STATISTICAL STUDY OF THE DELAY VARIANCE ESTIMATION FOR THE INDIVIDUAL AND GLOBAL METHODS

O. Meste, E. Bataillou, H. Rix
Laboratoire I3S-CNRS URA 1376
Bat.4, Les Lucioles, 250 Avenue Albert Einstein
Sophia Antipolis, 06560 Valbonne, FRANCE
e-mail: meste@essi.fr

ABSTRACT

When the variance of the delays is assumed to be relevant in a series of recurrent signals, two approaches are encountered. Either each delay is estimated allowing the computation of the sample variance (individual method) or the expected variance is directly estimated (global method). These two approaches are statistically compared using the global method introduced in a previous work and two individual methods: a Averaged Square Difference Function based estimator and a linear system based one. We finally show that the global method exhibits an interesting behaviour mainly due to its unbiasness.

1 INTRODUCTION

In this communication, we assume that a large number N of realizations of a recurrent signal $s(t)$ are observed with a random delay, in a noisy environment. To take examples in biomedical field, the deterministic and transient signal $s(t)$ may be an evoked potential, an myoelectric signal, an electrocardiogram. The random behaviour of these realizations is assumed to be due to the random delay and the additive noise. The study of another classical time-transform as the scaling instead of the delay has been achieved in previous works [3,2].
Each realization $x_i(t)$ is modelized using the following expression :

$$x_i(t) = s(t-d_i)+n_i(t) \qquad (i = 1 \dots N; 0 \leq t \leq P) \quad (1)$$

The random delays d_i are assumed to be small and to be realizations of zero mean identically distributed random variable (iidrv). The noise $n_i(t)$ is a zero mean white gaussian process.
In some circumstances, the knowledge on the statistical description of the delay is more important than the delay itself. We are interested in the second order statistics because it can be a relevant quantity in the case of physiological delays.

Figure (1) shows two different ways in estimating the delay variance. The first and obvious estimator is ${}_1\hat{\sigma}_d^2$ which needs all the estimated d_i ($\hat{d}_i$) to compute the

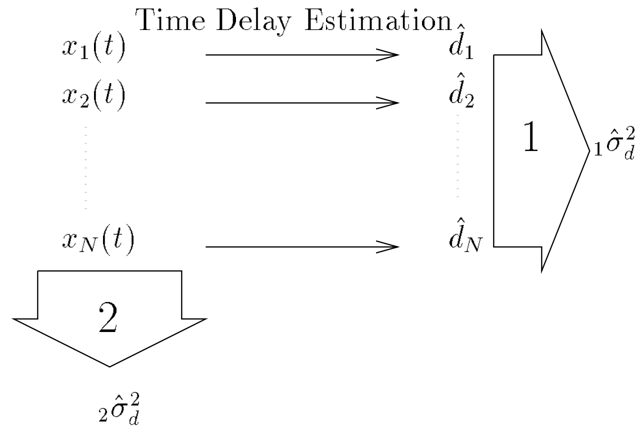

Figure 1: *two different ways for the σ_d^2 estimation*

classical sample variance (individual method). The second estimator ${}_2\hat{\sigma}_d^2$ is based on a previous work [4] where we showed that the delay variance can be estimated directly from the set of the N realizations (global method) under some hypothesis on the delays and the spectrum of signal $s(t)$. In [2] we have shown that a minimum variance estimation can be achieved even in the presence of colored noise. The global method was initialy developped to take into account the loss of a major hypothesis : for a given realization the delay is constant. An example of a variable delay is given in [4].
When this constancy is assumed, it is straightforward to compare the two approaches. So, we propose to characterize these two estimators (${}_1\hat{\sigma}_d^2$, ${}_2\hat{\sigma}_d^2$) studying their mean and variance. We will compare them using the same approach based on a series expansion which will be valid when delays are small enough.
In order to estimate each delay with the best accuracy, we will use the Average Square Difference Function (ASDF) because of its advantageous formulation [1]. We will show that the ASDF with parabolic interpolation [1] can be replaced by a close form using an explicit Least Square solution of a linear system when the delay is small enough.

2 THEORETICAL DEVELOPMENT

When the model of the delayed signal $x_i(t)$ is given by (1), the assumption of the sufficient smallness of the delay leads to use the following second order approximation of the signal variance [4]:

$$\hat{\sigma}_x^2(t) \approx s^{(1)^2}(t)\hat{\sigma}_d^2 - s^{(2)}(t)s^{(1)}(t)\overline{d^3} + \frac{1}{4}s^{(2)^2}(t)(\overline{d^4} - \overline{d^2}^2) + \hat{\sigma}_n^2 \quad (2)$$

where the symbols $\bar{}$ and $\hat{}$ correspond to the sample mean and the sample variance, respectively and $s^{(p)}(t)$ the p^{th} derivative of $s(t)$.

In a sampled time formulation (period equal to T), the previous equation becomes :

$$\hat{\boldsymbol{\sigma}}_x^2 \approx \mathbf{A}\boldsymbol{\theta} + \mathbf{b} \quad (3)$$

with

$$\mathbf{A} = \left(\mathbf{s}^{(1)} \cdot \mathbf{s}^{(1)} \mid -\mathbf{s}^{(2)} \cdot \mathbf{s}^{(1)} \mid \frac{1}{4}\mathbf{s}^{(2)} \cdot \mathbf{s}^{(2)} \mid \mathbb{1} \right) \quad (4)$$

and

$$\boldsymbol{\theta} = \begin{pmatrix} \hat{\sigma}_d^2 & \alpha & \beta & \hat{\sigma}_n^2 \end{pmatrix}^T \quad (5)$$

Where $\boldsymbol{\sigma}_x^2$ is the sample variance of $\mathbf{x}_i$, $\mathbb{1}$ the unit vector, $\hat{\sigma}_n^2$ the sample noise variance and $\mathbf{b}$ the random error due to the incertitude of the noise variance estimator. The symbol "$\cdot$" means the component wise product. The values α and β correspond to those in (2).
The sample variance being asymptotically unbiased and the number N large, the random vector $\mathbf{b}$ is zero mean so we can assume that the least square estimation ($\hat{\boldsymbol{\theta}}$) of $\boldsymbol{\theta}$ is unbiased. Let's now evaluate the covariance of $\boldsymbol{\theta}$:

$$cov(\hat{\boldsymbol{\theta}}) = (A^T A)^{-1} A^T E[\mathbf{b}\mathbf{b}^T] A (A^T A)^{-1} \quad (6)$$

The whiteness of the noise n_i leads to simplify $E[\mathbf{b}\mathbf{b}^T]$ by $2\frac{N-1}{N^2}\sigma_n^4\mathbf{I}$ [2] so (6) becomes :

$$cov(\hat{\boldsymbol{\theta}}) = 2\frac{N-1}{N^2}\sigma_n^4(\mathbf{A}^T\mathbf{A})^{-1} \approx \frac{2}{N}\sigma_n^4(\mathbf{A}^T\mathbf{A})^{-1} \quad (7)$$

The first component of $\hat{\boldsymbol{\theta}}$ will be noted ${}_2\hat{\sigma}_d^2$ and will be an estimate of $\hat{\sigma}_d^2$ (sample variance). From (4) its variance is defined by :

$$var({}_2\hat{\sigma}_d^2) \approx \frac{2}{N}\sigma_n^4 K_2 \quad (8)$$

The coefficient K_2 is a non linear function of the successive derivatives of $\mathbf{s}$. Some numerical values will be given by simulation but not its analytic expression.

We have shown that ${}_2\hat{\sigma}_d^2$ can be considered as an unbiased estimator of $\hat{\sigma}_d^2$ whose variance is given in (8). The next step will be to show that the ASDF with a parabolic interpolation [1] has a close form using a linear system. First of all, let's recall this time delay estimator ($\hat{d}_{i,A}$) of d_i:

$$\hat{d}_{i,A} = arg\min_{\tau} \hat{R}_{i,A}(\tau) \quad (9)$$

$$\begin{aligned} \hat{R}_{i,A}(\tau) &= \frac{1}{N}\sum_{k=1}^{N}[x_i(kT) - s(kT+\tau)]^2 \\ &= \frac{1}{N} < \mathbf{x}_i - \mathbf{s}_\tau, \mathbf{x}_i - \mathbf{s}_\tau > \\ &= \frac{1}{N}(< \mathbf{x}_i, \mathbf{x}_i > -2 < \mathbf{x}_i, \mathbf{s}_\tau > \\ &\quad + < \mathbf{s}_\tau, \mathbf{s}_\tau >) \end{aligned} \quad (10)$$

with $\mathbf{s}_\tau$ the vector corresponding to the sampled signal $s(t+\tau)$.
The parabolic interpolation of $\hat{R}_{i,A}(\tau)$ around its apex leads to the two typical steps:

- Locate the index m of the maximum cross-corelation lag $\hat{R}_{i,A}(mT)$ (coarse estimate)
- Evaluate the time delay $\hat{d}_{i,A}$ using the following equation (fine estimate)

$$\hat{d}_{i,A} = -\frac{T}{2}\frac{\hat{R}_{i,A}(mT+T) - \hat{R}_{i,A}(mT-T)}{\hat{R}_{i,A}(mT+T) - 2\hat{R}_{i,A}(mT) + \hat{R}_{i,A}(mT-T)} + mT \quad (11)$$

One can show that using the definition of $\hat{R}_{i,A}(\tau)$ (10), the previous equation can be rewriten using some scalar product operations:

$$\hat{d}_{i,A} = \frac{T}{2}\frac{< \mathbf{x}_i, \mathbf{s}_{mT-T} > - < \mathbf{x}_i, \mathbf{s}_{mT+T} >}{< \mathbf{x}_i, \mathbf{s}_{mT+T} > -2 < \mathbf{x}_i, \mathbf{s}_{mT} > + < \mathbf{x}_i, \mathbf{s}_{mT-T} >} + mT \quad (12)$$

The conditions needed to use a series expansion in order to give the expression of $\hat{\sigma}_x^2(t)$ (2), are easily fullfiled when the delay corresponds to $mT \pm T$. So, the delayed vectors $\mathbf{s}_{mT-T}$ and $\mathbf{s}_{mT+T}$ can be approximated using a second order series expansion:

$$\begin{aligned} \mathbf{s}_{mT-T} &\approx \mathbf{s} + (mT-T)\mathbf{s}^{(1)} + \frac{(mT-T)^2}{2}\mathbf{s}^{(2)} \\ \mathbf{s}_{mT+T} &\approx \mathbf{s} + (mT+T)\mathbf{s}^{(1)} + \frac{(mT+T)^2}{2}\mathbf{s}^{(2)} \end{aligned}$$

When we introduce these two approximation in (12) and after some calculus, we finally obtain:

$$\hat{d}_{i,A} = -\frac{< \mathbf{x}_i, \mathbf{s}^{(1)} >}{< \mathbf{x}_i, \mathbf{s}^{(2)} >} \quad (13)$$

Then using a parabolic interpolation and a second order series expansion leads to the ratio of two scalar products (13).

We are going to show that this result is almost the same than those obtained with a second order series expansion of the delayed signal $s(t-d_i)$.
Equation (1) is replaced by:

$$x_i(t) \approx s(t) - d_i s^{(1)}(t) + \frac{1}{2} d_i^2 s^{(2)}(t) + n_i(t) \qquad (14)$$

and expressed in a vector form (after sampling):

$$\mathbf{x}_i \approx \mathbf{B}\boldsymbol{\theta}_i + \mathbf{n}_i \qquad (15)$$

with

$$\mathbf{B} = \left(\mathbf{s} \mid \mathbf{s}^{(1)} \mid \mathbf{s}^{(2)}\right) \qquad (16)$$

and

$$\boldsymbol{\theta}_i = \begin{pmatrix} 1 & -d_i & d_i^2/2 \end{pmatrix}^T \qquad (17)$$

The least-squares solution of the system (15) is :

$$\hat{\boldsymbol{\theta}}_i = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{x}_i \qquad (18)$$

Using the definition of $\mathbf{B}$ (16) and the assumption that $\mathbf{s}$ vanishes at the limits of the observation interval, the product $\mathbf{B}^T\mathbf{B}$ is equal to :

$$\mathbf{B}^T\mathbf{B} = \begin{pmatrix} <\mathbf{s},\mathbf{s}> & 0 & <\mathbf{s},\mathbf{s}^{(2)}> \\ 0 & <\mathbf{s}^{(1)},\mathbf{s}^{(1)}> & 0 \\ <\mathbf{s},\mathbf{s}^{(2)}> & 0 & <\mathbf{s}^{(2)},\mathbf{s}^{(2)}> \end{pmatrix} \qquad (19)$$

One property of such a matrix is that its inverse has the form:

$$(\mathbf{B}^T\mathbf{B})^{-1} = \begin{pmatrix} \gamma_1 & 0 & \gamma_2 \\ 0 & \frac{1}{<\mathbf{s}^{(1)},\mathbf{s}^{(1)}>} & 0 \\ \gamma_2 & 0 & \gamma_3 \end{pmatrix} \qquad (20)$$

Then the product $(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$ is :

$$(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T = \begin{pmatrix} \gamma_1\mathbf{s}^T + \gamma_2\mathbf{s}^{(2)T} \\ \frac{1}{<\mathbf{s}^{(1)},\mathbf{s}^{(1)}>}\mathbf{s}^{(1)T} \\ \gamma_2\mathbf{s}_T + \gamma_3\mathbf{s}^{(2)T} \end{pmatrix} \qquad (21)$$

The expression of γ_i will not be given but can be easily evaluated. From (17) , (18) and (21) we can deduce the least-squares estimation $(\hat{d}_i)$ of d_i :

$$\hat{d}_i = -\frac{<\mathbf{x}_i,\mathbf{s}^{(1)}>}{<\mathbf{s}^{(1)},\mathbf{s}^{(1)}>} \qquad (22)$$

Using the properties of n_i, the unbiasness of the least-squares estimate $(\hat{\boldsymbol{\theta}}_i)$ is clearly stated and its covariance is :

$$cov(\hat{\boldsymbol{\theta}}_i) = \sigma_n^2(\mathbf{B}^T\mathbf{B})^{-1} \qquad (23)$$

So, using (20) the variance of $\hat{d}_i$ is :

$$var(\hat{d}_i) = \sigma_n^2\frac{1}{<\mathbf{s}^{(1)},\mathbf{s}^{(1)}>} = \sigma_n^2 K_1 \qquad (24)$$

At this stage we can compare the estimation of d_i following the two approaches, i.e the ASDF $(\hat{d}_{i,A})$ and the solution of the previous linear system $(\hat{d}_i)$. There is a direct equality when we use (14) and the following inequality:

$$<\mathbf{s}^{(1)},\mathbf{s}^{(1)}> \gg \frac{d_i^2}{2} <\mathbf{s}^{(2)},\mathbf{s}^{(2)}> + <\mathbf{n}_i,\mathbf{s}^{(2)}> \qquad (25)$$

In the following, we will use $\hat{d}_i$ as an estimate of d_i because its variance is easily expressed and the assumptions used in its statement are the same than those for the global method. Contrarily to the ASDF estimate, we will not assume that $\mathbf{s}$ is a strictly band limited random signal.
Let's assume that $\hat{d}_i$ is used to estimate each delay d_i $(i = 1\ldots N)$, and let's call ${}_1\hat{\sigma}_d^2$ the sample variance calculated with the $\hat{d}_i$ previously calculated. The sample variance of the true delays d_i will be called $\hat{\sigma}_d^2$.
Let's recall that ${}_2\hat{\sigma}_d^2$ is the estimation of $\hat{\sigma}_d^2$ using the global method (2) and that it is characterized by its unbiaseness and its variance (8). In (5), we can see that ${}_2\hat{\sigma}_d^2$ is the estimate of $\hat{\sigma}_d^2$ (sample variance) but not σ_d^2 (theoretical variance). So, to characterize ${}_1\hat{\sigma}_d^2$ we must substract $\hat{\sigma}_d^2$. We are going to characterize the mean and the variance of ${}_1\hat{\sigma}_d^2 - \hat{\sigma}_d^2$
Using the definition of ${}_1\hat{\sigma}_d^2$, $\hat{\sigma}_d^2$ and $\hat{d}_i$:

$${}_1\hat{\sigma}_d^2 = \frac{1}{N}\sum_{i=1}^{N}\hat{d}_i^2 - (\frac{1}{N}\sum_{i=1}^{N}\hat{d}_i)^2 \qquad (26)$$

$$\hat{\sigma}_d^2 = \frac{1}{N}\sum_{i=1}^{N}d_i^2 - (\frac{1}{N}\sum_{i=1}^{N}d_i)^2 \qquad (27)$$

and

$$\hat{d}_i = d_i + e_i \qquad (28)$$

where e_i is the error in the estimation of d_i.
The mathematical expectation of ${}_1\hat{\sigma}_d^2 - \hat{\sigma}_d^2$ can be evaluated by :

$$E[{}_1\hat{\sigma}_d^2 - \hat{\sigma}_d^2] = \frac{N-1}{N}\sigma_e^2 \qquad (29)$$

The variance of the error e_i has been previously evaluated (24) then (29) has its final form :

$$E[{}_1\hat{\sigma}_d^2 - \hat{\sigma}_d^2] = \frac{N-1}{N}\sigma_n^2 K_1 \approx \sigma_n^2 K_1 \qquad (30)$$

The variance of ${}_1\hat{\sigma}_d^2 - \hat{\sigma}_d^2$ has the form :

$$var[{}_1\hat{\sigma}_d^2 - \hat{\sigma}_d^2] = E[({}_1\hat{\sigma}_d^2 - \hat{\sigma}_d^2)^2] - E^2[{}_1\hat{\sigma}_d^2 - \hat{\sigma}_d^2] \qquad (31)$$

After some development, we get:

$$var[{}_1\hat{\sigma}_d^2 - \hat{\sigma}_d^2] = \frac{2(N-1)}{N^2}\sigma_e^4 + \frac{4(N-1)}{N^2}\sigma_d^2\sigma_e^2 \qquad (32)$$

and using (24), we finally get :

$$var[{}_1\hat{\sigma}_d^2 - \hat{\sigma}_d^2] = \frac{N-1}{N^2}(2\sigma_n^4 K_1^2 + 4\sigma_d^2\sigma_n^2 K_1) \approx \frac{1}{N}(2\sigma_n^4 K_1^2 + 4\sigma_d^2\sigma_n^2 K_1) \quad (33)$$

So, we have shown that when the noise variance cannot be estimated, the first estimator $({}_1\hat{\sigma}_d^2)$ is biased but not the second one $({}_2\hat{\sigma}_d^2)$.

3 SIMULATIONS

In order to correctly compare the two estimators, we must take into account their variance (33) and (8) for a fixed value of K_1 (24) and K_2 (8). As the expression of K_2 is not defined we will use a simulation based on N (equal to 1000) gaussian shape and randomly delayed signals with a gaussian jitter whose theoretical variance is equal to $\sigma_d^2 = 0.25$.

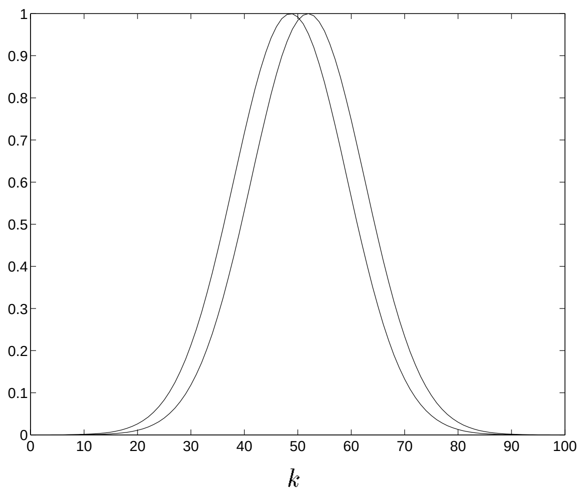

Figure 2: *two realizations of $s(t-d_i)$ corresponding to the min and max values of d_i*

Figure (2) shows the minimum and maximum delayed signal. For signals in fig. (2) , the coefficients K_1 and K_2 are: $K_1 = 12.02$, $K_2 = 7967$. Using these coefficients we calculate in fig. (3) the theoretical Mean Square Error (MSE) for different values of σ_n^2. In fig. (3), the simulation results, indicated by the dotted line, have been obtained from 500 independent runs.
We can notice a good agreement between the theoretical results and the simulation ones for low SNR (e.g. $\sigma_n^2 = 0.1$ corresponds to $SNR = 5dB$). The departure in results for high SNR (small σ_n^2 values) is due to the approximation errors in (2) which become preponderant when the noise effect is neglectable. These errors are mainly due to the approximation $E[nv] = E[n]E[v]$ for two independent random variables when the sample mean is used instead of the mathematical expectation.

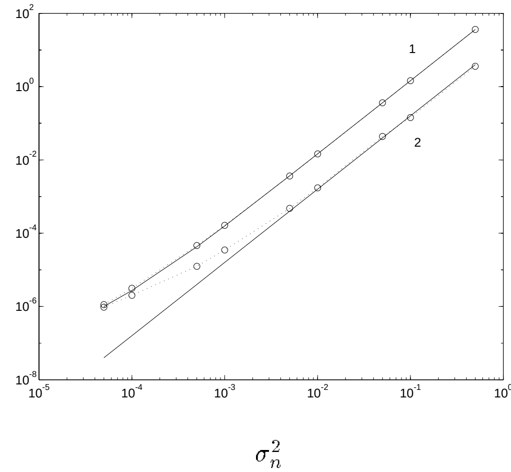

Figure 3: *Mean Square Error in function of σ_n^2. group 1: ${}_1\hat{\sigma}_d^2$; group 2: ${}_2\hat{\sigma}_d^2$. For each group, the solid line represents the theoritical value and the dotted line the simulation results*

4 CONCLUSION

The second approach of σ_d^2 estimation, called ${}_2\hat{\sigma}_d^2$ (global method), was initially proposed in a particular case where the jitter may vary in function of time leading to $\sigma_d^2(t)$. We have shown that when σ_d^2 is independent of time our approach can be advantageously compared to classical one. In a first time, the considered reference has been the ASDF and in a second time a simple linear system close to an expansion of the ASDF. Theoretical and simulation results have been given for a statistical comparison. The major drawback, highlighted by the comparison, in the ${}_1\hat{\sigma}_d^2$ estimator (individual method) is the presence of bias. In some circumstances, the knowledge of the noise statistics could be used to reduce it. Nevertheless, the assumptions encountered in the second estimator $({}_2\hat{\sigma}_d^2)$ statement reduce its application field.

References

[1] G. Jacovitti, G. Scarano, "Discrete time techniques for time delay estimation", IEEE Trans. Sig. Proc., Vol. 41, No.2, pp.525-533, 1993.

[2] O. Meste, E. Bataillou, H. Rix, "Comparaison de deux estimateurs de la variance relative à des fluctuations d'échelle de transitoires, en présence de bruit coloré", 15ieme Colloque GRETSI,Juan-Les-Pins, pp. 205-208, 1995.

[3] O. Meste, E. Bataillou, H. Rix, "Scaling factor and jitter characterization: an unified approach using mean and variance", Proc. 17th Annual Int. Conf. IEEE-EMBS, Montreal, CD-ROM, 1995.

[4] O. Meste, H. Rix, "Jitter statistics estimation in alignment processes", to be published *Signal Processing*, Vol. 51, No. 1, 1996.