Finite record effects of the errors-in-variables estimator for linear dynamic systems

Kurt Barbé, Rik Pintelon and Gerd Vandersteen
Vrije Universiteit Brussel, Dept. ELEC, Pleinlaan 2, B-1050 Brussels, Belgium
Email: Kurt.Barbe@vub.ac.be

Abstract – The frequency domain Errors-In-Variables (EIV) estimator for linear dynamic systems is generally formulated as a weighted least square estimator, where the weights used are the estimated noise (co)variances. In this paper, we study the finite record influences on the uncertainty of the EIV estimator for using the estimated weights. Furthermore, we discuss when a significant influence can occur and derive a practical method to circumvent this problem without the need of introducing a parametric noise model.

Keywords – Frequency domain, Errors-In-Variables, Weighted Least Squares, Leakage, windowing, Linear dynamic systems.

I. INTRODUCTION

Transfer function modeling is a key step in many practical engineering problems such as: the measurement and design of amplifiers, the calibration of sensors, the (physical) modeling of devices from (noisy) input-output data ...

Frequency domain system identification offers a tool to identify, from (noisy) input/output data, the transfer function $G_0$ of a linear time-invariant (LTI) dynamic system, as in Figure 1. The identification of such systems is generally formulated as a weighted least squares (WLS) problem, [1]. Besides the information of the input and output, a frequency dependent weighting matrix is used to enhance the statistical properties of the transfer function estimate. This weighting matrix depends on the noise characteristics of the input/output errors such that systematic errors are removed and the uncertainty of the transfer function estimate is reduced.

To estimate the transfer function $G_0(j\omega_k)$, measured at angular frequencies $\omega_k$, with $k = 0, \ldots, F - 1$, and $j = \sqrt{-1}$, a parametric estimate $\hat{G}(j\omega_k, \theta)$ is considered. The Errors-In-Variables (EIV) estimator of the parameters $\theta$ is found by minimizing the following quadratic cost function, with respect to the parameters $\theta$, [1],

$$
\sum_{k=0}^{F-1} \frac{|\hat{y}(k) - G(j\omega_k, \theta)\hat{u}(k)|^2}{\sigma_n^2(k) + |G(j\omega_k, \theta)|^2\sigma_e^2(k) - 2Re(\sigma_n^2(k)\hat{G}(j\omega_k, \theta))}
$$

where $\hat{u}(k), \hat{y}(k)$ are the sample mean of the discrete Fourier transform (DFT) spectra, when multiple periods are measured. $\sigma_n^2(k), \sigma_e^2(k), \sigma_{\hat{y}}^2(k)$ are the (co)variances at frequency bin $k$, and $\hat{A}$ denotes the complex conjugate of $A$.

A WLS strategy, [2], is applied to improve the statistical properties. Indeed, when there is (Gaussian) output-noise only, the weighted least squares estimator $\hat{G}(j\omega_k, \theta)$ is asymptotically, ($F \to \infty$), consistent and efficient, [3], [4]. Hence, the estimator converges to the true transfer function and has the lowest uncertainty, respectively. When both the input and output signals are disturbed by (Gaussian) noise, the WLS estimator $\hat{G}(j\omega_k, \theta)$ remains consistent, [1], but in general the efficiency property is lost. Fortunately, the loss in efficiency (i.e. the increase in uncertainty with respect to the lowest reachable bound) is small, [5].

The WLS estimator outperforms the classical least squares estimator when the numerators, in (1), have different variances (i.e. frequency dependent signal to noise ratio). By normalizing the numerators by the true (co)variances, the frequency bins with a poor signal to noise ratio are suppressed. In the case of Gaussian filtered input/output noise, one can show that the weighted least squares estimator is the Maximum Likelihood estimator, [1].

The good properties of the WLS estimator are only guaranteed asymptotically ($F \to \infty$). For finite record lengths, the denominator of (1) is not equal to the variance of the numerator due to the non-periodicity of the noise, [6]. This non-periodicity results in a leakage contribution in the numerator of the cost function (1). In practice, the leakage contribution is unknown, even in the case when the noise filters are known.

One approach to overcome this problem is by using a parametric input-output noise model, such that the leakage contribution is estimated in a parametric way, [7]. This approach is often abandoned in practice, since it is one of the hardest identification problems for linear systems, [7]:

**Figure 1** The simulation setup, where $u_0$ and $y_0$ are the respective input/output signals, $u$ and $y$ the measured input/output and $n_u$ and $n_y$ the (filtered) input/output noise.

978-1-4244-3353-7/09/$25.00 ©2009 IEEE 1525
one needs to treat the excitation as an arbitrary signal instead of a periodic signal. Besides the identification of the transfer function of system and the input/output noise, one needs to model the excitation as filtered noise, to keep the problem identifiable. Four model selections have to be performed one for the system model, two for the noise models and one for the signal model. The quality of the estimate of the system transfer function strongly depends on the quality of the estimated noise models. Generally initial values for the more complex optimization algorithm are in general not an easy task.

However, in many situations, the leakage contribution of the noise is assumed to be negligible. In this paper, we study the influence of this assumption on the finite record properties of the WLS estimator. Furthermore, we discuss when a significant influence can occur and derive a practical method to circumvent this problem without the need of introducing a parametric noise model.

II. PROBLEM STATEMENT

A. The measurement Set-up

In Figure 1, the measurement set-up is defined. We consider a normalized periodic excitation signal \( u_0 \), such that its standard deviation does not depend on the number of frequency lines \( F \). The steady state response to a periodic input signal is a periodic output signal \( y_0 \).

To avoid transient effects, the measurements only start when the transients become small. Leakage errors are avoided by measuring an integer number of periods \( P \) of \( u_0 \) where every measured period has \( L \) data points.

We assume that both the input signal \( u_0 \) and the output signal \( y_0 \) are disturbed by zero-mean additive filtered (Gaussian) noise \( n_u, n_y \) respectively. Both noise sources are allowed to be mutually correlated. The noise filters are assumed to be Bounded input-Bounded output (BIBO) stable. The measured input and output signals are respectively denoted by \( u, y \).

B. Formalizing the problem statement

We define the discrete Fourier transform at frequency bin \( k \) of the \( i \)th period, \( u[i] \), of the signal \( u \) as,

\[
U[i](k) = \frac{1}{\sqrt{F}} \sum_{n=0}^{L-1} u[i](n)e^{-j\frac{2\pi kn}{L}} \tag{2}
\]

A similar expression holds for the output Fourier coefficients \( Y[i](k) \). The set-up formulated in subsection A, implies that for the input and output Fourier coefficients the following holds, [1],

\[
U[i](k) = U_0(k) + H_0(k)E_u[i](k) + T_u[i](k) \tag{3}
\]

where \( U_0(k), H_0(k), E_u[i](k) \) and \( T_u[i](k) \) are the discrete Fourier coefficients of the true input signal, the input noise filter, the underlying white noise sequence of the \( i \)th sub-record and a leakage effect due to the non-periodicity of the noise.

It can be shown, [6], that the leakage effect \( T_u[i](k) \) vanishes from equation (3) proportional to \( L^{-\frac{1}{2}} \) with probability one. Furthermore, it is shown in [8], that

\[
\operatorname{Var}(\hat{P}(k) - G(j\omega_k \theta)\hat{U}(k)) = \sigma^2_{\epsilon}(k) + |G(j\omega_k \theta)|^2\sigma^2_{\epsilon}(k) - 2\Re(\sigma^2_{\epsilon}(k)G(j\omega_k \theta)) + O(L^{-1}) \tag{4}
\]

where \( O(L^{-1}) \) is due to the leakage terms in (3). Hence, it can be suspected that removing the leakage errors influences the statistical properties of the transfer function estimator for finite record lengths.

Taking the structure of the leakage errors \( T_u[i](k) \) where \( x = u, y \) into account, shows that the leakage errors become dominant in (4), for those frequencies where the noise filter is nearly zero. This follows from the fact that the leakage errors \( T_u[i](k) \) have the same poles as the noise filter but possibly different zeros. Therefore, for the frequency bin \( k_0 \) where \( H_u(k_0) \approx 0 \), we obtain that,

\[
\operatorname{Var}(U[i](k_0)) \approx \operatorname{Var}(T_u[i](k_0))
\]

However, in the cost function the leakage contribution is not taken into account.

III. SUPPRESSION OF THE LEAKAGE ERROR

The leakage of the noise is unknown in practice. Since we consider the situation where the leakage is not modeled, we can only suppress its contribution in the cost function (1). The general approach to suppress leakage errors is by introducing a smooth window, [9]. In this section, we shall restrict the analysis to the input signal only. Similar expressions hold for the output signal.

We define the windowed Fourier transform at frequency bin \( k \) of the \( i \)th period, \( u[i] \), as

\[
U_w[i](k) = \frac{1}{M} \sum_{n=0}^{L-1} u[i](n)w(n)e^{-j\frac{2\pi kn}{L}} \tag{5}
\]

where \( M = \sqrt{\sum_{n=0}^{L-1} w(n)^2} \).

A. Analysis of the leakage error

In the situation where the noise (co)variances are known, we want to suppress the leakage effect in the numerator of the cost function as good as possible. The application of a window function \( w(n) \) and computing the sample mean of the Fourier coefficients changes equation (3) to,

\[
\hat{U}_w(k) = W * U_0(k) + W * (H_0(k)\hat{E}_w(k)) + W * \hat{T}_w(k) \tag{6}
\]

In (6) the leakage contribution is further suppressed in the term \( W * \hat{T}_w(k) \), [10]. Unfortunately, an unwanted effect also appears in (6), namely the spectrum of the true signal \( U_0(k) \) is changed to \( W * U_0(k) \). For finite record lengths this introduces systematic errors in the transfer function estimate.

A good choice of window to suppress the leakage error is the Hanning window, [9]. In [10] it is shown that the
Hanning window suppresses the leakage error such that $W \ast \hat{U}_o(k) = O \left( \frac{L}{2} \right)$ with probability one. If the window is a Hanning window, the first term of equation (6) simplifies to, [8],

$$\bar{U}_w(k) = \frac{1}{2} U_o(k) - \frac{1}{4} \left[ (U_o(k-1) + U_o(k+1) \right]$$

(7)

where $H_{u,w}(k), \hat{E}_{u,w}(k)$ denotes the windowed DFT defined in (5). Equation (7) shows that the true spectrum $U_o(k)$ is changed to a weighted sum of the three consequent DFT lines, as we show in Subsection B. The bias introduced can easily be eliminated by considering two periods instead of one.

B. Data processing in blocks of two periods

Let us consider blocks of two periods, then equation (5) becomes,

$$U^{[i]}_{w_2}(k) = \frac{1}{M} \sum_{m=0}^{2L-1} u^{[i]}_{w_2}(n) w(n) e^{-j2\pi \frac{n}{L}}$$

(8)

where $u^{[i]}_{w_2}(n)$ denotes the ith block of 2 periods of the signal $u$ and $U^{[i]}_{w_2}(k)$ denotes the windowed DFT over the ith block of 2 consequent periods.

In Section II.B, it was explained that the true excitation signal $u_o$ is a periodic signal with $F$ frequencies. Clearly, considering blocks of two periods implies, that the spectrum of the true input signal $U_o$ is only present at the even frequency bins in equation (8).

Applying the Hanning window; we obtain, [8],

$$\bar{U}_{w_2}(2k) = \frac{1}{2} U_o(k) + H_{u,w_2}(2k) \hat{E}_{u,w_2}(2k)$$

$$+ O \left( \frac{L}{2} \right)$$

(9)

where $H_{u,w_2}(2k), \hat{E}_{u,w_2}(2k)$ denotes the windowed DFT defined in (8). In equation (9) the leakage error was suppressed without deforming the true spectrum $U_o$.

In the next section, we discuss the use of the Hanning window over blocks of two periods in the WLS cost function (1).

IV. WINDOWED WEIGHTED LEAST SQUARES COST FUNCTION

A. Known noise (co)variances

We first consider the case when the true noise (co)variances are known. In that situation the denominator of the cost function (1) is known. As explained in section II.B, the variance of the cost function is not equal to the denominator of the cost function, due to leakage errors.

To reduce the influence of the leakage errors, we use the modified DFT, (8), computed in blocks of two periods of the signal as explained in the previous section. The cost function becomes,

$$\sum_{k=0}^{F-1} \frac{|\hat{Y}_{w_2}(2k) - G(j\omega_k, \theta)\bar{U}_{w_2}(2k)|^2}{\sigma^2(k) + |G(j\omega_k, \theta)|^2 \sigma^2(k) - 2Re(\sigma^2(k)G(j\omega_k, \theta))}$$

(10)

The simulation in section V reveals that an improvement, with respect to the classical computation of the Fourier coefficients (5), up to 10 dB can be expected for the Root Mean Squared Error (RMSE) of the transfer function estimate if zeros are present in the noise transfer function. Unfortunately, if no transmission zeros are present in the noise filter the uncertainty of the transfer function estimate increases, with respect to the classical computation of the Fourier coefficients (5), with approximately 1.76 dB.

Indeed, it is easy to see that by using blocks of two periods instead of one, the uncertainty of $\bar{Y}_{w_2}(2k)$, the sample mean of the windowed Fourier coefficients, (8), is larger than of $\hat{Y}(k)$ (three noisy DFT lines are combined by the Hanning window). Furthermore, it can be shown, [8], if we denote $K_{WLS}$ as the cost function (1) and $K_{WWSLS}$ the cost function (10), where WWSLS indicates the use of the window weighted least squares,

$$\frac{\text{Var}(K_{WWSLS})}{\text{Var}(K_{WLS})} = \frac{3}{2}$$

which implies an approximate increase of the parameter uncertainty of 1.76 dB. When there are zeros present in the transfer function of the noise filters, this increase in uncertainty of 1.76 dB is compensated by the reduction of the leakage error. This is illustrated in section V.

B. Unknown noise (co)variances

In the case of unknown noise (co)variances, we have the cost function,

$$\sum_{k=0}^{F-1} \frac{|\hat{Y}_{w_2}(2k) - G(j\omega_k, \theta)\bar{U}_{w_2}(2k)|^2}{\sigma^2(k) + |G(j\omega_k, \theta)|^2 \sigma^2(k) - 2Re(\sigma^2(k)G(j\omega_k, \theta))}$$

(11)

In the cost function, (11), the numerator is equal to the numerator of (10). We do not need to modify the numerator since it is not a function of the true (co)variances. For the denominator, we cannot use the cost function (10), since the true (co)variances are unknown.

Classically, in the case of unknown noise (co)variances these are estimated from the data, [1], by using the sample (co)variance of the Fourier coefficients,

$$\hat{\sigma}^2(k) = \frac{1}{p-1} \sum_{i=1}^{p} |U^{[i]}(k) - \bar{U}(k)|^2$$

$$\hat{\sigma}^2(k) = \frac{1}{p-1} \sum_{i=1}^{p} |Y^{[i]}(k) - \bar{Y}(k)|^2$$

$$\hat{\sigma}^2(k) = \frac{1}{p-1} \sum_{i=1}^{p} \left( U^{[i]}(k) - \bar{U}(k) \right) \times \left( Y^{[i]}(k) - \bar{Y}(k) \right)$$

(12)

By using (12), the leakage errors (3) are also present in the estimates of the noise (co)variances. To overcome this problem, we can replace the Fourier coefficients $U^{[i]}(k)$,
\[ Y^{[i]}(k), \hat{U}(k), \hat{Y}(k) \] by \[ U^{[i]}_w(k), Y^{[i]}_w(k), \hat{U}_w(k), \hat{Y}_w(k) \]
respectively.

\[
\delta_\theta^2(k) = \frac{1}{p-1} \sum_{i=1}^{p} \left| \frac{U^{[i]}_w(k) - \bar{U}_w(k)}{U^{[i]}_w(k)} \right|^2
\]
\[
\delta_\gamma^2(k) = \frac{1}{p-1} \sum_{i=1}^{p} \left| \frac{Y^{[i]}_w(k) - \bar{Y}_w(k)}{Y^{[i]}_w(k)} \right|^2
\]
\[
\delta_\gamma_0^2(k) = \frac{1}{p-1} \sum_{i=1}^{p} \left( \frac{U^{[i]}_w(k) - \bar{U}_w(k)}{U^{[i]}_w(k)} \right) \times \left( \frac{Y^{[i]}_w(k) - \bar{Y}_w(k)}{Y^{[i]}_w(k)} \right)
\]

Please note that we did not compute the sample variances over blocks of 2 periods, since it is not needed as the difference \( U^{[i]}_w(k) - \bar{U}_w(k) \) is independent of the true spectrum \( U_0 \).

Let us denote \( K_{SWWLS} \) the cost function (11), where the estimated (co)variances (13) were used. The abbreviation SWWLS indicates the use of the Windowed Weighted Least Squares with Sample (co)variances. Let us denote \( K_{SWLS} \) the cost function (1), where the estimated (co)variances (12) were used. The abbreviation SWLS indicates the use of the Weighted Least Squares with Sample (co)variances.

The simulation in section V reveals that if the noise transfer function has zeros, the use of the cost function \( K_{SWWLS} \) decreases the uncertainty on the transfer function estimate up to 10 dB with respect to the uncertainty on the transfer function estimate by using the cost function \( K_{SWLS} \).

Similar to the case where the noise (co)variances are known, one can show, [8], that the uncertainty of the noise transfer function estimate increases by 1.76 dB, when the cost function \( K_{SWWLS} \) is used with respect to the case where the cost function \( K_{SWLS} \) is used when no zeros are present in the noise transfer function.

A remaining theoretical question is the consistency of the SWLS-estimator. It needs to be shown that the numerator and denominator in (11) are still stochastically independent.

When there are zeros present in the transfer function of the noise filters, this increase in uncertainty of 1.76 dB is compensated by the reduction of the leakage error. This is illustrated in section V.

V. NUMERICAL EXAMPLE

In this section, we shall compare the classical (S)WLS estimator (1) with the (S)WWLS estimator (10). In the two simulations the same transfer function \( G_0(j \omega) \) but the applied noise filters, \( H_x \) for \( x = u, y \), are different.

In the first example a Butterworth filter was used for the input noise and the output noise, in the second example both the input and output noise filter has two transmission zeros symmetrical positioned with respect to the resonance of the system transfer function \( G_0(j \omega) \).

A. Example

In this example the simulation set-up as in Figure 1 was used. The true system \( G_0 \) is a second order digital filter. The same second order Butterworth filter with cut-off frequency at 0.1 \( \times f_c \) was used for the input and output noise filters \( H_u, H_y \). In Figure 2, the configuration of the poles and zeros of the system transfer function in black and of the noise transfer function (note that \( H_u = H_y \)) in gray.

![Figure 2](image_url)

The true input signal \( u_0 \) is a random phase multisine,

\[ u_0(t) = \frac{1}{\sqrt{F}} \sum_{k=1}^{F} \cos\left( 2\pi \frac{k}{L} t + \phi_k \right) \]

where \( t = 0, \ldots, PL - 1 \), and where the phases \( \phi_k \) are drawn ad random from a uniform [0, 2\( \pi \] ] distribution. For the simulation we choose \( P = 6, L = 1000 \) and \( F = 333 \). The standard deviation of the driving noise sources is chosen in such a way that the input and output signal to noise ratio \( \frac{\sigma_u}{\sigma_{noise}} \) equals 35 dB.

In Figure 3 the Root Mean Squared Error (RMSE) is shown of the transfer function estimate following the set-up above. The Monte-Carlo simulation consists of 1000 experiments. In the top plot the classical WLS cost function (1) was used where in the bottom plot the WWLS cost function (10) where a Hanning window was used. The solid black curve indicates the use of the true (co)variances, the solid gray curve indicates the use of the sample (co)variances (S(W)WLS-estimator) and the dashed black curve is the predicted loss in efficiency based on linear system identification theory, [1]. Comparing the top plot, where the classical method was used, and the bottom plot where the WWLS was used; we observe the increase in RMSE of approximately 1.7 dB as predicted in section IV.

B. Example
In this example, we use the same system \( G_0(j\omega) \) and excitation signal as in Example B. The input and output noise filters in this example has two transmission zeros symmetrical positioned with respect to the resonance frequency of the system transfer function \( G_0(j\omega) \). The poles/zeros configuration of the noise filter with respect to the system is given in Figure 4.

The Monte-Carlo simulation consists of 1000 experiments. In the top plot the classical WLS cost function (1) was used where in the bottom plot the WWLS cost function (10) where a Hanning window was used. The solid black curve indicates the use of the true (co)variances, the solid gray curve indicates the use of the sample (co)variances (S(W)LSE-estimator) and the dashed black curve is the predicted loss in efficiency.

where a Hanning window was used. The solid black curve indicates the use of the true (co)variances, the solid gray curve indicates the use of the sample (co)variances and the dashed black curve is the predicted loss in efficiency.

Comparing the top plot, where the classical method was used, and the bottom plot where the WWLS was used; we observe a decrease in RMSE up to 10 dB when the WWLS estimator was used with respect to the WLS estimator.

VI. CONCLUSION

In this paper, we studied a finite record effect of the frequency domain EIV-estimator for LTI systems. We showed that using asymptotic weights in the WLS cost function can have a significant influence on the uncertainty of the transfer function estimate. In particular when transmission zeros are present in the noise filters, an significant increase on the uncertainty of the transfer function estimate is to be expected.

We derived an easy method to eliminate this increase in uncertainty on the transfer function estimate. The method is small extra data processing step which is easily implemented in practice.

Using a Hanning window in the WLS cost function, we showed that a decrease up to 10 dB in the uncertainty of the transfer function estimate can be observed, with respect to the classical WLS (without Hanning window), when
transmission zeros are present in the noise filter. However, if no transmission zeros are present in the noise filter, the WLS with Hanning window increases the uncertainty of the transfer function estimate with approximately 1.76 dB, with respect to the classical WLS estimate.

ACKNOWLEDGEMENT

This work was funded by the Fund for Scientific Research (FWO-Vlaanderen), the Methusalem grant of the Flemish Government (METH-1), and the Belgian program on interuniversity poles of attraction initiated by the Belgian State, Prime Minister’s Office, Science Policy programming (IUAP-VI/4) and the research council of the VUB (OZR).

REFERENCES