A Comparison of Transfer Function Estimators

Piet M. T. Broersen

***Abstract*—The response of a linear time-invariant process on a stochastic input signal is characterized by the transfer function. Unknown past inputs and future output are sources of inaccuracy in relating a finite segment of an output signal via an estimated transfer function to the corresponding input segment. These end effects are usually characterized with error bounds on the Fourier transform of the output signal, but the error in an estimated transfer function can be quantified more precisely in terms of bias and variance. The accuracy of three transfer function estimators is compared, showing an infinite variance for the Experimental Transfer Function Estimate (ETFE) and a better efficiency for the estimators which are based on the cross spectrum. The variance due to additive noise depends on whether the input is a stochastic or a deterministic signal.**

I. Introduction

TRANSFER functions can be estimated with parametric models or nonparametrically with Fourier analysis. The precision of transfer functions is important for hard error bounds in robust control applications. In a class of parametric models, the mean-square error between the true and the estimated transfer functions can be modeled as a sum of two terms which both depend on the order of the estimated model [1]. The bias term decreases with the model order, and a variance term increases with this order. By considering the bias to be exclusively caused by estimating low-order models from data that are generated by higher order processes, this behavior can be used in a criterion for model order selection in parametric estimation of the transfer function [1]. Earlier results [2] showed the possibilities of manipulating the bias in transfer function estimation.

The Empirical Transfer Function Estimate (ETFE) [3], [4] is a natural nonparametric estimate for transfer function estimation, based on deterministic viewpoints. With stochastic inputs, this transfer function estimate is accompanied by extra terms. One term is the additive noise; the other describes transients due to the input signal that was present before the interval, together with the output transients afterwards. Such end effects are usually characterized by bounds on the Fourier transform, as an infinite summation of absolute values [3]–[5]. Tighter bounds can be derived if additional information about the estimated process is taken into account [6]. Expressions for the additive noise variance in transfer functions with periodical input have been derived [4], [7]. Also an explicit expression for the bias due to the end effects has been given [8].

Raw ETFE estimates are generally not accurate enough, so smoothing is required. Three possibilities are: averaging of ETFE's, splitting of the data in subframes and using the average periodograms, and smoothing of long periodograms with a spectral window [3], [4]. This paper derives expressions for the variance of average quotients of real and complex stochastic variables, with application to transfer functions. A numerical simulation example shows that the proposed bias and variance expressions explain empirical results. Different ways of averaging the data in the Fourier domain give a different accuracy in the transfer functions estimated with stochastic inputs. When the input is a periodical or deterministic signal, no bias is found, and the three ways of averaging give the same, smaller, variance.

Manuscript received May 10, 1994; revised February 20, 1995.

The author is with the Department of Applied Physics, Delft University of Technology, 2600 GA Delft, The Netherlands.

IEEE Log Number 9411485.

II. Variance of Stochastic Quotients

The quotient of two real normally distributed zero-mean stochastic variables has a Cauchy distribution [9], which has no finite moments such as mean or variance. Introducing normally distributed, real, zero-mean x_i with variance σ_x^2, independent from arbitrarily distributed c_i with variance σ_c^2, it follows from (A1) in the Appendix by scaling the variance of the x_i to σ_x^2 that

$$\operatorname{var}\frac{c_1x_1+\cdots+c_Kx_K}{x_1^2+\cdots+x_K^2}=\frac{1}{(K-2)}\frac{\sigma_c^2}{\sigma_x^2}. \tag{1}$$

The expectation of a quotient of two zero-mean complex variables exists and is finite. For complex independent zero-mean normally distributed variables, it is given by [10, p. 98]

$$\mathcal{E}\left(\frac{a}{b}\right)=\mathcal{E}\left(\frac{ab^*}{bb^*}\right)=\frac{\mathcal{E}(ab^*)}{\mathcal{E}(bb^*)} \tag{2}$$

where the * denotes complex conjugate. The theoretical variance of a/b is ∞, which is a consequence of only two independent contributions of one complex number in the denominator; see also (A1) in the Appendix. Now, define b_i with independent normally distributed real and imaginary parts, each i.i.d. $N(0,\sigma_b^2/2)$. By considering the two independent real and imaginary contributions to the numerator whereas the denominator consists of $2K$ independent squares with variance $\sigma_b^2/2$ each, it follows elementarily from (A1) that, for $K\geq 2$

$$\operatorname{var}\frac{b_1}{\sum_{i=1}^{K} b_ib_i^*}=\frac{2}{2K(2K-2)}\frac{1}{\sigma_b^2/2}=\frac{1}{K(K-1)}\frac{1}{\sigma_b^2}. \tag{3}$$

Multiplication of (3) by a_i, independent of b_i, with $\mathrm{Re}\,(a_i)$ and $\mathrm{Im}\,(a_i)$ i.i.d. $N(0,\sigma_a^2/2)$ and taking the average of K terms

0018–9456/95$04.00

in the numerator gives

$$\operatorname{var} \frac{\sum_{i=1}^{K} a_i b_i^*}{\sum_{i=1}^{K} b_i b_i^*} = \frac{2K}{2K(2K-2)} \frac{\sigma_a^2}{\sigma_b^2/2} = \frac{1}{K-1} \frac{\sigma_a^2}{\sigma_b^2} \tag{4}$$

which is ∞ for $K = 1$, like the variance of a/b before.

Each complex variable can be seen as a realization of a Fourier transform for a single frequency. The distribution of a transform of N observations will tend to normality by the central limit theorem, and real and imaginary parts will be independent [5]. At least three real variables in (1) or two complex variables in (4) are required in the denominator of the variance expressions to obtain a finite result. Simulations with normal variables corroborate the theoretical variance expressions; uniform distributions yielded only slightly different outcomes.

III. Statistical Analysis of Transfer Functions

The processes in this comparison of transfer function estimators can be described as

$$y(t) = \sum_{k=0}^{\infty} g(k) u(t-k) + v(t) \tag{5}$$

with $y(t)$ the observed output signal, $u(t)$ a zero-mean normally distributed white noise with variance σ_u^2, and $g(k)$ the impulse response with the stable transfer function

$$G(e^{j\omega}) = \sum_{k=0}^{\infty} g(k) e^{-j\omega k}. \tag{6}$$

The additive output noise $v(t)$ is a stationary series of random variables independent of $u(t)$, with zero-mean value, finite variance, and power spectral density $h_v(\omega)$.

The Fourier transform of M observations of the output signal is defined as

$$Y_M(\omega) = \frac{1}{\sqrt{2\pi M}} \sum_{t=0}^{M-1} y(t) e^{-j\omega t}. \tag{7}$$

The values obtained for $\omega = 2\pi k/M$, $k = 0, 1, \cdots, M-1$ form the familiar discrete Fourier transform (DFT). Likewise $U_M(\omega)$ with $u(t)$ and $V_M(\omega)$ with $v(t)$ are DFT transform pairs. The relation between the transforms is given by [3], [5]

$$Y_M(\omega) = G(e^{j\omega}) U_M(\omega) + V_M(\omega) + R_M(\omega). \tag{8}$$

The term $R_M(\omega)$ represents the end effects, which are the response on the input signal prior to $t = 0$ and the continuation of the response on $u(t)$ for $t \geq M$

$$R_M(\omega) = \frac{1}{\sqrt{2\pi M}} \sum_{k=0}^{\infty} \left[g(k) e^{-j\omega k} \cdot \sum_{t=-k}^{-1} \{u(t) - u(t+M)\} e^{-j\omega t} \right]. \tag{9}$$

Several approximations for limits of this term have been derived as infinite sums of absolute values [3], [6]. It is easily seen that $R_M(\omega)$ equals zero for periodical inputs with period M.

Suppose that KN observations of input and output signals are available. Transfer functions are usually estimated by smoothing raw estimates. This can be realized in three ways. The average of the ETFE is the first estimator. It is defined as [3], [4]

$$G_1(e^{j\omega}) = \frac{1}{K} \sum_{i=1}^{K} \frac{Y_N^{(i)}(\omega)}{U_N^{(i)}(\omega)} \tag{10}$$

where $Y_N^{(i)}(\omega)$ is the Fourier transform of subframe i of length N. The second estimator is the quotient of the average cross spectrum between input and output divided by the input spectrum

$$G_2(e^{j\omega}) = \frac{\frac{1}{K} \sum_{i=1}^{K} Y_N^{(i)}(\omega) U_N^{(i)}(\omega)^*}{\frac{1}{K} \sum_{i=1}^{K} U_N^{(i)}(\omega) U_N^{(i)}(\omega)^*}. \tag{11}$$

The third alternative starts with a long Fourier transform of all KN observations. Afterwards, a spectral window is used to estimate the cross spectrum in the numerator and the input spectrum in the denominator. Many different windows have been described [3], [5]. For ease of notation but without loss of generality, the Daniell window is used in this paper. With that window the third estimate becomes, for odd window length K

$$G_3(e^{j\omega}) = \frac{\frac{1}{K} \sum_{i=-(K-1)/2}^{(K-1)/2} Y_{KN}(\omega + i\Delta) U_{KN}(\omega + i\Delta)^*}{\frac{1}{K} \sum_{i=-(K-1)/2}^{(K-1)/2} U_{KN}(\omega + i\Delta) U_{KN}(\omega + i\Delta)^*} \tag{12}$$

where Δ equals $2\pi/KN$. One effect of a window is a bias in $G_3(e^{j\omega})$, which is proportional to the second derivative of the transfer function with respect to its argument, if the input spectrum is flat [3], [5]. This type of bias is eliminated in all results of this paper. The three estimates for the transfer function coincide for periodical inputs that fit on the interval N. Differences exist for stationary stochastic input signals.

A. Bias and Variance of Transfer Function Estimators

For a single frequency, each Fourier transform is just a complex random variable. An approximation to the bias due to the end effects can be derived easily with (2) as [8]

$$\operatorname{bias}[G_{1,2,3}(e^{j\omega})] \approx \frac{1}{M} e^{j\omega} G'(e^{j\omega}) \tag{13}$$

where the prime denotes differentiation with respect to the argument $e^{j\omega}$. This will be a good approximation of the bias if the transform length M is such that the impulse response decays effectively to zero in this interval. That is anyhow a prerequisite for accurate nonparametric estimation of a transfer function.

The variance of an estimated transfer function has a component due to the additive noise [3], [5], [7] and a second one that depends on the end effects of the finite sample. With (8) it follows that

$$\frac{Y_M(\omega)}{U_M(\omega)} = G(e^{j\omega}) + \frac{R_M(\omega) + V_M(\omega)}{U_M\omega)}. \tag{14}$$

$G(e^{j\omega})$, without subscript, is the theoretical transfer function of (6). This is a complex constant for every frequency, that may be disregarded in variance computations where $R_M(\omega) + V_M(\omega)$ corresponds to a_i and $U_M(\omega)$ to b_i in (4).

In $G_1(e^{j\omega})$, the division takes place before the averaging. Hence, the variance of $G_1(e^{j\omega})$ is equal to $1/K$ times the variance of a/b, which has an infinite value because at least three degrees of freedom are required for a finite result. This means that the estimator $G_1(e^{j\omega})$ has no finite variance if the input is a stochastic process. Both $\mathcal{E}[Y_M(\omega)]$and $\mathcal{E}[U_M(\omega)]$ are zero for the signals considered. This causes the estimate obtained with only one subframe to be highly irregular. The sample variance does not converge to any fixed value even after a great number of simulation runs, because it is very strongly influenced by that single realization where $U_M(\omega)$ is closest to its expectation zero. This effect has the strongest influence for the frequencies $\omega = 0$ and π, because $U_M(\omega)$ and $Y_M(\omega)$ are real there; the absolute value for other frequencies has two independent contributions of the real and the imaginary parts and will less frequently become very close to zero. So the quotient of Fourier transforms ETFE is no sound basis for transfer function estimation with stochastic excitation.

The denominator term of the variance of $G_2(e^{j\omega})$ and $G_3(e^{j\omega})$ requires $\mathcal{E}[U_M(\omega)U_M(\omega)^*]$, the power spectral density$h_u(\omega)$ that equals $\sigma_u^2/2\pi$. Further, the additive noise is independent of the process, so $R_M(\omega)$ and $V_M(\omega)$ are also independent and

$$\begin{aligned}\mathcal{E}(\{R_M(\omega) + V_M(\omega)\}\{R_M(\omega) + V_M(\omega)\}^*)\\ = \mathcal{E}(\{R_M(\omega)R_M(\omega)^*\}) + h_v(\omega).\end{aligned} \tag{15}$$

With (9), the end-effect term can be written as (see (16) at the bottom of this page). Here $u(t)$ is assumed to be uncorrelated white noise, so the contributions of $u(t)$ and $u(t+M)$ in (9) to the product in (15) are equal, and cross-product contributions may be neglected. The total variance is approximated as twice the contribution of $u(t)$ and follows with (4) as

$$\mathrm{var}\,[G_{2,3}(e^{j\omega})] = \frac{1}{K-1}\left[\frac{h_v(\omega)}{h_u(\omega)} + \frac{\mathcal{E}\{R_M(\omega)R_M(\omega)^*\}}{h_u(\omega)}\right], \quad \omega \neq o, \pi \tag{17}$$

for the frequencies with complex transfer function. The divisor in front should be $1/(K-2)$ for the frequencies 0 and π, because the transfer function is real there. The end-effect variance includes the divisor M, the length of the Fourier transforms. For $G_3(e^{j\omega})$ this variance is K times smaller than the variance contribution to $G_2(e^{j\omega})$, because the end effect occurs only once in $G_3(e^{j\omega})$ and K times in $G_2(e^{j\omega})$, once for every subframe.

IV. Simulation Results

The sample variance of $G_1(e^{j\omega})$ and the accuracy of the variance formula (17) for $G_2(e^{j\omega})$ and $G_3(e^{j\omega})$ have been studied in simulations with many different processes, values of K ranging from 2 to 1000 and various signal-to-noise ratios. For stochastic input signals with normal distribution, the results are described accurately by using as divisor in (17) $K-1$ instead of the value K that is usually given in the theory for the reduction of the additive noise in subframes [3]. Especially for low values of K, it can be seen clearly from the simulations that the variance has $K-1$ as divisor.

This section presents an example showing to what extent the proposed bias and variance expressions explain simulation results. An Infinite Impulse Response process with one left-hand and one right-hand side term has been used, given by

$$\begin{aligned} x(t) + ax(t-1) &= u(t) + bu(t-1)\\ y(t) &= x(t) + v(t).\end{aligned} \tag{18}$$

The formulas for the bias and variance of this process can be derived analytically, and they will be compared with the average of 200 000 simulation runs for $N = 64, K = 9, a = 0.5, b = -0.8, h_u(\omega) = 1/2\pi$, $h_v(\omega) = 0.2/2\pi$. The theoretical results obtained with (13) and (17) are

$$\begin{aligned}\mathrm{bias}\,[G_{1,2,3}] &= \frac{a-b}{M}\frac{e^{-j\omega}}{(1+ae^{-j\omega})^2}\\ \mathrm{var}\,[G_{2,3}] &= \frac{1}{K-1}\left[\frac{h_v(\omega)}{h_u(\omega)} + \frac{2}{M}\frac{(b-a)^2}{1-a^2}\right.\\ &\qquad \left.\cdot\left(\frac{1}{|\,1+ae^{-j\omega}\,|^2}\right)\right].\end{aligned} \tag{19}$$

The bias in the imaginary part is multiplied by M in Fig. 1. The result of (19), multiplied by M, is presented by the drawn line for a comparison, showing that the bias is accurately described for all three estimators. The Fourier interval M in (19) is K times longer for $G_3(e^{j\omega})$: $M = KN$, so the bias in $G_3(e^{j\omega})$ is K times smaller than the bias in the other two estimators. As $G_1(e^{j\omega})$ and $G_2(e^{j\omega})$ use the same length in the Fourier transform, the same bias formula applies to both. It is according to theory [10] that the sample mean of $G_1(e^{j\omega})$ converged, being the quotient of two zero-mean

$$\begin{aligned}\mathcal{E}(R_M(\omega)R_M(\omega)^*) &\approx \frac{2}{2\pi M} \times \mathcal{E}\left[\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{t=-k}^{-1}\sum_{s=-m}^{-1} g(k)e^{-j\omega k}g(m)e^{j\omega m}u(t)e^{-j\omega t}u(s)e^{j\omega s}\right]\\ &= \frac{2\sigma_u^2}{2\pi M}\left[\sum_{k=0}^{\infty} g(k)e^{-j\omega k}\left(\sum_{m=0}^{k-1} mg(m)e^{j\omega m} + \sum_{m=k}^{\infty} kg(m)e^{j\omega m}\right)\right]\end{aligned} \tag{16}$$

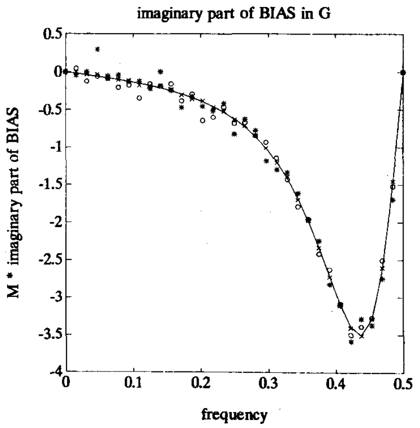


Fig. 1. Imaginary part of the bias of the three transfer function estimators from 64 observations of the example process with stochastic input. The drawn line represents the theoretical result of (19) with $M = N$; oo and $\times\times$ correspond to N times the bias in simulations in G_1 and G_2, respectively, and $**$ is KN times the bias of G_3 in simulations.

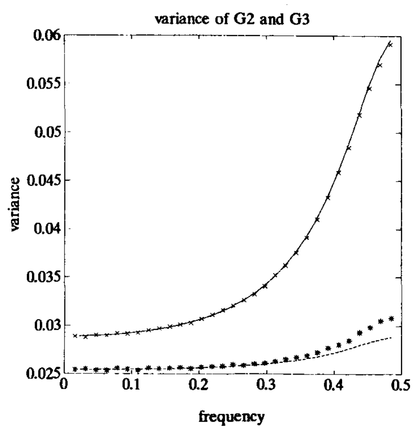


Fig. 2. Variance of G_2 and G_3 with the theoretical results of (19) as drawn lines and $\times\times$ and $**$ for the simulations of G_2 and G_3, respectively.

complex numbers. Averaging has no effect on the bias, and additive noise will not influence the bias.

The variance of $G_1(e^{j\omega})$ did not converge to any fixed value, not even without additive noise. The sampling inaccuracy of $G_1(e^{j\omega})$ turned out to be much greater than that of the other two estimators, which are given in more detail in Fig. 2. Equation (19) is accurate for the variance of $G_2(e^{j\omega})$, and also other simulations showed the same sort of accuracy. As K equals 9 in this example, the difference between $1/K$ and $1/K-1$ will be 11%, so the agreement between the drawn theoretical lines and the marks of the simulations shows the accuracy of $K-1$ as divisor. The variance of $G_3(e^{j\omega})$ is roughly K times smaller than that of $G_2(e^{j\omega})$ in the case without additive noise. The accuracy of the formula is best for small K and for large N.

Summarizing it can be stated that the formulas for bias and variance describe the actual behavior in simulations with stochastic input. Averaging with $G_1(e^{j\omega})$ is not advisable, because the theoretical variance of the estimated transfer function is infinite, with the practical consequence that an estimate will often exhibit some very inaccurate points in the frequency domain. Averaging with $G_2(e^{j\omega})$ yields better results, but the most accurate results are definitely obtained with $G_3(e^{j\omega})$. For a periodical input signal, all three estimators give identical results: no bias is present, and the smaller variance becomes h_v/Kh_u.

V. Conclusions

Expressions have been derived for the bias and variance due to end effects in transfer function estimates with stationary stochastic excitations. It turns out that the equivalent of one degree of freedom is lost in the estimation of an average transfer function with stochastic excitation. Hence, the result of averaging the quotient of single Fourier transforms of output and input gives a poor estimate. The quotient of the average cross-spectrum estimate over subframes and the estimated input spectrum yields a better result, but the most accurate result is found by doing the averaging with a window, after the Fourier transform. It is remarkable that averaging in transfer function estimation loses the equivalent of one degree of freedom for stochastic inputs in contrast with periodical or deterministic inputs.

Appendix

Given are K independent, real, zero-mean, normally distributed random variables x_i, $i = 1, \cdots, K$, each with unit variance. The variance of $x_1/(x_1^2 + x_2^2 + \cdots + x_K^2)$ can be written as

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (2\pi)^{-K/2} \left(\frac{x_1}{x_1^2 + \cdots + x_K^2} \right)^2 e^{-(x_1^2 + \cdots + x_K^2)/2} \, dV$$

with $dV = dx_1 dx_2 \cdots dx_K$. By transforming the integral to spherical coordinates and using the mathematical results:

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dV = \int_0^{\infty} dR \int_0^{2\pi} d\phi_0 \int_0^{\pi} \cdots \int_0^{\pi} d\phi_1 \cdots d\phi_{K-2} R^{K-1} \cdot \prod_{j=1}^{K-2} (\sin\, \phi_j)^{K-1-j}$$

$$\int_0^{\pi} \sin^K \phi d\phi = \sqrt{\pi} \frac{\Gamma(\frac{K+1}{2})}{\Gamma(\frac{K+2}{2})} \qquad z\Gamma(z) = \Gamma(z+1)$$

$$\int_0^{\infty} R^K e^{-R^2/2} dR = 2^{\frac{K-1}{2}} \Gamma(\frac{K+1}{2})$$

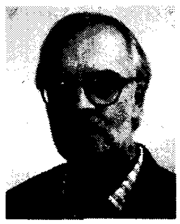
it follows for $K \geq 3$ that

$$\operatorname{var} \frac{x_1}{x_1^2 + \cdots + x_K^2} = \frac{1}{K(K-2)}. \tag{A1}$$

For $K = 2$, $\operatorname{var}\{x_1/(x_1^2 + x_2^2)\}$ becomes ∞ because the integrand for dR contains R^{-1} which gives an infinite contribution when integrated from zero.

REFERENCES

[1] G. C. Goodwin, M. Gevers, and B. Ninness, "Quantifying the error in estimated transfer functions with application to model order selection," *IEEE Trans. Automat. Contr.*, vol. 37, no. 7, pp. 913–928, 1992.
[2] B. Wahlberg and L. Ljung, "Design variables for bias distribution in transfer function estimation," *IEEE Trans. Automat. Contr.*, vol. AC-31, no. 2, pp. 134–144, 1986.
[3] L. Ljung, *System Identification: Theory for the User.* Englewood Cliffs, NJ: Prentice-Hall, 1987.
[4] ——, "On the estimation of transfer functions," *Automatica*, vol. 21, no 6, pp. 677–696, 1985.
[5] M. B. Priestly, *Spectral Analysis and Time Series.* London, UK: Academic Press, 1981.
[6] R. L. Kosut, "Adaptive control via parameter set estimation," *Int. J. Adaptive Contr. Sig. Process.*, vol. 2, no. 4, pp. 371–399, 1988.
[7] P. E. Wellstead, "Non-parametric methods of system identification," *Automatica*, vol. 17, no. 1, pp. 55–69, 1981.
[8] P. M. T. Broersen, "On bias in transfer functions estimated with stochastic excitation," in *Proc. ICASSP 1994 Conf.* (Adelaide, Australia, 1994), vol. IV, pp. 49–52.
[9] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, vol 1. London, UK: Charles Griffin & Co., 1963.
[10] K. S. Miller, *Complex Stochastic Processes, An Introduction to Theory and Application.* Reading, MA: Addison-Wesley, 1974.



Piet M. T. Broersen was born in Zijdewind, The Netherlands, in 1944. He received the M.Sc. degree in applied physics in 1968 and the Ph.D. degree in 1976, both from the Delft University of Technology, Delft, The Netherlands.

He is with the Department of Applied Physics of the Delft University of Technology. His research interests are the application of signal processing and parameter estimation to model building.