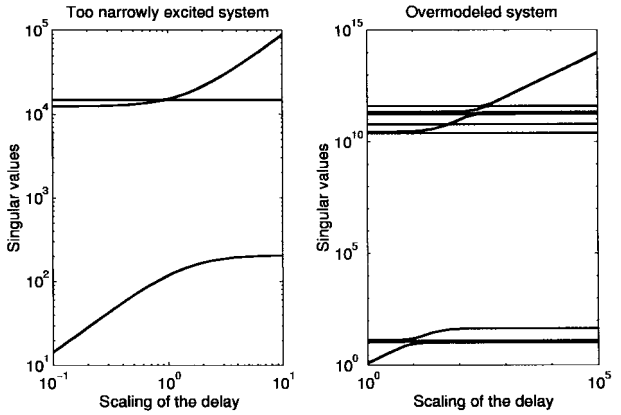

Fig. 5. Two unidentifiable lowpass filter models.

10^{-15}. The right plot corresponds to a second order highpass filter with two poles close to 1. It is excited in the band from 0 Hz to 10^{6} Hz at 11 equally spaced frequencies. Variances of the input and output noises are 10^{-7}.

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 exciting the system in a too narrow frequency band from 0 Hz to 0.01 Hz as compared to the corner frequency of 0.8 Hz. The transfer function parameters are estimated accurately while the delay is not.

The right plot corresponds to identification in a proper frequency band from 0 Hz to 1 Hz, but as an overmodeled 4/3 system. The SV's are distributed in two distinct bands. The estimation is poor. It may give a hint why scaling of the delay does not improve in this situation: if parameters cannot be identified at all, or some parameters appear as a delay due to poor experiment design, scaling makes no difference.

C. Independent Scaling 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 two estimators to determine an optimal delay scaling factor. The first is developed for *on-line* estimation during system identification sessions, as it is for the rule of thumb suggested for global scaling. The second is developed for *off-line* estimation. Somewhere, both estimators suffer from the same drawback: in order to make a prediction about an optimal scaling factor, some assumptions are made concerning the general behavior of the SV's.

A. Pairing Estimator

The name of this estimator stems from the *pairing routine* which it exploits: starting with $\mathbf{J}$, a matrix $\mathbf{J_{fix}}$ is formed by removing the last column which corresponds to the delay. The SV's of $\mathbf{J}$ and $\mathbf{J_{fix}}$ are paired and only the SV associated with the delay remains alone. Based on this value, the algorithm proposes a scaling factor to bring this SV to among the others. Pairing of SV's of $\mathbf{J}$ and $\mathbf{J_{fix}}$ is done with the so-called Hungarian method. It is based on looking for alternating paths in bipartite graphs [4, pp. 2-69–2-70].

B. Iterative Estimator

A proper scaling factor yields the minimal ratio between the largest and smallest SV, present in $\mathbf{J}$. There is a global minimum in all systems which have been studied and it decreases monotonically at both sides of this minimum, toward it. Based on this observation, robust algorithms can be used to approach this minimum iteratively, like *hunt* and *bisectioning* algorithms [7, 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 estimator.

VI. Conclusions

This paper shows that in frequency-domain models, a delay can be scaled independently of the transfer function to which it is associated, both for ordinary *s*-domain descriptions and for descriptions with orthogonal polynomials. Only one SV associated to the system equations is significantly affected by scaling of the delay.

The predominantly 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 a number of purposes:

1) to improve numerical conditioning during estimation of arbitrary models;
2) to enhance identifiability for some models with an extreme spread in accuracy among the estimated parameters;
3) as an analytical tool to provide the experimenter with some insight about the model under estimation.

References

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

[2] R. Pintelon, "Phase correction of linear time invariant systems with digital allpass filters," *IEEE Trans. Instrum. Meas.*, vol. 39, pp. 324–330, Apr. 1990.
[3] R. Pintelon and L. van Biesen, "Identification of transfer functions with time delay and its application to cable fault location," *IEEE Trans. Instrum. Meas.*, vol. 39, pp. 479–484, June 1990.
[4] I. Kollár, *Frequency-Domain System Identification Toolbox for Use with MATLAB*. Natick, MA: The Mathworks, Inc., 1994.
[5] Y. Rolain, R. Pintelon, K. Q. Xu, and H. Vold, "Best conditioned parametric identification of transfer function models in the frequency-domain," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1954–60, Nov. 1995.
[6] G. H. Golub and C. F. van Loan, *Matrix Computations*. London, U.K.: Johns Hopkins Univ. Press, 1989.
[7] W. H. Press, B. P. Flamery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes—The Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press, 1986.

Jeroen Strompf (S'98) was born in The Netherlands in 1973. He studied electrical engineering at Delft University of Technology, Delft, The Netherlands, and graduated in September 1998.

From November 1996 to April 1997, he was a Visiting Graduation Work Student at the Technical University of Budapest, Budapest, Hungary. Since his graduation, he has been involved in the start-up of several high-tech information technology firms.

István Kollár (M'87–SM'93–F'97) was born in Budapest, Hungary, in 1954. He graduated in electrical engineering from the Technical University of Budapest in 1977 and in 1985 received the degree of "Candidate of Science," (the equivalent of Ph.D.) from the Hungarian Academy of Sciences, and the degree Dr.Tech from the Technical University of Budapest.

From September 1993 to June 1995, he was a Fulbright Scholar and Visiting Associate Professor in the Department of Electrical Engineering, Stanford University, Stanford, CA. He is Associate Professor of Electrical Engineering, Technical University of Budapest. His research interests span the areas of digital and analog signal processing, measurement theory, and system identification. He has published about 50 scientific papers and is coauthor of the book *Technology of Electrical Measurements* (New York: Wiley, 1993, L. Schnell, Ed.). He authored the *Frequency-Domain System Identification Toolbox* for Matlab.

Yves Rolain, for a photograph and biography, see this issue, p. 326.