Cross-Spectrum versus Spectrum Average

F. Vernotte, A. Baudiquez, E. Rubiola
FEMTO-ST, Time and Frequency Department
Observatory THETA, UBFC
Besançon, France
francois.vernotte@femto-st.fr

E. Lantz
FEMTO-ST, Département d'Optique P.M. Duffieux
UBFC
Besançon, France

***Abstract*—This paper compares the efficiencies of the cross-spectrum and of a simple spectrum average to estimate the frequency stability of an oscillator.**

***Keywords*–frequency stability; cross-spectrum; Bayesian statistics; inverse problem**

I. Introduction

The main purpose of this paper is the comparison of the efficiencies of the cross-spectrum (c-s) estimator, i.e. the covariance of 2 spectra, regarding a simple average of these spectra. We assume that each spectrum is composed of a common red noise, that we call signal, and a white noise. The white noises of both spectra have the same level σ_N^2 but are uncorrelated. Such a comparison imposes the perfect knowledge of this parameter σ_N^2 in order to be able to evaluate the signal variance σ_S^2 with the spectrum average (s.a) method. However, this latter estimator exhibits a lower estimate variance than the other one suggesting that it could be more efficient even if it is biased in essence. Ultimately, this paper addresses the question about the efficiency of the Dike radiometer versus that of the correlation radiometer [1], [2], [3]. In fact, one bin of the FFT is equivalent to the IF filter of the receiver. The question is then: which estimator is the most efficient?

II. Variance of the Estimates

At a given Fourier frequency f, the spectra may be modeled as

$$\left\{ \begin{array}{l} X = A + C \\ Y = B + C \end{array} \right. \tag{1}$$

where A, B are uncorrelated normal complex random variables (rv) of variance σ_N^2, i.e. the white level, and C a normal complex rv of variance σ_S^2/f^α, with σ_S^2 the signal level and α the red noise exponent.

The estimators are then

$$\left\{ \begin{array}{l} \hat{S}_{cs} = X \cdot \tilde{Y} \\ \hat{S}_{av} = \Re\left[\dfrac{X+Y}{2}\right]^2 + \Im\left[\dfrac{X+Y}{2}\right]^2 - \dfrac{\sigma_N^2}{2} \end{array} \right.$$

where $\tilde{}$ stands for the complex conjugate of the quantity which is below and $\Re[\cdot], \Im[\cdot]$ stands respectively for the real and the imaginary parts of the quantities within the brackets. One can easily verify that these estimators are unbiased[1] [4].

[1] We stated in the introduction that the s.a estimator is "biased in essence" because, in order to get an unbiased estimator of the signal level, we must subtract half the noise level from it, i.e. its bias.

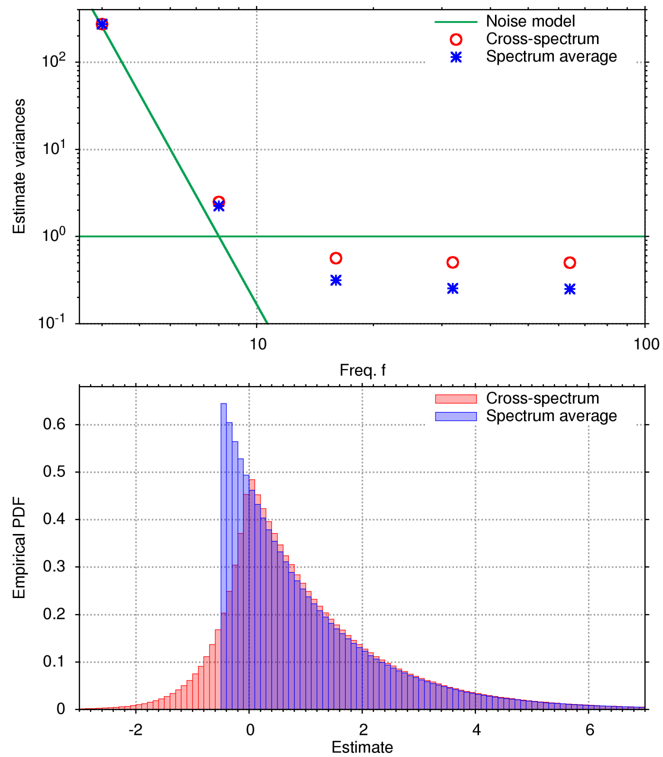

Fig. 1. Variance (above) and histogram (below) of the estimates. These data were obtained from a set of 1 million simulated spectra. Below is the histogram of the estimates at $f = 8$ a.u where $\sigma_S^2 = \sigma_N^2 = 1$ a.u.

On the other hand, denoting $\mathbb{V}[\cdot]$ the variance of the quantity within the brackets, we can demonstrate that

$$\left\{ \begin{array}{lll} \mathbb{V}\left[\hat{S}_{cs}\right] \approx \mathbb{V}\left[\hat{S}_{av}\right] & \text{for} & \sigma_S^2 \gg \sigma_N^2 \\ \mathbb{V}\left[\hat{S}_{cs}\right] \approx 2\mathbb{V}\left[\hat{S}_{av}\right] & \text{for} & \sigma_S^2 \ll \sigma_N^2. \end{array} \right.$$

This is confirmed by Fig. 1 (top) which exhibits the variance of the estimates of both estimators applied to a signal composed of a mixture of uncorrelated white noise of level 1 arbitrary unit (a.u) and a common f^{-8} noise which crosses the white noise level at $f = 8$ a.u. At $f = 4$ a.u, the signal level is 256 times higher than the white level and the variances of both estimators coincide. On the other hand, for frequencies higher than 16 a.u, the signal level is less than 256 times lower than the white level and the variance of the c-s estimates is 2

TABLE I
COMPARISON OF THE 95% BOUNDS GIVEN BY THE C-S AND S.A ESTIMATORS FOR A NOISE LEVEL $\sigma_N^2 = 1$ A.U.

σ_S^2	Best %	Min %	Mean %	Max %
0.0	20	-86	17	28
0.1	24	-94	16	28
0.2	27	-140	16	28
0.3	28	-150	16	28
0.5	31	-146	16	28
0.7	32	-168	17	28
1.0	34	-210	14	28
1.4	36	-210	13	28
2.0	38	-169	13	28
3.2	35	-203	11	28
5.0	37	-187	9	28
10.0	40	-133	5	27

The second column ("Best") exhibits the percentage of cases where the 95% bound from the c-s estimator is more stringent (lower) than the 95% bound obtained by the s.a estimator. The following columns show respectively the minimum, the mean and the maximum values of the relative differences between the 95% bounds of both estimators: $(B95_{cs} - B95_{av})/B95_{cs}$.

times higher than the variance of the s.a estimates. This seems to indicate a better efficiency of the s.a estimator.

However, the histograms of the estimates of both estimators exhibit a very different shape (see Fig. 1 bottom): a variance-Γ double exponential for the c-s estimates [5] and the χ^2 decreasing exponential for the s.a estimates. Then, although the negative part are very different, the positive part of the histograms are quite similar. However, for low positive estimates, the c-s histogram is a little bit higher. Therefore, this shows the probability to get a positive estimate, i.e. an informative estimate, is very slightly higher with the c-s estimator than with the s.a estimator.

The only way to get an objective answer about the efficiency of the methods is to search for the Bayesian upper limit, e.g. at 95 % confidence, of the σ_S^2 estimation knowing one estimate: the most efficient method is the one which provides the most stringent upper limit.

III. SEARCHING FOR THE MOST STRINGENT ESTIMATOR

A. *Inverse problem*

We have have first to adress the direct problem, i.e. the statistics of the c-s (or s.a) estimates knowing the signal level (the noise level is assumed to be known). Then, we have to deduce the inverse problem from the direct problem, i.e. the statistics of the signal level from a c-s (or s.a) estimate. The direct and inverse problem resolutions will be detailed in the full paper.

B. *Comparison of the 95 % confidence limits*

An example of result of such a process is given in Table I: for different signal level values (the noise level is set to 1 a.u), this table gives the proportion of most stringent bound given by the c-s estimator. This percentage increases from 20% when there is no signal to 40% when the signal level is 10 times the noise level, i.e. when the number of negative estimates decreases. It is highly probable that the percentage tends toward 50% when $\sigma_S^2 \gg \sigma_N^2$, i.e. when the probability to get a negative $\hat{S}_{cs}$ estimate tends toward 0. On the other hand, Table I shows that the mean deviation between the bounds obtained by both estimator is as low as $10 \sim 15\%$ and that the maximum deviation, i.e. when the c-s bound is higher the s.a bound, does not exceed 28%. However, the minimum deviation can be of the order of -200% meaning that in some rare cases the s.a 95% bound can be 2 times higher than the c-s bound.

Nevertheless, these results definitely show that the a-v estimator is more efficient than the c-s estimator, even if the difference is small.

IV. CONCLUSION

In order to compare the efficiency of the cross-spectrum estimator and the spectrum average estimator to assess the signal level, we first calculated the variances of these estimators and observed that the variance of the latter estimator is lower than the variance of the former one, suggesting that the spectrum average estimator is the best. We decided then to compare the Bayesian limit at 95% of confidence of the signal level given by both estimators. Here also we found a slight advantage for the spectrum average estimator. This estimator is then the winner of this trial, at least for 2 degrees of freedom. An extension of this study for higher degrees of freedom will be given in the full paper.

Nevertheless, considering on one hand the very small differences between the efficiencies of these estimator and on the second hand the significant discrepancies between the 95% bounds that may occasionally appear, the wiser solution could be to systematically compute the bounds given by both estimators and choose the lower one.

Finally, we must also remind that the second estimator implies a perfect knowledge of the noise level since half of it has to be subtracted from the spectrum average. The least uncertainty about this noise level knowledge could drastically decrease the efficiency of the spectrum average method. In this connexion, the cross-spectrum estimator is definitely the most robust estimator since it does not need an explicit knowledge of the noise level.

ACKNOWLEDGEMENT

This work was partially funded by the ANR Programmes d'Investissement d'Avenir (PIA) Oscillator IMP (Project 11-EQPX-0033) and FIRST-TF (Project 10-LABX-0048).

REFERENCES

[1] R. H. Dicke, "The measurement of thermal radiation at microwave frequencies," *Rev. Sci. Instrum.*, vol. 17, no. 7, pp. 268–275, Jul. 1946.
[2] E. J. Blum, "Sensibilit des radiotelescopes et recepteurs correlation," *Ann. Astrpohysique*, vol. 22, no. 2, pp. 140–163, 1959.
[3] M. E. Tiuri, *Radio-Telescope Receivers,* Chap. 7 of J. D. Kraus, *Radio Astronomy*. New York: McGraw Hill, 1966.
[4] E. Rubiola and F. Vernotte, "The cross-spectrum experimental method," March 2010, arXiv:1003.0113v1.
[5] F. Vernotte and E. Lantz, "Three-cornered hat and Groslambert covariance: A first attempt to assess the uncertainty domains," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 66, no. 3, pp. 643–653, March 2019.