IEEE Instrumentation and Measurement
Technology Conference
Budapest, Hungary, May 21–23, 2001

Generalization of a Total Least Squares Problem in Frequency Domain System Identification

László Balogh, István Kollár
Department of Measurement and Information Systems
Budapest University of Technology and Economics
phone: + 36 1 463-[3583,1774], fax: + 36 1 463-4112, email: [lbalogh,kollar]@mit.bme.hu

***Abstract** – In this paper a solution to the frequency domain system identification of a linear time-invariant system is investigated. A generalization of the total least squares (TLS) algorithm is shown and analyzed. Some simulation examples on real measured data are given, in order to illustrate the properties of the new method in practice.*

***Keywords** – system identification, total least squares, generalized eigenvalues, TLS, frequency domain, initial value setting*

I. INTRODUCTION

Parametric system identification usually concludes in the estimation of unknown parameters in a model ([1],[2],[3]). The estimation of the parameters can be done in many different ways. For the sake of short computing time and numerical simplicity, our goal is usually to cast the problem in the form of a set of linear equations. Because of the distortions and noises in the measurement process, we consider an over-determined set of linear equation set. Therefore, we have to use an approximation which makes the linear equations compatible. One of these, the TLS method, is very effective for frequency domain system identification. However, in the TLS solution some inherent constraints have to be fulfilled, which are sensitive to linear transformations (frequency scaling, etc.). Therefore, it is important to understand what happens during transformations, and formulate how the constraints can be transformed.

The structure of this paper is the following:

- *II. Preliminaries and foundations* discusses the notations and assumptions. Furthermore it contains the basic theorems and statements.
- *III. Generalization of the TLS problem* contains the theoretical result which is a generalization of the TLS problem.
- *IV. Simulation examples* contains verification and illustration of the practical usage of the new algorithms on real measured data.

This work was supported by the Flemish and Hungarian Governments' Bilateral Cooperation Agreement (BIL99/18, TeT Nr. 99/1) and by the Belgian National Fund for Scientific Research, the Flemish Government (GOA-IMMI), and the Belgian government as a part of the Belgian programme on Interuniversity Poles of attraction (IUAP4/2) initiated by the Belgian State, Prime Minister's Office, Science Policy Programming.

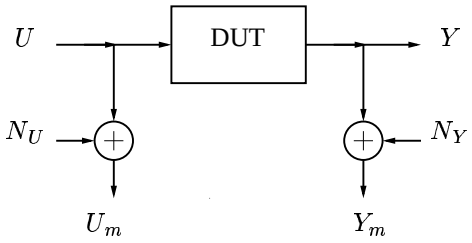

Fig. 1. The measurement setup.

II. PRELIMINARIES AND FOUNDATIONS

In the model the description of the system with its transfer function is

$$H(j\omega,\mathbf{p}) = \frac{N(j\omega,\mathbf{p})}{D(j\omega,\mathbf{p})} = \frac{\beta_{do}(j\omega)^{do} + \ldots + \beta_1(j\omega) + \beta_0}{\alpha_{no}(j\omega)^{no} + \ldots + \alpha_1(j\omega) + \alpha_0}, \quad (1)$$

$$\mathbf{p}^T = [\alpha_0, \ldots, \alpha_{no}, \beta_0, \ldots, \beta_{do}]$$

where ω is the angular frequency, α_i, β_i are the coefficients of the transfer function polynomials, $\mathbf{p}$ is the collection of α_i, β_i, and no, do are the orders of the numerator and the denominator, respectively. A similar expression can be used in z-domain if $j\omega$ is replaced by $z^{-1} = e^{-j\omega T_s}$, where T_s is the sampling time.

The model of the measurement process can be seen in Fig. 1.

We are using the following notations:

- U, Y: the exact, but unknown input and output,
- N_U, N_Y: additive noises on the input and output,
- U_m, Y_m: the measured data (Fourier amplitudes at different frequencies).

The following equations describe this stochastic model of the measurement:

$$\begin{aligned} U_m &= U + N_U \\ Y_m &= Y + N_Y. \end{aligned} \quad (2)$$

0-7803-6646-8/01/$10.00 ©2001 IEEE

The measured input and output are known at discrete frequencies denoted by $\omega_1, \dots, \omega_F$. (If we have time domain samples, the discrete Fourier spectra can be calculated by using the discrete Fourier transform or its fast version, the FFT). We assume we know the variances of the additive noises, and that the noises have zero mean, they are uncorrelated over the frequency, and they have bounded moments.

If we collect the variables into vectors, we can write:

$$\mathbf{U}_m = \mathbf{U} + \mathbf{N}_U$$
$$\mathbf{Y}_m = \mathbf{Y} + \mathbf{N}_Y,$$

where for example $\mathbf{Y}_m$ is

$$\mathbf{Y}_m = \begin{bmatrix} Y_m(j\omega_1) & Y_m(j\omega_2) & \dots & Y_m(j\omega_F) \end{bmatrix}^T.$$

Using (1), the model equation is obtained:

$$N(j\omega, \mathbf{p})U(j\omega) - D(j\omega, \mathbf{p})Y(j\omega) = 0.$$

This equation is true for every frequency, and these are linear in f, therefore we can write it in matrix form:

$$\mathbf{A}\mathbf{p} = \mathbf{0}, \tag{3}$$

where the rows of the matrix $\mathbf{A}$ belong to the corresponding ω_k. This equation is linear in $\mathbf{p}$ and the elements of $\mathbf{A}$ are

$$\mathbf{A}_{m,n} = U(j\omega_m)(j\omega_m)^{n-1} \text{ if } \quad n \leq do + 1$$

$$\mathbf{A}_{m,n} = -Y(j\omega_m)(j\omega_m)^{n-1-(do+1)} \quad \text{if} \quad n > do + 1$$

Using (2) and (3) we can introduce the noisy $\mathbf{A}_m$.

$$\mathbf{A}_m = \mathbf{A} + \mathbf{N}_\mathbf{A} \tag{4}$$

From the noise assumptions it follows that

- $\mathbf{N}_\mathbf{A}(j\omega_k)$, $k = 1, \dots F$ are zero mean, mixing ([2]), complex random variables,
- $E\left\{\mathbf{N}_\mathbf{A}\mathbf{N}_\mathbf{A}^H\right\} = \begin{bmatrix} \mathbf{C}_U & 0 \\ 0 & \mathbf{C}_Y \end{bmatrix}$,
- $E\{\mathbf{N}_\mathbf{A}\mathbf{N}_\mathbf{A}^T\} = 0$,
- the errors $\mathbf{N}_\mathbf{A}(j\omega_k)$ are independent over the frequency.

For more details, see [2] and [4].

The weighted total least squares

Using (4) we can formulate the parameter estimation as a total least squares problem ([4]), looking for a solution of $\mathbf{A}_m\mathbf{p} = \mathbf{0}$, where the solution for $\mathbf{A}_m$ may contain errors in all elements. The definition of the TLS problem is the following ([5]):

$$\min \parallel \mathbf{W}(\mathbf{A}_m - \widehat{\mathbf{A}})\mathbf{C}^{-1} \parallel_\mathrm{F}^2 \tag{5}$$

subject to

$$\widehat{\mathbf{A}}\mathbf{p} = 0 \quad \text{and} \quad \mathbf{p}^T\mathbf{p} = 1.$$

Here $\mathbf{W}$ is a left weighting matrix, and $\mathbf{C}$ is a square root of the "column covariance matrix" of $\mathbf{WA_m}$: $\mathbf{C}^T\mathbf{C} = E\{\mathbf{N}_\mathbf{A}^T\mathbf{W}^T\mathbf{W}\mathbf{N}_\mathbf{A}\}$ and F denotes the Frobenius norm. $\widehat{\mathbf{A}}$ is the estimation of $\mathbf{A}$. The properties of $\widehat{\mathbf{A}}$ and connection with LS can be found in more details in [5].

Elimination of $\widehat{\mathbf{A}}$ in (5) gives the equivalent cost function minimized by the WGTLS estimator ([4])

$$K = \min \text{ trace } \left((\mathbf{WA}_m\mathbf{p})[\mathbf{p}^T\mathbf{C}^T\mathbf{C}\mathbf{p}]^{-1}(\mathbf{WA}_m\mathbf{p})^T\right) \tag{6}$$

subject to

$$\mathbf{p}^T\mathbf{p} = 1.$$

Table I. contains the possible choices of $\mathbf{W}$ and $\mathbf{C}$.

TABLE I
POSSIBLE COMBINATIONS FOR $\mathbf{W}$ AND $\mathbf{C}$.

name	weight	covariance
TLS	$\mathbf{I}$	$\mathbf{I}$
WTLS	$\mathbf{W}$	$\mathbf{I}$
GTLS	$\mathbf{I}$	$\mathbf{C}$
WGTLS	$\mathbf{W}$	$\mathbf{C}$

Transformation of the parameter vector

In many cases, we have to transform the parameter vector into a new base. This can be described by multiplying the parameter vector with a transformation matrix and continuing the estimation algorithm with the vector obtained as the result. The applications of this can be seen in the next section. The transformation of the parameter vector can be written in the following form:

$$\widetilde{\mathbf{p}} = \mathbf{T}\mathbf{p}.$$

and that of $\mathbf{A}_m$ as $\mathbf{A}_{mt} = \mathbf{A}_m\mathbf{T}^{-1}$.

Hence we should rephrase the TLS problem. We usually do it like this:

$$\min \parallel \mathbf{W}(\mathbf{A}_{mt} - \widehat{\mathbf{A}}_t)\mathbf{C}^{-1} \parallel_\mathrm{F}^2 \tag{7}$$

subject to

$$\widehat{\mathbf{A}}_t\mathbf{T}\mathbf{p} = \widehat{\mathbf{A}}_t\widetilde{\mathbf{p}} = 0 \quad \text{and} \quad \widetilde{\mathbf{p}}^T\widetilde{\mathbf{p}} = 1,$$

and the corresponding cost function:

$$K = \min \text{ trace } \left((\mathbf{WA}_m\widetilde{\mathbf{p}})[\widetilde{\mathbf{p}}^T\mathbf{C}^T\mathbf{C}\widetilde{\mathbf{p}}]^{-1}(\mathbf{WA}_m\widetilde{\mathbf{p}})^T\right)$$

subject to

$$\widetilde{\mathbf{p}}^T\widetilde{\mathbf{p}} = 1. \tag{8}$$

Here the problem is that the known algorithms cannot account for the fact that by transforming the parameter vector, the constraint $\| \mathbf{p} \|_{\mathrm{F}}^2 = 1$ should be transformed, too. If we use constraint (8), we solve not the original WTLS problem in the new base.

Consider a two-variable parameter vector. This example is a very simple case but it helps us to imagine what happens in the higher dimension spaces. In Fig. 2a. we can see the original space of the parameter vector, the unit circle as a constraint and the assumed solution of the TLS problem. What will happen if we transform the problem into a new base? The unit circle is usually transformed into an ellipse. The points of this ellipse are the possible solutions of the original minimization problem. If we use the known algorithm, we will search the solution not on this ellipse, but on the unit circle (see Fig. 2b). It is important to note that in this case after the transformation of $\mathbf{Ap} = 0$ we do not transform the constraint ($\mathbf{p}^T\mathbf{p}$). In the *new algorithm* we suggest, we transform the minimization problem together with the constraint. Hence in the new base we solve the original problem. The new algorithm is discussed in the next section.

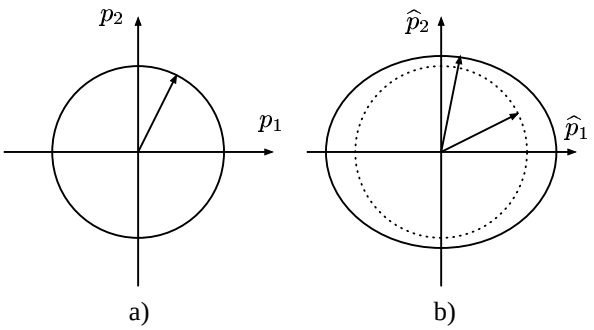

Fig. 2. The original and the transformed space of the parameter vectors.

III. GENERALIZATION OF THE TLS PROBLEM

We can generalize the WTLS problem in the following way:

$$\min \| \mathbf{W}(\mathbf{A}_m - \widehat{\mathbf{A}})\mathbf{C}^{-1} \|_{\mathrm{F}}^2$$

subject to

$$\widehat{\mathbf{A}}\mathbf{p} = 0 \quad \text{and} \quad \mathbf{p}^T\mathbf{B}^T\mathbf{B}\mathbf{p} = 1. \tag{9}$$

Now the constraint is a bilinear expression[1]. Hence the corresponding cost function is:

$$K = \min \operatorname{trace}\left((\mathbf{W}\mathbf{A}_m\mathbf{p})[\mathbf{p}^T\mathbf{C}^T\mathbf{C}\mathbf{p}]^{-1}(\mathbf{W}\mathbf{A}_m\mathbf{p})^T\right)$$

[1] Note: as a matter of fact (9) can be interpreted that the norm of $\mathbf{p}$ equals one, when we define the scalar product of vector $\mathbf{x}_1$ and $\mathbf{x}_2$ as $\mathbf{x}_1^T\mathbf{B}^T\mathbf{B}\mathbf{x}_2$.

subject to

$$\mathbf{p}^T\mathbf{B}^T\mathbf{B}\mathbf{p} = 1.$$

This problem leads to a generalized eigenvalue problem. Therefore this problem can be solved very effectively with generalized singular value decomposition (GSVD). We may use GSVD($\mathbf{WAC}^{-1}, \mathbf{B}$) or GSVD($\mathbf{WA}, \mathbf{CB}$) (see [6],[5]). The corresponding Matlab program is the following:

```
[U1,U2,X,S1,S2]=gsvd(W*A,C*B);
Xi=inv(X');
p=Xi(:,1);
```

This generalization of the constraint allows us to compensate for the transformation of the parameter vector. If we choose matrix $\mathbf{B}$ so that

$$\mathbf{B} = \mathbf{T}^{-1},$$

then we solve the problem mentioned the previous section. Picturesquely it means that we are searching the solution of the transformed WTLS problem on the transformed unit circle (the ellipse in Fig. 2b.).

The maximum likelihood cost function

In order to compare the different estimators, we need a measure of quality. The maximum likelihood (ML) cost function is a possible candidate for this. Maximum likelihood (ML) estimation is the best we can do in many cases if identification of a system is required. Unfortunately, in frequency domain system identification the maximum likelihood method leads a nonlinear problem ([3],[2]). Therefore, we cannot apply efficient numerical algorithms such as WTLS. Nevertheless the maximum likelihood has good statistically properties (see [3],[1]).

To obtain the ML cost function from (5), the matrix $\mathbf{W}$ has to be the following:

$$\begin{aligned}\mathbf{W}_{\mathrm{ML}}^{-2}(j\omega_k, \mathbf{p}) = \mathbf{N}(j\omega_k, \mathbf{p})\mathbf{C}_U(j\omega_k)\mathbf{N}^H(j\omega_k, \mathbf{p}) \\ +\mathbf{D}(j\omega_k, \mathbf{p})\mathbf{C}_Y(j\omega_k)\mathbf{D}^H(j\omega_k, \mathbf{p}).\end{aligned} \tag{10}$$

One can see that the matrix $\mathbf{W}$ depends on the parameter vector $\mathbf{p}$. This causes the minimization of this cost function be nonlinear in parameter vector $\mathbf{p}$.

$$K_{\mathrm{ML}}(\mathbf{p}) = \frac{1}{2}\sum_{k=1}^{F} \frac{|N(j\omega_k, \mathbf{p})U(j\omega_k) - D(j\omega_k, \mathbf{p})Y(j\omega_k)|^2}{\sigma_U^2(j\omega_k)|N(j\omega_k, \mathbf{p})|^2 + \sigma_Y^2(j\omega_k)|D(j\omega_k, \mathbf{p})|^2} \tag{11}$$

A possible way to compare the results of different estimators is to compare the values of the maximum likelihood cost function. Practically this means that the parameter vectors obtained as results of the different WTLS estimators have to be substituted in (11).

IV. SIMULATION EXAMPLES

In this section we will discuss the applications considering the theoretical results mentioned above. The focus is on the transformations of parameter vector.

In practice we use the transformation of the parameter vector in many cases. Here we will analyze three occurrences:

- frequency scaling,
- orthogonal polynomial base and
- known subsystem.

It is possible to combine the different cases, as mentioned later.

A. Frequency scaling

To avoid the calculation with numbers of different orders of magnitude, which is an ill-conditioned numerical way, first we scale the frequencies before the estimation algorithms will be started ([3],[1],[7],[2]). This means that the frequencies are divided by a scale factor which is generally computed in the following way:

$$\omega_{\text{scale}} = \frac{\omega_{\text{min}} + \omega_{\text{max}}}{2}$$

One can consider this as if the bandpass spectrum was moved to the center radian frequency 1.

Therefore, we have to scale the parameter vector. To obtain the final result we have to eliminate the effect of the frequency scaling. This means

$$\alpha_i (j\omega_{\text{scale}})^i \quad \text{for } i = 0, \dots, do.$$

Similarly in the case of the denominator:

$$\beta_i (j\omega_{\text{scale}})^i \quad \text{for } i = 0, \dots, no.$$

It can be seen that frequency scaling is equivalent with a transformation of the parameter vector.

$$\widetilde{\mathbf{p}} = \mathbf{T}\mathbf{p},$$

where

$$\mathbf{T}_{\text{scale}} = \text{diag}[(j\omega_{\text{scale}})^{no}, \dots, 1, (j\omega_{\text{scale}})^{do}, \dots, 1].$$

Consequently, if we would like to solve the original TLS problem we have to set $\mathbf{B}$ as

$$\mathbf{B} = \mathbf{T}_{\text{scale}}^{-1}.$$

Now let us consider a real-life experiment, mechanical measurement of a robot arm. In Fig. 3. we can see the transfer function and the variance at the measured frequencies. We will estimate a model with orders 4/6.

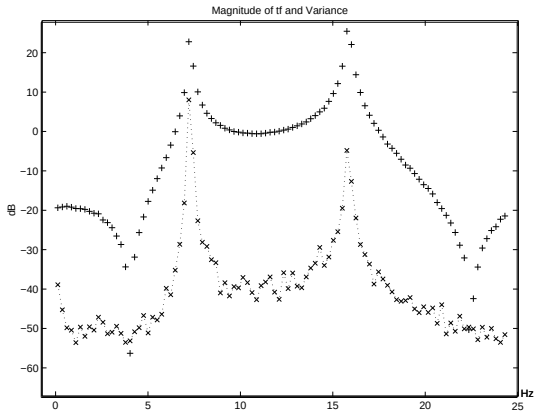

Fig. 3. The measured transfer function and the variance of the robot arm

TABLE II
THE LENGTH OF THE PARAMETER VECTORS IN THE CASE OF THE ROBOT ARM.

scaling	bilin. comp.	TLS	GTLS
no	no	1.0	1.0
yes	no	$4.26 \cdot 10^9$	$7.89 \cdot 10^9$
yes	yes	1.0	1.0

Table II. contains the estimation results. The first row of the table is the solution of the original problem. The first column is related to the TLS algorithm and the second one is related to the GTLS algorithm. It can be seen that in the case of the first and the third columns the lengths of the parameter vectors are the same. These vectors are even equal. But the parameter vector in the second row differs from the others. The cause is that in this case we did not apply the bilinear compensation for the frequency scaling.

The comparison of the results with the maximum likelihood cost function can be seen in the Table III. It is interesting to

TABLE III
THE MAXIMUM LIKELIHOOD COST FUNCTION OF THE PARAMETER VECTORS IN THE CASE OF THE ROBOT ARM.

scaling	bilin. comp.	TLS	GTLS
no	no	293	293
yes	no	274	351
yes	yes	293	293

observe that the values of the ML cost function in Table III do not seem to follow any particular order. The reason is that the original constraint $\mathbf{p}^T\mathbf{p} = 1$ is by itself arbitrary: it is not better or worse than $\widehat{\mathbf{p}}^T\widehat{\mathbf{p}} = 1$. Therefore, then is no 'best' method. This paper establishes the *equivalence* between dif-

ferent $\mathbf{p}$-domains.

B. Orthogonal polynomials

Orthogonal polynomials are used to enhance the numerical conditioning of the problem. Without details we note that using orthogonal polynomials is equivalent to a transformation ([7],[2]). If $\widetilde{\mathbf{p}}$ denotes a parameter vector in the new base computed with Gram-Schmidt orthogonalization, we can write:

$$\widetilde{\mathbf{p}} = \mathbf{T}_{\text{orth}}\mathbf{p},$$

where $\mathbf{T}_{\text{orth}}$ the transformation matrix mentioned above. In this case we have to set $\mathbf{B}$ as

$$\mathbf{B} = \mathbf{T}_{\text{orth}}^{-1}.$$

Considering frequency scaling in addition, we obtain

$$\mathbf{B} = \mathbf{T}_{\text{orth}}^{-1}\mathbf{T}_{\text{scale}}^{-1}. \tag{12}$$

We use the example demonstrated in the previous subsection (robot arm). We apply (12) as the bilinear constraint.

TABLE IV
THE LENGTHS OF THE PARAMETER VECTORS IN THE CASE OF THE ROBOT ARM (ORTHOGONAL POLYNOMIALS).

representation	bilin. comp.	TLS	GTLS
polynomial	no	1.0	1.0
orthogonal pol.	no	$5.2 \cdot 10^4$	$5.2 \cdot 10^4$
orthogonal pol.	yes	1.000	1.000

TABLE V
THE MAXIMUM LIKELIHOOD COST FUNCTION OF THE PARAMETER VECTORS IN THE CASE OF THE ROBOT ARM (ORTHOGONAL POLYNOMIALS).

representation	bilin. comp.	TLS	GTLS
polynomial	no	293	293
orthogonal pol.	no	447	447
orthogonal pol.	yes	293	293

In this case the length of the parameter vector is smaller if we do not use bilinear compensation for orthogonal polynomials (see table IV.). Moreover the value of the maximum likelihood cost function at the point $\mathbf{p}$ without orthopol compensation is larger than if we use (12).

C. Known subsystem

If we know the transfer function of a part of the system, we can incorporate it into the identification process. Up to now only one method was published to achieve this (see [3]). We will show that there is another way, too.

In this case the known subsystem is given by the numerator and denominator of its transfer function. The identification process is executed with fixed degrees of numerator and denominator. If we have a known subsystem given by its transfer function, then its degrees have to be reduced by the corresponding degrees of the numerator and the denominator of the transfer function of the subsystem, respectively.

Let N_f and D_f be the numerator and denominator of the transfer function of known part of the system, respectively. By applying these formulas, we can write both parts of the whole transfer function:

$$\begin{aligned} N(j\omega_k, \mathbf{p}) &= N_f(j\omega_k, \mathbf{p}_f)\widehat{N}(j\omega_k, \widetilde{\mathbf{p}}) \\ D(j\omega_k, \mathbf{p}) &= D_f(j\omega_k, \mathbf{p}_f)\,\widehat{D}(j\omega_k, \widetilde{\mathbf{p}}), \end{aligned}$$

where $\widehat{N}$ and $\widehat{D}$ denote the numerator and denominator of the unknown part of the transfer function, respectively. $\mathbf{p}_f$ is a vector of known coefficients.

So far if we had to make an identification with a known subsystem, we would consider this by modifying the measured data ([3]). It means that in the identification process we use the following expressions:

$$\begin{aligned} H_f(j\omega_k, \mathbf{p}) &= \frac{N_f(j\omega_k, \mathbf{p}_f)}{D_f(j\omega_k, \mathbf{p}_f)} \\ \overline{U}_m(j\omega_k) &= U_m(j\omega_k) \\ \overline{Y}_m(j\omega_k) &= \frac{Y_m(j\omega_k)}{H_f(j\omega_k, \mathbf{p})} \\ \overline{\sigma}_U^2(j\omega_k) &= \sigma_U^2(j\omega_k) \\ \overline{\sigma}_Y^2(j\omega_k) &= \frac{\sigma_Y^2(j\omega_k)}{|H_f(j\omega_k, \mathbf{p}_f)|^2}, \end{aligned} \tag{13}$$

where the overbar denotes the modified data used in the identification method. Hence we make identification with new 'measured' data and reduced-degree polynomials.

If we would like to solve the problem defined by (13) for example with the total least squares method with fixed $\mathbf{p}$-norm, then the solution will be different from that of the original problem. The contradiction can be resolved as follows.

We can construct a block Toeplitz matrix which is the transformation between $\mathbf{p}$ and $\widetilde{\mathbf{p}}$. With this formula

$$\mathbf{p} = \mathbf{T}_{\text{known}}\widetilde{\mathbf{p}} = \begin{bmatrix} \mathbf{T}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_D \end{bmatrix} \widetilde{\mathbf{p}}.$$

In details, one can write this out:

$$\mathbf{T}_N = \begin{bmatrix} \gamma_{nf} & & & \\ \gamma_{nf-1} & \gamma_{nf} & & \\ \vdots & \gamma_{nf-1} & \ddots & \\ \gamma_0 & \vdots & \ddots & \gamma_{nf} \\ & \gamma_0 & & \gamma_{nf-1} \\ & & \ddots & \vdots \\ & & & \gamma_0 \end{bmatrix}$$

$$\mathbf{T}_D = \begin{bmatrix} \delta_{df} & & & \\ \delta_{df-1} & \delta_{df} & & \\ \vdots & \delta_{df-1} & \ddots & \\ \delta_0 & \vdots & \ddots & \delta_{df} \\ & \delta_0 & & \delta_{df-1} \\ & & \ddots & \vdots \\ & & & \delta_0 \end{bmatrix}$$

where

$$N_f(j\omega_k) = \sum_{r=0}^{n_f} \gamma_r (j\omega_k)^r \quad \text{and} \quad D_f(j\omega_k) = \sum_{r=0}^{d_f} \delta_r (j\omega_k)^r.$$

Instead of the correction of the measured data (13), we use transformation of the parameter vector. Hence we can write:

$$N(j\omega_k, \mathbf{p}) = N(j\omega_k, \mathbf{T}_{\text{known}}\widetilde{\mathbf{p}}) = \widehat{N}(j\omega_k, \widetilde{\mathbf{p}})$$
$$D(j\omega_k, \mathbf{p}) = D(j\omega_k, \mathbf{T}_{\text{known}}\widetilde{\mathbf{p}}) = \widehat{D}(j\omega_k, \widetilde{\mathbf{p}}).$$

It can be seen that $\widetilde{\mathbf{p}}$ has reduced degree, because of the known subsystem. By using the bilinear expression in solving the total least squares problem, we can arrive exactly at the same solution of the original problem as with the methods in [3]. In this case

$$\mathbf{B} = \mathbf{T}_{\text{known}}.$$

It is important to see that the rank of $\mathbf{B}^T\mathbf{B}$ is smaller than the length of $\mathbf{p}$ (it equals the length of $\widetilde{\mathbf{p}}$). Hence we can solve the generalized eigenvalues problem.

Continuing the examples, let us consider that we know two poles and two zeros of transfer function of the robot arm. The amplitude diagram of known part of the system is shown in Fig. 4.

After the estimation process we obtain that with this bilinear correction the results are the same in both cases (TLS, GTLS). Table VI. contains the lengths of the parameter vector. It is illustrated that the method is correct.

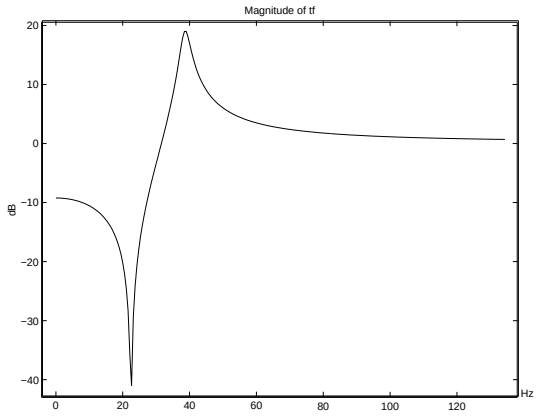

Fig. 4. The known part of the transfer function of the robot arm

TABLE VI
THE LENGTH OF THE PARAMETER VECTORS IN THE CASE OF THE ROBOT ARM WITH KNOWN SUBSYSTEM.

	TLS	GTLS
without bilin. comp.	$2.46 \cdot 10^7$	$3.73 \cdot 10^7$
with bilin. comp.	1.0	1.0

NOVELTIES

In this paper a generalization of the total least squares problem is discussed, by using a bilinear expression as a constraint of the parameter vector, instead of fixing the norm. Furthermore, three applications of this result are shown. All are important because by using the bilinear constraint, we can solve exactly the original problem in the new basis of the parameter vector. The transformation formula of the parameter vector in the case of a known subsystem is also new.

REFERENCES

[1] Pintelon et al., "Parametric identification of transfer functions in the frequency domain - a survey," *IEEE Transactions on Automatic Control*, vol. 39, no. 11, pp. 2245–60, November 1994.

[2] R. Pintelon and J. Schoukens, "System identification - a frequency domain approach," IEEE Press, to appear.

[3] István Kollár, *Frequency Domain System Identification Toolbox*, Gamax, 2001, http://elecwww.vub.ac.be/fdident/.

[4] Rik Pintelon et al., "Analyses, development, and applications of TLS algorithms in frequency domain system identification," *SIAM J. Matrix Anal. Appl.*, vol. 19, no. 4, pp. 983–1004, October 1998.

[5] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem - Computational Aspects and Analysis*, Society for Industrial and Applied Mathematics, Philadelphia, USA, 1991.

[6] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The John Hopkins University Press, Baltimore, USA, 1989.

[7] Yves Rolain R. Pintelon K. Q. Xu and H. Vold, "On the use of orthognal polynomials in high order frequency domain system identification and its application to modal parameter estimation," in *Proceedings of the 33rd IEEE Conference on Decision and Control*, Lake Buena Vista, Florida, USA, December 14-16 1994, pp. 3365–73.