Statistical error due to finite-time averaging

B. Picinbono
Laboratoire des Signaux et Systèmes, ESE, Plateau du Moulon 91190 Gif-Yvette, France

(Received 19 May 1977)

Previous results concerning statistical errors due to finite-time T averaging in many experiments are made more precise. In particular, we show that the error does not always decrease according to a $T^{-1/2}$ law as $T\to\infty$. In the case of power or correlation function measurements, the departure from a $T^{-1/2}$ law may be due to an unbounded spectral density for some frequencies, which appears for some physical noises.

The statistical errors due to the replacement of ensemble averages by averages over finite-time intervals (T) in the evaluation for the correlation functions of stochastic processes were recently given.[1] These results were presented in the context of computer experiments and applied to the case of liquid sodium.[2] They are, however, very general. According to Refs. 1 and 2, the statistical error is of the order of $T^{-1/2}$ for large T, with a coefficient of the square root related to some characteristic relaxation time.

Zwanzig and Ailawadi[1,2] considered Gaussian processes and stated the following with respect to non-Gaussian processes: "Because this is not necessarily true when $A(t)$[originally assumed to be a Gaussian random variable] is a dynamical quantity, our results are expected to be plausible estimates, but not rigorous. At the present time we do not know of any way to correct for non-Gaussian behavior."[1] In the present paper, we show that the error is not always of the order of $T^{-1/2}$, *even in the Gaussian case*. Our calculations are applicable to any moment of the process.

Let us consider a real second-order stationary stochastic process (SP) $X(t)$. This SP is, for example, the dynamical quantity $A(t)$ considered in previous papers[1,2], or any function of this quantity. At this state of the calculation, physical interpretations of $X(t)$ are not required. In what follows, the mean value $\langle X(t)\rangle$ of $X(t)$ is denoted by m_x. The average of $X(t)$ over a time interval T is

$$X_T(t)=\frac{1}{T}\int_{t-T}^{t} X(\theta)\,d\theta\,. \tag{1}$$

This function is obtained from $X(t)$ by linear filtering. This is a perfect time averager, sometimes called a cardinal filter because the modulus of its complex gain is[3]

$$G_T(\nu)=(\sin\pi\nu T)/\pi\nu T\,. \tag{2}$$

It is clear that

$$\langle X_T(t)\rangle = m_x\,, \tag{3}$$

meaning that $X_T(t)$ is an unbiased estimator of the mean value m_x of $X(t)$. Using standard methods, the error squared defined by

$$\epsilon^2(T)\equiv\langle[X_T(t)-m_x]^2\rangle \tag{4}$$

is found to be[4]

$$\epsilon^2(T)=\int_{-\infty}^{+\infty}\left(\frac{\sin\pi\nu T}{\pi\nu T}\right)^2\gamma_x(\nu)\,d\nu\,, \tag{5}$$

where $\gamma_x(\nu)$ denotes the power spectrum of the SP $X(t)$-m_x.

We are actually interested in the behavior of $\epsilon^2(T)$ for large values of T. Since the analysis of this behavior has been previously reported,[5,6] we shall present here only the main results.

It follows from Eq. (5) that for large values of T, the only important frequencies are those in the vicinity of $\nu=0$. Thus let us suppose that, for $|\nu|\le a$,

$$\gamma_x(\nu)=|\nu|^m f(\nu), \tag{6}$$

where $f(\nu)$ is a continuous nonzero bounded function. Because $X(t)$ is a second-order SP, we have $m>-1$. The case $m=-1$ is beyond the scope of this study, even though it may correspond to physical noise, e.g., the "flicker noise." We write

$$\epsilon^2(T)=\epsilon_a^2(T)+\epsilon_a'^2(T), \tag{7}$$

where ϵ_a^2 and $\epsilon_a'^2$ are the integrals of $\gamma_x(\nu)$ over the intervals $|\nu|\le a$ and $|\nu|>a$, respectively. Since the integral of $\gamma_x(\nu)$ over all frequencies is, by definition, the variance $\sigma_x^2<+\infty$, we have

$$\epsilon_a'^2(T)\le\sigma_x^2(\pi a)^{-2}T^{-2}. \tag{8}$$

The calculation of $\epsilon_a^2(T)$ is performed in Ref. 5 or 6. Combining this result with Eqs. (7) and (8), we obtain the behavior of $\epsilon^2(T)$ for large values of T. If $-1<m<+1$, we obtain

$$\epsilon^2(T)\simeq c(m)T^{-(m+1)} \tag{9}$$

and, if $m\ge1$,

$$\epsilon^2(T)\simeq c(m)T^{-2}\,. \tag{10}$$

Thus the result presented in Ref. 1 is valid only when if $m=0$, that is, when $\gamma_x(0)$ is finite and nonzero.

Let us briefly consider this special case. Starting from Eq. (5), we have, for large T,

$$\epsilon^2(T) \simeq \gamma_x(0)T^{-1} . \tag{11}$$

Moreover, since $\gamma_x(\nu)$ is the Fourier transform of the correlation function $\Gamma(\tau) \equiv \langle X(t)X(t-\tau)\rangle$, we have

$$\gamma_x(0) = \int_{-\infty}^{+\infty} \Gamma_x(\tau)\, d\tau = \sigma_x^2 \int_{-\infty}^{+\infty} \frac{\Gamma_x(\tau)}{\Gamma_x(0)}\, d\tau = \sigma_x^2 t_c . \tag{12}$$

The expression on the right-hand side (rhs) of Eq. (12) can be viewed as the definition of the correlation time t_c of the process $X(t)$. This definition is particularly convenient in the case of exponentially decreasing correlation functions. Thus the error squared is

$$\epsilon^2(T) \simeq \sigma_x^2 t_c T^{-1} . \tag{13}$$

It is important to note that the decrease of ϵ^2 as $T\to\infty$ does not necessarily behave as T^{-1}. From Eq. (9), it follows that, as $m \to -1$, the error in fact becomes independent of T. In that case, time averaging is useless. Conversely, it is not possible to obtain an error decreasing faster than T^{-2}.

Let now the previous results be applied to the measurement of the correlation function of a dynamical quantity $A(t)$.[1,2] An unbiased estimate of its correlation function is obtained by using Eq. (1), where the function $X(t)$ is expressed in terms of $A(t)$ by

$$X(t) = A(t)A(t-\tau) . \tag{14}$$

In order to apply the previous results, it is necessary to evaluate the behavior of the power spectral density $\gamma_x(\nu)$ for $\nu \sim 0$. As done in Ref. 1, we simplify the discussion by considering the case $\tau = 0$. We thus estimate the power, or variance, of $A(t)$.

Suppose first that $A(t)$ is Gaussian. The correlation function of $X(t) - \langle X(t)\rangle$ is given by

$$\Gamma_x(\tau) = 2\Gamma_A^2(\tau) , \tag{15}$$

where $\Gamma_A(\tau)$ is the correlation function of $A(t)$. By Fourier transformation we obtain

$$\gamma_x(\nu) = 2\int_{-\infty}^{+\infty} \gamma_A(n)\gamma_A(\nu - n)\, dn , \tag{16}$$

and $\gamma_x(0)$ can be written

$$\gamma_x(0) = 2\int_{-\infty}^{+\infty} \gamma_A^2(\nu)\, d\nu \tag{17}$$

because $\gamma(\nu) = \gamma(-\nu)$ for real SP. This integral cannot be zero, but it can be infinite. Thus the parameter m of Eq. (6) must satisfy the inequalities

$$-1 < m \le 0 . \tag{18}$$

If $m = 0$, $\gamma_x(0)$ is finite and the error is that given in Eq. (11). That is the result of Ref. 1.

Let us now consider the case where the integral on the rhs of (17) is infinite. Because $A(t)$ is a second-order SP, $\gamma_A(\nu)$ is integrable and $\gamma_x(\nu)$ can become infinite only if, for some frequencies ν_i, the power spectral density of $A(t)$ is unbounded. This situation is encountered in reality. It is encountered, for example, in the flicker noise, in which case $\nu_i = 0$.

To make the above discussion more precise, let us consider the case where there is only one singular frequency. Let us assume further that, for $|\nu - \nu_i| \le a$,

$$\gamma_A(\nu) = |\nu - \nu_i|^{m_i} f_i(\nu) , \tag{19}$$

where $f_i(\nu)$ is a nonzero continuous bounded function. Evidently, since there is only one frequency ν_i, we have $\gamma_A(x) \le B$ for $|\nu - \nu_i| > a$.

Using this expression to calculate the integral in (17), we see easily that $\gamma_x(0)$ is infinite if $m_i < -\frac{1}{2}$. Because $\gamma_A(\nu)$ is integrable as $A(t)$ is of second order, we have $m_i > -1$. We now study the behavior of $\gamma_x(\nu)$ as $\nu \to 0$ when

$$-1 < m_i < -\tfrac{1}{2} . \tag{20}$$

If $m_i > -\frac{1}{2}$, $\gamma_x(0)$ is finite and the error is that given by Eq. (11). The detailed calculations presented in Appendix A show that

$$\gamma_x(\nu) \simeq c_i |\nu|^{2m_i+1} \tag{21}$$

as $\nu \to 0$ with condition (20).

Using Eqs. (6) and (9), we conclude that the error in the measurement of the power of $A(t)$ is

$$\begin{aligned} \epsilon^2(T) &\simeq c(m_i)T^{-2(m_i+1)}, \quad -1 < m_i < -\tfrac{1}{2} \\ &\simeq cT^{-1}, \quad m_i > -\tfrac{1}{2} \end{aligned} \tag{22}$$

if there is one frequency ν_i where the power spectrum has the behavior in Eq. (19). When there is more than one frequency where $\gamma_A(\nu)$ is infinite, the result follows from Eq. (22) by addition.

As suggested in Ref. 1, it is interesting to estimate the importance of the Gaussian assumption. Unfortunately, the non-Gaussian behaviors which appear in many physical problems do not uniquely specify the SP, and it is therefore impossible to provide general results. Nevertheless, results can be given for some non-Gaussian statistics, in particular for spherically invariant processes,[7] which are natural generalizations of Gaussian processes.

We can deduce two kinds of information from Eqs. (13) or (22): first, the exponent of T on the behavior of $\epsilon^2(T)$ for large T, and second, the coefficient c, which is $\sigma_x^2 t_c$ for T^{-1}. As far as the exponent of T is concerned, the Gaussian assumption is not essential. Indeed, consider a dynamical quantity $A(t)$, not necessarily Gaussian. We study the measurement of the mean value of $X(t) = A^2(t)$ by time averaging. Following the results of the first part of this Comment, the error $\epsilon^2(t)$ in this measurement for large T depends only on the behavior of $\gamma_x(\nu)$ as $\nu \to 0$. In particular, if $\gamma_x(0) \neq 0$, the error is proportional to T^{-1}.

If we know the fourth-order moment of $A(t)$, it is possible to calculate $\gamma_x(0)$. This moment can generally[8] be decomposed into two terms: the first one coincides with that obtained for Gaussian SP; the second one, called "cumulant," is due to the non-Gaussian character of the process. We show in Appendix B with the help of fairly general assumptions concerning this cumulant that the behavior of $\epsilon^2(T)$ for large values of T is the same as in the Gaussian case. Thus for this kind of problem, the Gaussian assumption is unnecessary.

From a practical point of view it is not always convenient to perform time averagings as indicated in Eq. (1). Indeed, the process $X(t)$ may become nonstationary during the long time intervals needed to reduce the error. It is then necessary to use more sophisticated methods, such as those used in the study of oscillators.[4]

Finally we can indicate that there is a connection between the behavior of the power spectrum as in Eq. (6) or (19) and the structure of the correlation function for large values of τ. In computer experiments it is often this point of view which is used, and we discuss briefly this point in Appendix C.

APPENDIX A

We study the behavior for $\nu \to 0$ of $\gamma_x(\nu)$ [defined by Eq. (16)], with $\gamma_A(\nu)$ given by Eq. (19). Because $\gamma_A(\nu) = \gamma_A(-\nu)$, the convolution appearing in Eq. (16) can be written

$$g(\nu) = \int_{-\infty}^{+\infty} \gamma_A(n)\gamma_A(n-\nu)\,dn\,. \tag{A1}$$

We will study this function for $|\nu| < a$. For this purpose, let us suppose $\nu > 0$ and introduce the intervals I and I' defined by

$$\begin{aligned} &n \in I \leftrightarrow \nu + \nu_i - a \le n \le \nu_i + a\,, \\ &n \in I' \leftrightarrow n \notin I\,. \end{aligned} \tag{A2}$$

Evidently $g(\nu) = g_I(\nu) + g_{I'}(\nu)$, where these functions are restrictions of the integral (A1) to the intervals I and I', respectively.

Consider first the integral $g_{I'}(\nu)$. If $n \in I'$, $\gamma_A(n) \times \gamma_A(n-\nu)$ is bounded, because there is only one frequency where $\gamma_A(\nu)$ is unbounded. This means that $\gamma_A(n)\gamma_A(n-\nu)$ becomes infinite only for $n = \nu_i$ and $n = \nu + \nu_i$, which are outside I'. Moreover, for large values of n, $\gamma_A(n) \simeq \gamma_A(n-\nu)$, because $|\nu| < a$. But $\gamma_A(n)$ is integrable from $-\infty$ to $+\infty$, because $A(t)$ is a second-order SP. Thus $\gamma_A^2(n)$ is integrable for high frequencies and $g_{I'}(\nu)$ is bounded.

Consider next $g_I(\nu)$, which we write, using Eq. (19),

$$g_I(\nu) = \int_{\nu+\nu_i-a}^{\nu_i+a} |n-\nu_i|^{m_i} |n-\nu_i-\nu|^{m_i} \times f_i(n) f_i(n-\nu)\,dn\,. \tag{A3}$$

By taking $n - \nu_i = \nu f$, we obtain

$$g_I(\nu) = \nu^{2m_i+1} \int_{(\nu-a)/\nu}^{a/\nu} |f|^{m_i} |f-1|^{m_i} k_i(f)\,df\,, \tag{A4}$$

where $k_i(f) < M$. The integral I appearing in Eq. (A4) is bounded. Indeed, we have

$$I \le M \int_{-\infty}^{+\infty} |f|^{m_i} |f-1|^{m_i}\,df \tag{A5}$$

because this integral exists. Indeed for $|f| \to \infty$, the integrand is of the order of $|f|^{2m_i}$, which is integrable according to Eq. (20). For $f \to 0$ or 1, the integral is regular because $m_i > -1$. The same calculation can be performed for $\nu < 0$. The final result is Eq. (21) of the main text.

APPENDIX B

We calculate the power spectrum $\gamma_x(\nu)$ of $X(t)$, which we define as the square of a non-Gaussian dynamical quantity $A(t)$. Because $\gamma_x(\nu)$ is the Fourier transform of the correlation function of $A^2(t)$, we can write

$$\gamma_x(\nu) = \int \langle A(t+\tau)A(t+\tau)A(t)A(t)\rangle e^{-2\pi i\nu\tau}\,d\tau\,. \tag{B1}$$

We introduce the Fourier transform $\gamma_A(\vec{\nu})$ of the fourth-order moment

$$\Gamma_A(\vec{t}) = \langle A(t_1)A(t_2)A(t_3)A(t_4)\rangle\,. \tag{B2}$$

Because $A(t)$ is stationary, $\gamma_A(\vec{\nu}) = 0$ outside a stationary manifold S,[9] defined by $\nu_1 + \nu_2 + \nu_3 + \nu_4 = 0$. It can be written

$$\gamma_A(\vec{\nu}) = \gamma_G(\vec{\nu}) + \gamma_c(\vec{\nu})\,. \tag{B3}$$

The first term in (B3) is the Gaussian term. It is the only one if $A(t)$ is Gaussian. Its contribution to the error $\epsilon^2(T)$ has already been calculated.

Thus we shall only consider the contribution of the second term in (B3), which can be written

$$\gamma_c(\vec{\nu}) = \gamma_c(\nu_1, \nu_2, \nu_3,)\,\delta(\nu_1 + \nu_2 + \nu_3 + \nu_4)\,. \tag{B4}$$

By Fourier transformation we calculate the contribution of this term in $\gamma_x(\nu)$. Taking this expression into Eq. (B1) we find

$$\gamma_{xc}(\nu)=\int\int\gamma_c(\nu-\nu_2,\nu_2,\nu_3)\,d\nu_2\,d\nu_3\ . \tag{B5}$$

In general, we assume that the integral of $|\gamma(\nu_1, \nu_2, \nu_3)|$ is bounded. This assumption secures the boundedness of $\gamma_{xc}(0)$. The contribution of the non-Gaussian term to the error is then of the order of T^{-1}. If the error decreases more slowly than T^{-1}, this is due to the Gaussian term.

APPENDIX C

The calculation of Refs. 1 and 2 were presented in the context of computer experiments on statistical mechanics. The aim of this comment is to explain that the problem is very general and appears in all the experiments where ensemble averages are measured by using time averages and assuming some ergodic properties of the process.

Some examples of spectral densities such as in Eq. (6) were presented in the study of statistical models of contact noise in semiconductors.[10]

In the case of computer experiments in molecular dynamics, it seems that the only well-known exception of a bounded spectral density was given by Alder and Wainwright.[11] It was particularly found that the correlation function of hard disks in two dimensions decays asymptotically as τ^{-1}, which is equivalent to a logarithmic singularity in the spectral density. This example is interesting because it cannot be described by Eq. (6), and we present shortly the result corresponding in this case to Eq. (9).

Let us suppose that

$$\begin{aligned}\gamma(\nu)&=-\ln\alpha\,|\nu|,\quad |\nu|<\nu_0<1/\alpha\\ &\le m,\quad |\nu|\ge\nu_0\,.\end{aligned} \tag{C1}$$

The error squared is given by Eq. (15), which becomes a sum of two integrals from 0 to ν_0 and from ν_0 to ∞. The second integral gives a contribution as in Eq. (10) which decreases as T^{-2}.

The first one can be written

$$I=\int_0^{\nu_0}\gamma_x(\nu)\left(\frac{\sin\pi\nu T}{\pi\nu T}\right)^2 d\nu\,. \tag{C2}$$

After simple calculation, we obtain that, for large values of T,

$$I\simeq C(\ln T)/T\,, \tag{C3}$$

which converges to zero, but more slowly than $\epsilon^2(T)$ of Eq. (11). Thus a logarithmic singularity in the spectrum leads to a slow convergence of the error to zero.

[1] R. Zwanzig and N. K. Ailawadi, Phys. Rev. 182, 280 (1969).
[2] N. K. Ailawadi, Phys. Rev. A 5, 1968 (1972).
[3] B. Picinbono, *Introduction à l'Étude des Signaux et Phénomènes Aléatoires* (Dunod, Paris, 1971), p. 53.
[4] E. Boileau and B. Picinbono, IEEE Trans. Instrum. Meas. 25, 66 (1976).
[5] B. Picinbono, Phys. Rev. A 4, 2398 (1971).
[6] B. Picinbono, Adv. Appl. Prob. 6, 512 (1974).
[7] B. Picinbono, IEEE Trans. Inf. Theory IT-16, 77 (1970).
[8] T. W. Anderson, *The Statistical Analysis of Times Series* (Wiley, New York, 1958), p. 438.
[9] M. Rousseau, J. Opt. Soc. Am. 61, 1307 (1971).
[10] G. G. Mac Farlane, Proc. Phys. Soc. 59, 366 (1947).
[11] B. J. Alder and T. E. Wainwright, Phys. Rev. A 1, 18 (1970).