

# Variance of Fourier Coefficients Calculated from Overlapped Signal Segments for System Identification

László Balogh\*, István Kollár\* and Gregory Gueret\*\*

\*Department of Measurement and Information Systems  
Budapest University of Technology and Economics  
Budapest, Hungary, H-1521

\*\*Ecole Supérieure d'Electronique de l'Ouest, Angers, France  
Phone: +36 1 463-[3583,1774], +33 241 810-692, fax: +36 1 463-4112  
Email: lbalogh@mit.bme.hu, kollar@mit.bme.hu, walrus@free.fr

**Abstract** – In frequency domain system identification, the Fourier coefficients and their variances are used to calculate the estimated parameters. They are calculated from the measured signals. Therefore, good estimation of the Fourier coefficients and of their variance is of utmost importance. It is a common practice to segment the periodic signal into periods, calculate the Fourier coefficients in each period, average them, and perform variance analysis, too. Here an important question arises: do we use all the information present in the signal? Can we improve this procedure? This paper analyzes the information utilization in the process, examines when it is optimal, explains when processing of overlapped segments can help and why, and suggests an improved preprocessing procedure, increasing the quality of the calculated Fourier coefficients.

**Keywords** – parameter estimation, variance reduction, overlapped segments, frequency domain system identification, Fourier coefficient, variance analysis.

## I. INTRODUCTION

The widely accepted way of preprocessing data for frequency domain system identification is [1] to

- measure the steady-state response and the input signal of a system using periodic excitation,
- slice up the input and output signals into periods,
- calculate the Fourier coefficients of each period at the excited frequencies,
- perform averaging and variance analysis over the calculated Fourier coefficients.

The natural desire is to obtain the best possible results by this, at least in terms of bias and variance of both the Fourier coefficients and of the variances. Therefore, the questions arise

- whether we utilize the full information present in the signal,

<sup>1</sup> This work was supported by the grant FKFP 0074/2001, furthermore 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 Ministers Office, Science Policy Programming.

- whether by using more periods by processing overlapped segments can help, and
- whether we can somehow improve this preprocessing.

The questions are even more challenging if we consider the well-known paper of Welch [2] which advocates the use of overlapped segments for the calculation of the periodogram. Since the estimation of the periodogram is equivalent to the estimation of the variance of Fourier coefficients, it seems to be logical that if the use of overlapped segments brings benefit in the estimation of the periodogram, it is beneficial in our case, too. This question will be discussed in detail.

In this paper first of all the basic notations and assumptions are discussed. Then the maximum likelihood estimates are studied for some cases where calculations are relatively simple (coherent sampling, additive white noise). After this, consequences for more general cases are discussed, and an improved procedure for colored noise is introduced. Last, simulation results verify the theoretical results.

## II. BASIC NOTATIONS

We assume that we have a measured periodic signal, with additive observation noise:

$$x(i) = x_0(i) + n(i),$$

where  $x_0(i)$  is a sample of the noiseless signal, and  $n(i)$  is a sample of the observation noise.

The analyzed periodic signal  $x_0$  is a multisine, that is, it contains certain harmonics of the basic frequency. As a rule, we measure several periods, but at least two periods of the input. The periodic signal components are measured properly, that is, the sampling theorem is fulfilled, and no overlapping occurs.

*Coherent sampling* means that the samples of  $x_0$  cover an integer number of periods of the signal. In the simplest case we will further assume that  $x_0(i)$  is periodic with the same period

length as the continuous-time periodic signal. In the general case,  $x_0(i)$  is not periodic as a sequence (although it is taken from a continuous-time periodic signal), and does not cover an integer number of periods of the continuous-time signal.

Concerning the a priori knowledge, we assume here that the period length (basic frequency) of the signal is exactly known, so we need not extract this from the measured signal. This is not always true: an effective approach to the processing of noisy periodic signals with unknown period length is discussed at this conference in [3].

As for the noise, in the simplest case we assume that  $n(i)$  is white. Later, we will generalize the results for colored noise. In the paper we assume that the noise is Gaussian, because this is the most often used noise model. For other distributions a similar path of thoughts can be followed as here.

### III. GENERAL CONSIDERATIONS

If we have the samples  $x(i)$ , in order to estimate the Fourier coefficients, we need to fit the sequence with the sum of the harmonics present. As a rule of thumb, maximum likelihood estimation is to be used.

If the noise is Gaussian, for the maximum likelihood estimation we need to minimize the following least squares cost function [4]

$$\sum_{i=0}^{L-1} W_i \left( x(i) - \sum_{k=1}^F A_k \cos(2\pi f_k i T) + B_k \sin(2\pi f_k i T) \right)^2 \quad (1)$$

in terms of  $A_k, B_k$ , where  $T$  is the sampling interval. If the excitation is missing at certain frequencies, the corresponding coefficients  $A_k, B_k$  are set to zero.

The weights  $W_i$  are uniform for white noise, and are determined by the autocorrelation function of the noise if the noise is colored. Estimation of the Fourier coefficients is equivalent to the determination of  $2F$  real numbers or  $F$  complex numbers.

It was shown by [5] that for uniform weighting the minimum of (1) is reached by just the DFT (FFT) of the noisy series at the given frequencies ( $A_k = 2\text{Real}(X_k), B_k = 2\text{Imag}(X_k)$ ).

Since the above estimate is a Gauss-Markov estimate of the Fourier coefficients [4], it is also efficient: it is the minimum variance unbiased estimate, so we cannot perform better in our terms.

To compute the variance of the Fourier coefficients is equivalent to the computation of the periodogram of the noise, since the variance of the noise is essentially the same as the power of

the noise at the given frequencies. We do this by taking the periodogram of the difference between the measured signal and the estimated multisine – this is not precisely equal to the above periodogram, since the signal model is only an estimate, but it is close to it for  $L \gg F$ .

Below we will show that – at least for a selected simple case – this is again the ML estimate of the variance, so we cannot do better. But before doing this, let us discuss what is the effect of segmentation.

#### A. Segmentation without overlap

Let  $x(i), i = 0, \dots, L$  denote the samples as before. These are obtained via sampling a stationary, stochastic series. Let us assume here that original signal is periodic with period length  $T_p = NT$ , and  $L$  is an integer multiple of  $N$ . We slice up the signal into  $M = L/N$  segments (periods), and process each segment separately. This means that we perform DFT (FFT) on each segment, and take the average of the corresponding points to obtain the Fourier coefficients, and last we calculate the periodogram period by period and take the average of these periodograms.

It is straightforward to see that – apart from a scaling factor – the obtained Fourier coefficients are exactly the same as above (cf.  $L = MN$ ):

$$\begin{aligned} X_{kM} &= \sum_{i=0}^{L-1} x(i) e^{-2\pi \frac{(kM)i}{L}} \\ &= \sum_{m=0}^{M-1} \sum_{i=0}^{N-1} x(i + mN) e^{-2\pi \frac{k(i+mN)}{N}} \\ &= M \left( \frac{1}{M} \sum_{m=0}^{M-1} X_{k,m} \right). \end{aligned} \quad (2)$$

This means that *we cannot make anything better than averaging the corresponding Fourier coefficients from no-overlapped segments.*

For the case of colored noise (weighted LS, see (1)) this is not that straightforward, but common sense says the two ways of calculation very closely correspond again to each other. We will not dwell on this any further.

The calculation of the variance needs further simplifications.

#### B. Determination of the Variance for White Noise

As we stated above, we need to evaluate the periodogram of the residuals. The problem we face is that for continuous-spectrum signals, like the noise here, the periodogram of the  $L$ -point periodogram contains spectrum estimates at  $L/2 + 1$

points (from 0 to  $0.5/T$ , the end points inclusive), while the  $M$  periodograms of length  $L/M$  contain spectrum estimates at  $L/M/2 + 1$  points (also from 0 to  $0.5/T$ ). Therefore, the two estimations are incompatible, so we cannot directly compare them. We can argue that in the latter case we estimate a more smoothed spectrum than in the first one, but this does not bring us much closer. Therefore, we will make a further simplification. Let us assume that the noise is white, and its spectrum is described by a single parameter (e.g. let the constant value of the discrete-time spectrum is  $S$ ). Let us determine the ML estimate of  $S$  in both cases.

The maximum likelihood estimates of  $\mu$  and  $\sigma^2$  of a white Gaussian series  $n_i, i = 0 \dots L - 1$  are known to be

$$\hat{\mu} = \bar{n} = \frac{1}{L} \sum_{i=0}^{L-1} n_i, \quad (3)$$

$$\widehat{\sigma^2} = s^2 = \frac{1}{L} \sum_{i=0}^{L-1} (n_i - \bar{n})^2. \quad (4)$$

It is also known that (4) has  $\sigma^2 \chi_{L-1}^2$  distribution, with expected value  $\frac{L-1}{L} \sigma^2$ .

Therefore, since by neglecting the points at  $k = 0$  and  $k = L/2$  we have  $L/2 - 1$  complex points ( $X_k, k = 1 \dots L/2 - 1$ ), and since both the real and imaginary parts behave like  $n_i$  above, we can form the estimate for the real and for the imaginary parts, and average the two. This will be analyzed further.

We have the possibility to use two different estimates: one with non-segmented processing (long series), and one with segmented processing. The first one is our reference value (the ML one). We are going to prove that the second one is close to it, therefore, it is also ‘almost’ optimal.

For the long series, the maximum likelihood estimates are:

$$\widehat{C}_{k_i 1} = X_{k_i M}, \quad k_i = k_1 \dots k_F. \quad (5)$$

This gives the Fourier coefficients at the selected frequencies.

The variance is somewhat more complicated to determine. If we do not segment, the only reasonable choice is to assume that  $L \gg F$ , and to determine the Fourier amplitudes from the excited lines (see (5)), while determining the noise variance from the non-excited lines:

$$\widehat{S}_1^2 = \frac{1}{2(L/2 - 1) - 2F} \sum_{k \neq k_i M} |X_k|^2, \quad (6)$$

with  $X_k$  being the DFT result at index  $k$ .

The sum has  $\sigma^2 \chi_{L-2-2F}^2$  distribution, the expected value is exactly  $\sigma^2$ .

For the segmented case, similarly:

$$\widehat{C}_{k_i 2} = \frac{1}{M} \sum_{m=0}^{M-1} X_{k_i m} = \widehat{C}_{k_i 1}, \quad (7)$$

the estimate is the same as (5), like we discussed above, and

$$\begin{aligned} \widehat{S}_2^2 &= \frac{1}{M(N-2)} \sum_{m=1}^M \left( \sum_{i=1}^F |X_{k_i, m} - \widehat{C}_{k_i 2}|^2 \right. \\ &\quad \left. + \sum_{k \neq k_i} |X_{k, m}|^2 \right), \end{aligned} \quad (8)$$

The sum has  $\sigma^2 \chi_{M(N-2)-2F}^2 = \sigma^2 \chi_{L-2M-2F}^2$  distribution, the expected value is  $\frac{M(N-2)-2F}{M(N-2)} \sigma^2 = \frac{L-2M-2F}{L-2M} \sigma^2$ .

Now, at least for this simple case, we are ready to answer the basic questions. From the segments covering one period without overlapping, we can determine maximum likelihood estimates of the Fourier coefficients, and estimates with asymptotically 0 bias and approximately the same variance ( $L - 2 - 2F \approx L - 2M - 2F$ ) as the maximum likelihood estimate of  $S$ . Therefore, at least in this case, *no benefit can be obtained from any other procedure, including processing of overlapped segments*.

How are these results compatible with Welch’s ones? He strongly advocates overlapping. Is there any important difference between the approaches? The answer is yes. In our case we used no windowing, since we concentrated to the determination of the Fourier coefficients. But it is also possible to extract the periodic component, and THEN make a noise analysis with windowed and overlapped segments. The benefit is achieved by using a different model: by using short, windowed segments Welch estimates the points of a smooth spectrum, that is, it uses the information concerning the neighboring points, too. This is certainly better than using the variance *at the excitation frequencies only*.

In order to discuss this, we will discuss shortly how to process the Fourier coefficients for overlapped segments.

### C. Processing non-synchronized segments

As we have seen above, during the preprocessing of data for frequency domain system identification, we need to average

estimates of Fourier coefficients. This is straightforward when each segment corresponds to a period, since then the Fourier coefficients all have the same phase (apart from the effect of the noise). However, when

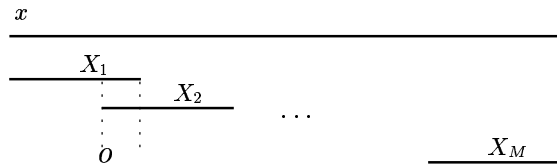
- overlapped segments are processed, or
- the segment length differs from an integer multiple of the period length,

the periodic content in each segment will have a different phase. Since the Fourier coefficients are complex numbers, the changing phase destroys the result of averaging. In Welch's method this was not a problem because we used the absolute square of the coefficients, so the phase was eliminated. However, we perform here linear averaging, so the phase is crucial. Therefore, we need to correct the phase of each Fourier coefficient before averaging, in order to assure that the phases of the corresponding coefficients are identical.

When the period length is  $T_p$ , and the beginning time of the  $m$ th segment is  $T_{0m}$ , correction is needed. It is easy to see that synchronization occurs when

$$(T_{0m} - T_{01}) \bmod(T_p) = 0. \quad (9)$$

If this is not true, let  $g_i$  be



$$\varphi_m \equiv \frac{(T_{0m} - T_{01}) \bmod(T_p)}{T_p} 2\pi.$$

Then the corrected  $X_{km}^*$  (the  $k$ th coefficient in the  $m$ th segment) is

$$X_{km}^* = X_i(k) e^{j\varphi_m k}.$$

Now we can average the corrected Fourier coefficients, using  $X_k^*$  instead of  $X_k$ .

#### D. Non-coherent sampling

In this case we cannot assume that the sampling frequency is the integer multiple of the fundamental frequency component of the signal. The properties of the DFT (leakage, picket-fence)

show that it is better to use the least squares (LS) solution. The LS estimation of the Fourier coefficients is the solution of the following linear equation:

$$U \begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_N \end{bmatrix} = x$$

where the matrix  $U$  is defined as

$$U = \begin{bmatrix} 1 & 1 & \dots & \dots \\ 1 & \cos(-2\pi \frac{1}{N}) & \sin(-2\pi \frac{1}{N}) & \dots \\ 1 & \cos(-2\pi \frac{2}{N}) & \sin(-2\pi \frac{2}{N}) & \dots \\ \dots & \dots & \dots & \dots \\ 1 & \cos(-2\pi \frac{N-1}{N}) & \sin(-2\pi \frac{N-1}{N}) & \dots \end{bmatrix}$$

Here vector  $p$  contains the parameters  $A_k$  and  $B_k$  in equation (1). This matrix  $U$  contains also a DC term because it helps to eliminate the offset, but in this paper we will not make use of this.

From the theoretical point of view we can determine the analytical solution of the equation above. Furthermore, the stochastic properties of the solution can also be derived. We know that the model of the measurement and signal is the following:

$$x = Up + n$$

where  $n$  is the noise. The LS solution minimizes the following cost function:

$$(x - Up)^T (x - Up),$$

which is the vector form of the equation (1). The solution of the above problem can be expressed analytically:

$$\hat{p} = (U^T U)^{-1} U^T x. \quad (10)$$

Its variance is:

$$(U^T U)^{-1} U^T \Sigma U (U^T U)^{-1}. \quad (11)$$

In this section we assume that the distribution of the noise is white Gaussian, therefore the solution of the minimization of the cost function above is Gaussian, too, with the following properties:

- the expected value of the estimate is  $p$  (unbiased estimation, moreover the minimum variance linear unbiased estimate),
- the variance is  $\Sigma_{\hat{p}} = \sigma^2 (U^T U)^{-1}$  ( $\sigma^2$  is the variance of the white Gaussian noise).

It is possible to find a formula for the variance estimation, too.

$$\widehat{\sigma^2} = \frac{1}{N} x^T (I - U(U^T U)^{-1} U^T) x.$$

The distribution of  $\widehat{\sigma^2}$  is independent of  $\hat{p}$  and is  $\sigma^2 \chi^2/N$ .

#### IV. COLORED GAUSSIAN NOISE

In this section we assume that we have additive Gaussian colored noise. Therefore we can estimate the variance with using Welch's method. This is a very efficient method still in the case of non-coherent sampling.

In this case there are problems with theoretical expressions. Because of the adjacent correlated samples it is not easy to specify some of similar expression as above. The colored noise causes that the adjacent segments are correlated, too. From this fact it follows that in the case of the estimation of either the periodogram or the Fourier coefficients we lose the advantage of adding up independent random variables.

##### *The algorithm*

The proposed algorithm to compute the Fourier coefficients and its variances is the following:

1. Determine the Fourier coefficients without using overlap.
2. Calculate the noise.
3. Estimate the power spectra of the noise using Welch's method.
4. Compute the auto covariance function of the noise.
5. Determine the Fourier coefficients with using the estimated auto covariance of the colored noise (maybe use overlap, as in Welch's method).
6. Maybe repeat from 2.

Consequently, in the  $i$ th iteration step the following relationships are valid:

- the noise is (1)

$$\hat{n} = x - U\hat{p}_i,$$

- the auto covariance function is (4)

$$\hat{\Sigma}_{i+1}(k) = \left[ \frac{k}{M} \right] \sum_{l,m, |l-m|=k} \hat{n}_l \hat{n}_m,$$

where  $\lfloor \cdot \rfloor$  denotes the integer part,

- the parameter is (5)

$$\hat{p}_{i+1} = (U^T \hat{\Sigma}_i^{-1} U)^{-1} U^T \hat{\Sigma}_i^{-1} x. \quad (12)$$

Some remarks to the algorithm:

- 'Calculate the noise' means that the estimated signal is subtracted from the original signal.
- In step 3. we use Welch's method with the Fourier coefficients computed beforehand either in step 1. or in step 5. (depends on the situation: if we are the first or later cycle).

- In Welch's method it is very important that a windowing function is used. Thus the periodogram estimation can be more efficient than without windowing. Hence the variance of the estimation of the auto covariance function can be reduced. The reason is that in this case we make use of the fact that the original spectra is continuous.
- It is not necessary to wait for the small change of the Fourier coefficients, the algorithm can also be used with fixed iteration number.
- At the end of the algorithm small a change means that the sum of the absolute values of difference of complex numbers is small.
- (12) is better for colored noise than using (10) (FFT): its variance  $(U^T \Sigma^{-1} U)^{-1}$  is smaller than (11).

The advantage of this algorithm is that it works even with non-coherent sampling. Similarly to the white noise, if we use Welch's method, the problem of averaging of complex numbers does not arise.

Furthermore, by estimating the covariance of the noise the estimation of the Fourier coefficients can be more accurate. Thus the estimation of the variance may be better, too.

The disadvantage of the algorithm are that it is more difficult to implement than FFT. It means that for an on-line application, for example on a DSP processor, where we have time limit for the procedures, may not be practicable. But it is not a real limitation for the system identification process where the signal processing and the identification algorithms are off-line.

#### V. SIMULATION RESULTS

In this section we show simulation results to confirm the theoretical statements above. The section is split into two main parts. In the first one we study the case when the signal contains only noise. Thus we need only variance estimation because we assume that the expected value is known, it is exactly 0. In the second part we show simulation results when the studied signal contains a periodic component and noise. In this case the estimation of the Fourier coefficients has to be done.

##### *A. Only second order moments estimation*

If we study only the noise (noise analysis of the system) then it is not necessary to estimate the Fourier coefficients corresponding to the assumption of the first part of the paper.

##### *Only white noise*

In this simulation we set the following parameters:

$$M = 1024 \quad N = 64 \quad \text{var}\{n\} = 1 \quad o = 0$$

We have obtained from the Monte-Carlo simulations that the expected value of the variance estimation is about 1, as we

expected. Furthermore, the Monte-Carlo simulations showed that the estimated Fourier coefficients tend to zero as  $M$  is increased.

Simulation results showed that in the case of white noise the variance of the Fourier coefficients and the periodogram cannot be decreased by increasing the overlap.

#### Only colored noise

In this case we generated the noise by filtering a white noise process. The other parameters were

$$M = 512 \quad N = 64 \quad \text{var}\{n\} = 1 \quad o = 0.$$

Hence we concluded that the method works well, i.e. the auto covariance function estimation gives asymptotically the correct result.

#### B. First and second order estimation

In this subsection we study the case when the  $x$  contains both signal and additive noise. In the first example we consider a simple sine wave and coherent sampling.

The parameters of the simulations:

$$M = 512 \quad N = 64 \quad \text{var}\{n\} = 1 \quad o = 0$$

Simulations confirmed the theoretical results. Applying Monte-Carlo simulations we observed that the variance estimation can be more accurate with using our proposed algorithm. Our experience is that about three iteration steps enough to get acceptable result.

## VI. CONCLUSIONS

In this paper different possible ways of the analysis of periodic signals have been studied. The main goal is to determine the Fourier coefficients and its variances for the system identification process. We saw that different assumptions lead to different estimators. We could obtain theoretical results only for the simple cases (independent Gaussian white noise). For other cases we have presented an iterative algorithm. Thus, for system identification we have a chance to increase the accuracy, so that the system identification algorithm will give better result. This paper contains a proposed algorithm which can decrease the variance both of the Fourier coefficient and of the estimated variance.

## References

[1] I. Kollár, *Frequency Domain System Identification Toolbox for Matlab*, The MathWorks, Inc., Natick, MA, 1994, current version: <http://elecwww.vub.ac.be/fdident/>.

[2] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. on Audio and Electroacoustics*, vol. AU-15, no. 2, pp. 70–73, June 1967.

[3] J. Schoukens, Y. Rolain, G. Simon, and R. Pintelon, "Fully automated spectral analysis of periodic signals," in *IEEE Instrumentation and Measurement Technology Conference*, Anchorage, AK, USA, May 2002.

[4] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, vol. 2, Edward Arnold, 1991.

[5] A. van den Bos, "Estimation of Fourier coefficients," *IEEE Trans. on Instrumentation and Measurement*, vol. 38, no. 5, 1989.

[6] D. R. Brillinger, *Time Series, Data Analysis and Theory*, McGraw-Hill, 1981.

[7] L. Schnell, *Technology of Electrical Measurements*, Wiley, 1993.

[8] A. van den Bos, "Estimation of complex parameters," in *Sysid'94, X. IFAC/IFORS International Symposium on System Identification and Parameter Estimation*, Copenhagen, 1994, vol. 3, pp. 495–9.

[9] M. B. Priestley, *Spectral Analysis and Time Series*, Academic Press, 1981.