

Cross-Spectrum Measurement Statistics: uncertainties and detection limit

Antoine Baudiquez, Éric Lantz, Enrico Rubiola, François Vernotte

Abstract—The cross-spectrum method consists in measuring a signal $c(t)$ simultaneously with two independent instruments. Each of these instruments contributes to the global noise by its intrinsic (white) noise, whereas the signal $c(t)$ that we want to characterize could be a (red) noise.

We first define the real part of the cross-spectrum as a relevant estimator. Then, we characterize the probability density function (PDF) of this estimator knowing the noise level (direct problem) as a Variance-Gamma (VG) distribution. Next, we solve the “inverse problem” thanks to Bayes’ theorem to obtain an upper limit of the noise level knowing the estimate. Checked by massive Monte Carlo simulations, VG proves to be perfectly reliable for any number of degrees of freedom (dof).

Finally we compare this method with an other method using the Karhunen-Loève transform (KLT). We find an upper limit of the signal level slightly different as the one of VG since KLT better takes into account the available information.

Index Terms—Bayesian statistics, confidence interval, cross-spectrum, Karhunen-Loève transform, probability density function.

I. INTRODUCTION

THE measurement of power spectra is a classical problem, ubiquitous in numerous branches of physics, as explained below. Power spectra are efficiently measured using Fourier transform methods with digitized data. Relevant bibliography is now found in classic books [1], [2], [3], [4].

We are interested in the measurement of weak statistical phenomena, which challenge the instruments and the mathematical tools, using the cross-spectrum method. This method consists of the simultaneous measurement of the signal with two separate and independent instruments [5]. The other approach, consisting on the observation of the spectral contrast in a chopped signal, broadly equivalent to the Dicke radiometer [6], is not considered here. Regarding the duration of the data record used to evaluate the Fast Fourier Transform (FFT), two asymptotic cases arise.

The first case is that of the measurement of fast phenomena, where a large number of records denoted m is possible in a reasonable observation time. At large m the central limit theorem rules and the background noise can be rejected by

a factor approximately equal to $1/\sqrt{m}$, depending on the estimator. Numerous examples are found in the measurement of noise in semiconductors [7], phase noise in oscillators and components [8], [9], [10], [11], frequency fluctuations and relative intensity noise in lasers [12], [13], electromigration in thin films [14], etc. Restricting to one bin of the Fourier transform, the power spectral density integrated over a suitable frequency range is used in radiometry [15], [16], Johnson thermometry [17] and other applications.

The second case is that of slow phenomena, where the fluctuations are very long term or non ergodic. On one hand the background noise is still rejected as before but with a very low m which can actually be equal to one. On the other hand, the central limit theorem does not apply and the statistical uncertainties are not trivial. This case is of great interest in radio astronomy, where the observations are limited by the available resources and take a long time. As instance millisecond pulsars (MSP) can be used as very stable clocks at astronomical distances [18]. The radio pulses times of arrival (TOA) of MSP are affected by numerous physical process, one of them are gravitational-wave (GW) perturbations [19], [20]. Red noise originated from GW perturbations in the signal path common to the radio-telescopes can be detected [21], [22]. Like the analysis of the signals provided by the LIGO/VIRGO interferometers which use cross correlation methods [23], [24], the LEAP experiment (i.e. Large European Array for Pulsars) [25] could use such methods in order to access lower frequencies and observe imperceptible phenomena such as early phases well before the coalescence of black holes or GW of cosmological origin (for example cosmic strings, inflation, primordial black holes).

This article is intended to put an upper limit on the uncertainty of the cross-spectrum estimate. The method proposed here is totally general regardless the power law type of noise. Indeed, even if the pulsar signal would be constituted by white noise, the realisations of this white noise would be the same, for low frequency, at different observatories while the realisations of measurement white noise are independent. However this paper shows a particular interest to red noise. Gravitational-waves have not been yet discovered in the TOA of millisecond pulsars. Thanks to the very long line of sight between the pulsars and us (several thousand light-years), we could access very low frequencies, inaccessible to LIGO/VIRGO, thus revealing much slower astrophysical phenomena. It is therefore important to develop statistical tools to improve measurement sensitivity to pulsar timing observations [26]. In this respect, we propose in Section II to state the cross-spectrum problem to define a proper

A. Baudiquez and E. Rubiola are with FEMTO-ST, Department of Time and Frequency, UMR 6174, Université Bourgogne Franche-Comté, France. Antoine’s ORCID is 0000-0002-7007-5273

E. Lantz is with FEMTO-ST, Département d’Optique P.M. Duffieux, UMR 6174 CNRS, Université Bourgogne Franche-Comté, France.

E. Rubiola is also with the Division of Quantum Metrology and Nanotechnology, Istituto Nazionale di Ricerca Metrologica (INRiM), 10135 Turin, Italy. Enricos ORCID is 0000-0002-5364-1835

F. Vernotte is with FEMTO-ST, Department of Time and Frequency, Observatory THETA, UMR 6174 CNRS, Université Bourgogne Franche-Comté, France. François’s ORCID is 0000-0002-1645-5873

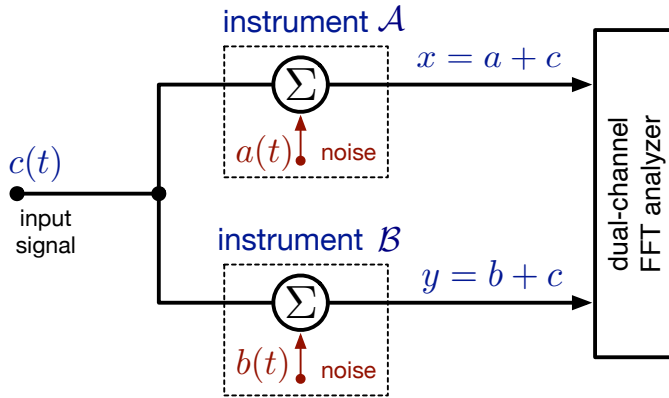


Fig. 1. Basics of the cross-spectrum method.

estimate. Based on the principle that the experiment is repeated m times, it is important to note that the estimation of the measurement uncertainty is A-Type as defined by the VIM [27]. Then in Section III we define the probability density function (i.e. ‘‘direct problem’’) of the cross-spectrum estimate which is used in Section IV to compute an upper limit by using a bayesian inference approach (i.e. ‘‘inverse problem’’). The results obtained are compared to an another method using the Karhunen-Loève transform developed in [28] and the conclusions are presented in Section V.

II. STATEMENT OF THE PROBLEM

A. Spectral Measurement

This paper aims to measure the mathematical expectation of the amplitude for a frequency bin. This amplitude obviously depends on the frequency and this is what we are looking for as shown in Fig. 2. Let us remind at that the Fourier transform of a white noise is a white noise. Fourier transform, random signals have generally infinite energy so the real Fourier transform can not be generally defined. A description of a realistic white noise actually corresponds to a Markov process of the first order and the reader can report to [29] for discrete simulation of colored noise and stochastic processes. The random part of the amplitude of each frequency bin is then uncorrelated from each bin. Considering a red noise as a filtered white noise means that the red noise spectrum is the product of the white spectrum by a deterministic function and, therefore, the uncorrelation property of the random part of the red noise spectrum is preserved. Working directly in the frequency domain therefore changes absolutely nothing in terms of a random variable. Consequently it does not matter to know if the signals are stationary, uncorrelated or not since we are interested in the spectrum and in a bin of frequency in particular.

B. Cross-spectrum Method

Let us consider 3 statistically independent signals: $a(t)$, $b(t)$ and $c(t)$ as shown in Fig. 1. On one side the two first $a(t)$ and $b(t)$ are respectively the instrument noise of \mathcal{A} and \mathcal{B} . On the other side $c(t)$ is an input signal which we want to characterize. This signal is stochastic and not necessarily

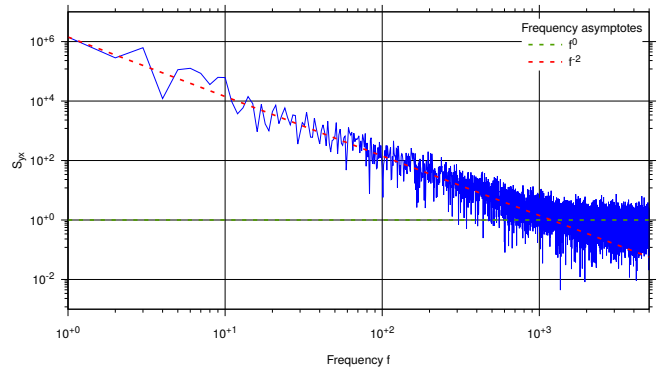


Fig. 2. This is a plot of a PSD estimate magnitude obtained via a discrete Fourier transform. The red and green dashed lines correspond to the power law frequency asymptotes related to the signal.

stationary. In the case of pulsar measurement, this input signal is generally a red noise. The output of each channel is

$$\begin{aligned} x(t) &= a(t) + c(t) \\ y(t) &= b(t) + c(t). \end{aligned} \quad (1)$$

Processing experimental signals, we can assume that white noise is a continuous function of time, and that the Fourier Transform (FT) always exists if we look at a very short interval regarding the sampling period. The reader can refer to [30] for a more detailed model of realistic white noise. Applying the FT on each channel gives

$$\begin{aligned} X(f) &= A(f) + C(f) \\ Y(f) &= B(f) + C(f) \end{aligned} \quad (2)$$

where f is the frequency, $X(f)$, $Y(f)$, $A(f)$, $B(f)$ and $C(f)$ stand respectively for the Fourier Transform of $x(t)$, $y(t)$, $a(t)$, $b(t)$ and $c(t)$. Our interest is carried out on the power spectral density (PSD) rather than the spectrum. The cross-spectrum is defined as

$$S_{yx}(f) = \frac{1}{T} \mathbb{E} [Y(f)X^*(f)], \quad (3)$$

where the cross-spectrum is actually a cross-PSD and $\mathbb{E}[\cdot]$ stands for the mathematical expectation of the quantity within the brackets. The factor T is the measurement time which is necessary for because the power calculated in time domain and in frequency domain must be the same (Parseval theorem), and also for $S_{yx}(f)$ to have the dimension of power per unit of frequency. The $*$ denotes the complex conjugate of the quantity placed before it. Eq. 3 is the two-sides PSD, which contains positive and negative frequencies. Experimentally averaging over m spectra realizations leads to the following cross-PSD estimator

$$\langle S_{yx} \rangle_m = \frac{1}{T} \langle Y(f)X^*(f) \rangle_m. \quad (4)$$

C. Cross Power Spectral Density

Averaging on a large number of observations, the mathematics is made simple by the central limit theorem, by which all the probability density functions (PDF) become Gaussian.

More interesting for us is the case of a small number of realizations, each of which taking long observation time-up to several years in the case of the millisecond pulsars.

The random variables (rv) $a(t)$, $b(t)$ and $c(t)$ follow a centered normal distribution whatever the kind of noise. Even red noise (e.g. random walk) follows a normal distribution not on the time average but regarding its ensemble average over the probability space which means it is a non ergodic process. Moreover a stochastic process with zero-mean Gaussian distribution has a FT which is also a random process with centered Gaussian distribution.

The rv $A(f)$, $B(f)$ and $C(f)$ can then be decomposed into a real and imaginary part

$$\begin{aligned} A(f) &= A'(f) + iA''(f) \\ B(f) &= B'(f) + iB''(f) \\ C(f) &= C'(f) + iC''(f). \end{aligned} \quad (5)$$

The real and imaginary part are statistically independent rv with equal variance following a zero-mean Gaussian distribution. For an ensemble average, or a time sequence sufficiently long to ensure a good spectral resolution, the samples at different frequencies are independent of each other. Hence, all the results of this paper are given for a given frequency that we do not mention explicitly. Of course, there is a deterministic relation between the results and this frequency, except in the case of a signal constituted by white noise, which is not the most common assumption in the envisioned applications. Let us now expand Eq. 4,

$$\begin{aligned} \langle S_{yx} \rangle_m &= \frac{1}{T} \langle YX^* \rangle_m \\ &= \frac{1}{T} [\langle A'B' + B'C' + C'A' + C'^2 \rangle_m \\ &\quad + \langle A''B'' + B''C'' + C''A'' + C''^2 \rangle_m \\ &\quad + i \langle A'B'' + B''C' + C''A' \rangle_m \\ &\quad - i \langle A''B' + B'C'' + C'A'' \rangle_m]. \end{aligned} \quad (6)$$

The terms in the imaginary part have a zero expectation, while the expectation in the real part is proportional to the PSD of the signal, i.e. what we are looking to characterize. As a consequence, the next sections focus solely on the real part $\Re\{S_{yx}\}$,

$$\langle \Re\{S_{yx}\} \rangle_m = \frac{1}{T} \langle (A^k + C^k)(B^k + C^k) \rangle_\nu \quad (7)$$

where $\nu = 2m$ the number of degree of freedom (dof). The superscript k means real or imaginary part because they are independent rv.

D. Statement of the Problem

1) *Measurements, and Estimates:* In the following, in order to simplify the notation, we will omit the superscript k . Thereby the real and imaginary part will be treated as 2 dof. Moreover to simplify the notations, we will omit the factor $\frac{1}{T}$ which does not affect the PDF. The **estimates** will be noted with a “*hat*” and we refer the cross-spectrum measurement for a given frequency to

$$\hat{Z} = (\hat{A} + \hat{C})(\hat{B} + \hat{C}) \quad (8)$$

where all \hat{A} , \hat{B} , \hat{C} are rv which are independent, centered and normal. In the following, we will assume that \hat{A} , \hat{B} , \hat{C} have only 1 dof, their real or their imaginary part, and that \hat{Z} does not come from the average of different spectra. A generalization of this problem to 2 dof (real and imaginary parts) and then $2m$ dof (average of m spectra) will be given.

2) *Direct and Inverse Problem:* In order to assess the uncertainty over the estimator σ_C^2 , called the *signal* level, we will have to distinguish to main issues:

- The **direct problem** consists in calculating the statistics of the cross-spectrum measurement \hat{Z} , knowing the model parameters σ_A^2 , σ_B^2 , σ_C^2 .
- The **inverse problem**, conversely consists in calculating a confidence interval over the unknown model parameter σ_C^2 , from the known parameters σ_A^2 , σ_B^2 and the cross-spectrum measurement \hat{Z} .

III. DIRECT PROBLEM

In Sections III-A to III-E we will omit the “*hat*” for estimates since we deal with the mathematical models.

A. Vector Formalization of the Problem

We will reuse here the formalism we developed in [31], i.e. a vector space of normal laws. Since we have 3 normal rv, we are in a vector space of 3 dimensions that we will denote $\mathbb{L}\mathbb{G}^3$ and which has the basis $(\vec{E}_A, \vec{E}_B, \vec{E}_C)$ defined as

$$\begin{cases} \vec{E}_A = \text{LG}_A(0, 1) \\ \vec{E}_B = \text{LG}_B(0, 1) \\ \vec{E}_C = \text{LG}_C(0, 1) \end{cases}$$

where $\text{LG}(0, 1)$ stands for a Laplace-Gauss (or normal) rv. with zero-mean (centered) and unity standard deviation ($\sigma = 1$). We assume that $\text{LG}_A(0, 1)$, $\text{LG}_B(0, 1)$, $\text{LG}_C(0, 1)$ are independent. We can define the scalar product between the basis vectors of $\mathbb{L}\mathbb{G}^3$ in such a way:

$$\begin{cases} \|\vec{E}_A\|^2 = \vec{E}_A \cdot \vec{E}_A = \text{LG}_A \cdot \text{LG}_A = \chi_A^2 \\ \|\vec{E}_B\|^2 = \vec{E}_B \cdot \vec{E}_B = \text{LG}_B \cdot \text{LG}_B = \chi_B^2 \\ \|\vec{E}_C\|^2 = \vec{E}_C \cdot \vec{E}_C = \text{LG}_C \cdot \text{LG}_C = \chi_C^2 \end{cases}$$

$$\begin{cases} \vec{E}_A \cdot \vec{E}_B = \text{LG}_A \cdot \text{LG}_B = \text{V}\Gamma_{AB} \\ \vec{E}_B \cdot \vec{E}_C = \text{LG}_B \cdot \text{LG}_C = \text{V}\Gamma_{BC} \\ \vec{E}_C \cdot \vec{E}_A = \text{LG}_C \cdot \text{LG}_A = \text{V}\Gamma_{CA} \end{cases}$$

where $\chi_{A,B,C}^2$ are 3 independent χ^2 rv with 1 dof and $\text{V}\Gamma_{AB,BC,CA}$ are 3 variance-Gamma (V Γ) rv [31], [32]. Any vector \vec{U} may be written as

$$\vec{U} = \begin{pmatrix} u_A \\ u_B \\ u_C \end{pmatrix} = u_A \vec{E}_A + u_B \vec{E}_B + u_C \vec{E}_C$$

where u_A, u_B, u_C are 3 constant scalars since all the random part is carried by the basis vectors. The scalar product between 2 vectors \vec{U} and \vec{V} is then:

$$\begin{aligned} \vec{U} \cdot \vec{V} &= \left(u_A \vec{E}_A + u_B \vec{E}_B + u_C \vec{E}_C \right) \cdot \\ &\quad \left(v_A \vec{E}_A + v_B \vec{E}_B + v_C \vec{E}_C \right) \\ &= u_A v_A \vec{E}_A \cdot \vec{E}_A + u_B v_B \vec{E}_B \cdot \vec{E}_B + u_C v_C \vec{E}_C \cdot \vec{E}_C \\ &\quad + (u_A v_B + u_B v_A) \vec{E}_A \cdot \vec{E}_B \\ &\quad + (u_B v_C + u_C v_B) \vec{E}_B \cdot \vec{E}_C \\ &\quad + (u_C v_A + u_A v_C) \vec{E}_C \cdot \vec{E}_A. \end{aligned}$$

On the other hand, if we consider the mathematical expectation of these expressions, we obtain

$$\mathbb{E} \left[\vec{E}_P \cdot \vec{E}_Q \right] = \delta_{P,Q} \quad \text{with } P, Q \in \{A, B, C\}$$

where $\delta_{P,Q}$ is the Kronecker delta. We see that we obtain the classical scalar product by using the mathematical expectation:

$$\mathbb{E} \left[\vec{U} \cdot \vec{V} \right] = u_A v_A + u_B v_B + u_C v_C.$$

Therefore, we will define that 2 vectors \vec{U} and \vec{V} are orthogonal if $\mathbb{E} \left[\vec{U} \cdot \vec{V} \right] = 0$.

B. From a normal Random Variable Product to a Chi-squared RV Difference

Following this formalism, Eq. (8) may be rewritten as

$$\begin{aligned} Z &= \left(\vec{A} + \vec{C} \right) \cdot \left(\vec{B} + \vec{C} \right) = \begin{pmatrix} a \\ 0 \\ c \end{pmatrix} \cdot \begin{pmatrix} 0 \\ b \\ c \end{pmatrix} \quad (9) \\ &= ab\text{V}\Gamma_{AB} + ac\text{V}\Gamma_{AC} + bc\text{V}\Gamma_{BC} + c^2\chi_C^2 \end{aligned}$$

where a, b, c are respectively the standard deviations of the rv A, B, C . As a consequence, $\mathbb{E}[Z] = c^2$. In the following, we will use the noise variances $\sigma_A^2 = a^2, \sigma_B^2 = b^2$ and the signal variance $\sigma_C^2 = c^2$.

As demonstrated in [33], a product of independent normal rv may be expressed as a difference of χ^2 rv. For this purpose, although we know that $(A + C)$ and $(B + C)$ are not independent, we introduce the rv $V_1 = (A + B)/2 + C$ and $V_2 = (A - B)/2$ in such a way that $A + C = V_1 + V_2$, $B + C = V_1 - V_2$ and therefore $(A + C)(B + C) = V_1^2 - V_2^2$. In this vectorial formalism:

$$\vec{V}_1 = \begin{pmatrix} a/2 \\ b/2 \\ c \end{pmatrix}, \quad \text{and} \quad \vec{V}_2 = \begin{pmatrix} a/2 \\ -b/2 \\ 0 \end{pmatrix}.$$

Therefore, (\vec{V}_1, \vec{V}_2) is the basis of the 2-dimensional subspace of $\mathbb{L}\mathbb{G}^3$ in which lies our whole problem. Since the squared modulus of \vec{V}_1, \vec{V}_2 are:

$$\begin{cases} \|\vec{V}_1\|^2 &= \frac{a^2}{4}\chi_A^2 + \frac{b^2}{4}\chi_B^2 + c^2\chi_C^2 \\ &\quad + \frac{ab}{2}\text{V}\Gamma_{AB} + ac\text{V}\Gamma_{AC} + bc\text{V}\Gamma_{BC} \\ \|\vec{V}_2\|^2 &= \frac{a^2}{4}\chi_A^2 + \frac{b^2}{4}\chi_B^2 - \frac{ab}{2}\text{V}\Gamma_{AB}, \end{cases}$$

their difference is consistent with Eq. (9) and then $Z = (\vec{A} + \vec{C}) \cdot (\vec{B} + \vec{C}) = \|\vec{V}_1\|^2 - \|\vec{V}_2\|^2$. Moreover, we can calculate the mathematical expectations of these squared modulus:

$$\begin{aligned} v_1^2 &= \mathbb{E} \left[\|\vec{V}_1\|^2 \right] = \frac{a^2 + b^2}{4} + c^2 \\ v_2^2 &= \mathbb{E} \left[\|\vec{V}_2\|^2 \right] = \frac{a^2 + b^2}{4}. \end{aligned} \quad (10)$$

On the other hand, since

$$\mathbb{E} \left[\vec{V}_1 \cdot \vec{V}_2 \right] = \frac{a^2 - b^2}{4} \quad (11)$$

the vector \vec{V}_1 and \vec{V}_2 are not orthogonal unless $a = b$, i.e. A and B have the same variance.

C. A Particular Case: A and B have the Same Variance

Let us define $\sigma_N^2 = \sigma_A^2 = \sigma_B^2 = n^2$, i.e. $n = a = b$. In this case

$$\mathbb{E} \left[\vec{V}_1 \cdot \vec{V}_2 \right] = \frac{n^2}{4} - \frac{n^2}{4} = 0,$$

\vec{V}_1, \vec{V}_2 are orthogonal which means that their squared modulus are 2 independent χ^2 rv:

$$\|\vec{V}_1\|^2 = v_1^2 \chi_{v_1}^2 \quad \text{and} \quad \|\vec{V}_2\|^2 = v_2^2 \chi_{v_2}^2$$

Thanks to [31, Appendix A], we know that this χ^2 rv difference is a $\text{V}\Gamma$ rv with a Probability Density Function (PDF), introduced by [34]:

$$p(x) = \frac{\gamma^{2\lambda} |x - \mu|^{\lambda-1/2} K_{\lambda-1/2}(\alpha|x - \mu|) e^{\beta(x-\mu)}}{\sqrt{\pi}\Gamma(\lambda)(2\alpha)^{\lambda-1/2}} \quad (12)$$

where $\gamma = \sqrt{\alpha^2 - \beta^2}$, $\Gamma(\lambda)$ is the gamma function, $K_w(z)$ is a hyperbolic Bessel function of the second kind ($w \in \mathbb{R}$ and $z \in \mathbb{C}$) and with the following parameters:

$$\mu = 0, \quad \alpha = \frac{v_1^2 + v_2^2}{4v_1^2 v_2^2}, \quad \beta = \frac{v_1^2 - v_2^2}{4v_1^2 v_2^2}, \quad \lambda = \frac{1}{2} \quad (13)$$

where λ is the number of dof divided by 2. In this particular case, since $a^2 = b^2 = n^2$, v_1^2 and v_2^2 becomes

$$v_1^2 = \mathbb{E} \left[\|\vec{V}_1\|^2 \right] = \frac{n^2}{2} + c^2 \quad \text{and} \quad v_2^2 = \mathbb{E} \left[\|\vec{V}_2\|^2 \right] = \frac{n^2}{2},$$

and we obtain

$$\alpha = \frac{n^2 + c^2}{n^2(2n^2 + c^2)} \quad \text{and} \quad \beta = \frac{c^2}{n^2(2n^2 + c^2)}$$

D. General Case

If $\sigma_A^2 \neq \sigma_B^2$, \vec{V}_1 and \vec{V}_2 are no longer orthogonal and therefore they are 2 correlated rv. We have then to search another set of basis vectors which are orthogonal. For this purpose, let us use the Gram-Schmidt process.

1) *Gram-Schmidt Orthogonalization*: Let us keep \vec{V}_1 unchanged. Let \vec{V}_{2P} be the projection of \vec{V}_2 onto \vec{V}_1 . Denoting θ the angle¹ between \vec{V}_1 and \vec{V}_2 , it comes

$$\vec{V}_{2P} = v_2 \cos(\theta) \frac{\vec{V}_1}{v_1}.$$

with

$$\cos(\theta) = \frac{\mathbb{E}[\vec{V}_1 \cdot \vec{V}_2]}{v_1 v_2}$$

and then

$$\vec{V}_{2P} = \frac{\mathbb{E}[\vec{V}_1 \cdot \vec{V}_2]}{v_1^2} \vec{V}_1. \quad (14)$$

Therefore, we can build the vector \vec{V}_{2N} which is the component of \vec{V}_2 orthogonal to \vec{V}_1 :

$$\vec{V}_{2N} = \vec{V}_2 - \vec{V}_{2P} = \vec{V}_2 - \frac{\mathbb{E}[\vec{V}_1 \cdot \vec{V}_2]}{v_1^2} \vec{V}_1.$$

Using Eq. (10) and (11) yields

$$\begin{aligned} \vec{V}_{2N} &= \begin{pmatrix} a/2 \\ -b/2 \\ 0 \end{pmatrix} - \frac{a^2 - b^2}{a^2 + b^2 + 4c^2} \begin{pmatrix} a/2 \\ b/2 \\ c \end{pmatrix} \\ &= \begin{pmatrix} \frac{a(b^2 + 2c^2)}{a^2 + b^2 + 4c^2} \\ -\frac{b(a^2 + 2c^2)}{a^2 + b^2 + 4c^2} \\ -\frac{c(a^2 - b^2)}{a^2 + b^2 + 4c^2} \end{pmatrix} = \begin{pmatrix} v_{2nA} \\ v_{2nB} \\ v_{2nC} \end{pmatrix}. \end{aligned}$$

We have now to express the measurement vectors $\vec{A} + \vec{C}$ and $\vec{B} + \vec{C}$ as linear combinations of the new basis of orthogonal vectors \vec{V}_1 and \vec{V}_{2N} . In order to do this, we must project these 2 measurement vectors onto the 2 basis vectors in the same way that we have projected \vec{V}_2 onto \vec{V}_1 in Eq. (14):

$$\begin{cases} \vec{A} + \vec{C} &= k_{AC1} \vec{V}_1 + k_{AC2n} \vec{V}_{2N} \\ \vec{B} + \vec{C} &= k_{BC1} \vec{V}_1 + k_{BC2n} \vec{V}_{2N} \end{cases}$$

with

$$\begin{aligned} k_{AC1} &= \frac{\mathbb{E}[(\vec{A} + \vec{C}) \cdot \vec{V}_1]}{\mathbb{E}[|\vec{V}_1|^2]} & k_{AC2n} &= \frac{\mathbb{E}[(\vec{A} + \vec{C}) \cdot \vec{V}_{2N}]}{\mathbb{E}[|\vec{V}_{2N}|^2]} \\ k_{BC1} &= \frac{\mathbb{E}[(\vec{B} + \vec{C}) \cdot \vec{V}_1]}{\mathbb{E}[|\vec{V}_1|^2]} & k_{BC2n} &= \frac{\mathbb{E}[(\vec{B} + \vec{C}) \cdot \vec{V}_{2N}]}{\mathbb{E}[|\vec{V}_{2N}|^2]} \end{aligned}$$

Therefore, $Z = (\vec{A} + \vec{C}) \cdot (\vec{B} + \vec{C})$ may be written as

$$\begin{aligned} Z &= k_{AC1} k_{BC1} |\vec{V}_1|^2 + k_{AC2n} k_{BC2n} |\vec{V}_{2N}|^2 \\ &\quad + (k_{AC1} k_{BC2n} + k_{AC2n} k_{BC1}) \vec{V}_1 \cdot \vec{V}_{2N} \\ &= k_{AC1} k_{BC1} \chi^2 + k_{AC2n} k_{BC2n} \ddot{\chi}^2 \\ &\quad + (k_{AC1} k_{BC2n} + k_{AC2n} k_{BC1}) \text{V}\Gamma \end{aligned} \quad (15)$$

¹In the same way as the orthogonality between 2 vectors is defined by the null mathematical expectation of their scalar product, the angles as well as the other relationships between vectors must be taken into account as mathematical expectation since they are valid on average but not for only one particular realization of these vectors.

where χ^2 and $\ddot{\chi}^2$ are independent χ^2 rv corresponding respectively to the squared norm of \vec{V}_1 and \vec{V}_{2N} . Thus, this relationship involves the difference of 2 χ^2 rv (it can be proved that $k_{AC2n} k_{BC2n} < 0$), which is well known [31], [33], plus a $\text{V}\Gamma$ rv, which makes the problem more complex. In order to simplify this problem, we should find a representation of Eq. (15) in which the cross term is identically null.

2) *Normalization and Rotation of the Basis Vectors*: Let (\vec{V}'_1, \vec{V}'_2) be the normalized equivalent of the basis $(\vec{V}_1, \vec{V}_{2N})$:

$$\vec{V}'_1 = \frac{\vec{V}_1}{\mathbb{E}[|\vec{V}_1|]} \quad \text{and} \quad \vec{V}'_2 = \frac{\vec{V}_{2N}}{\mathbb{E}[|\vec{V}_{2N}|]}.$$

With this new basis, Eq. (15) may be rewritten as

$$\begin{aligned} Z &= k_{AC1} k_{BC1} \mathbb{E}[|\vec{V}_1|]^2 \|\vec{V}'_1\|^2 \\ &\quad + k_{AC2n} k_{BC2n} \mathbb{E}[|\vec{V}_{2N}|]^2 \|\vec{V}'_2\|^2 \\ &\quad + k_{AC1} k_{BC2n} \mathbb{E}[|\vec{V}_1|] \mathbb{E}[|\vec{V}_{2N}|] \vec{V}'_1 \cdot \vec{V}'_2 \\ &\quad + k_{AC2n} k_{BC1} \mathbb{E}[|\vec{V}_1|] \mathbb{E}[|\vec{V}_{2N}|] \vec{V}'_1 \cdot \vec{V}'_2 \\ &= k'_{11} \|\vec{V}'_1\|^2 - k'_{22} \|\vec{V}'_2\|^2 + k'_{12} \vec{V}'_1 \cdot \vec{V}'_2 \end{aligned} \quad (16)$$

with

$$\begin{cases} k'_{11} &= k_{AC1} k_{BC1} \mathbb{E}[|\vec{V}_1|]^2 \\ k'_{22} &= -k_{AC2n} k_{BC2n} \mathbb{E}[|\vec{V}_{2N}|]^2 \\ k'_{12} &= k_{AC1} k_{BC2n} \mathbb{E}[|\vec{V}_1|] \mathbb{E}[|\vec{V}_{2N}|] \\ &\quad + k_{AC2n} k_{BC1} \mathbb{E}[|\vec{V}_1|] \mathbb{E}[|\vec{V}_{2N}|]. \end{cases}$$

We can then consider Eq. (16) as the expression of a quadratic form Q which associate a scalar w_0 to any vector $\vec{W} = w_1 \vec{V}'_1 + w_2 \vec{V}'_2$. Such a quadratic form may be described as

$$w_0 = \vec{w}^T [Q] \vec{w} \quad \text{with} \quad [Q] = \begin{pmatrix} k'_{11} & k'_{12}/2 \\ k'_{12}/2 & -k'_{22} \end{pmatrix}. \quad (17)$$

The simplification of our problem relies then in a rotation of the basis vectors in such a way that the quadratic form matrix $[Q]$ is diagonal. The eigenvalues of $[Q]$ are given by

$$\ell_1 = \frac{k'_{11} - k'_{22} - \sqrt{\Delta}}{2} \quad \text{and} \quad \ell_2 = \frac{k'_{11} - k'_{22} + \sqrt{\Delta}}{2}.$$

with $\Delta = (k'_{11} + k'_{22})^2 + k'_{12}{}^2$. Thanks to this rotation of the basis vectors, Eq. (15) and (16) become

$$Z = \ell_1 \chi^2 + \ell_2 \ddot{\chi}^2.$$

As already stated in § III-C, Z is a $\text{V}\Gamma$ rv with the following PDF parameters:

$$\mu = 0, \quad \alpha = \frac{\ell_1^2 + \ell_2^2}{4\ell_1^2 \ell_2^2}, \quad \beta = \frac{\ell_1^2 - \ell_2^2}{4\ell_1^2 \ell_2^2}, \quad \lambda = \frac{1}{2}. \quad (18)$$

E. Generalization to Larger Degrees of Freedom

In the case of $2m$ dof, i.e. real part + imaginary part multiplied by m averaged uncorrelated spectra, the only change to apply concerns the parameter λ in Eq. (13) and (18) which becomes $\lambda = m$.

According to [35, Eq. 12 p.80] we have the following relation:

$$K_{n+\frac{1}{2}}(z) = \left(\frac{\pi}{2z}\right)^{\frac{1}{2}} e^{-z} \sum_{r=0}^n \frac{(n+r)!}{r!(n-r)!(2z)^r} \quad (19)$$

with $n \in \mathbb{N}$ and $z \in \mathbb{C}$. Moreover $m \in \mathbb{N}^*$ which leads to the relation $n = m - 1$. Therefore let us expand Eq. (12) using Eq. (19):

$$p(x) = \frac{\kappa(\alpha, \beta)^m \epsilon(x, \mu, \alpha, m)}{\Gamma(m)} e^{-\alpha|x-\mu|+\beta(x-\mu)} \quad (20)$$

with the following parameters:

$$\kappa(\alpha, \beta) = \frac{\alpha^2 - \beta^2}{2\alpha} \quad \Gamma(m) = (m-1)!$$

$$\epsilon(x, \mu, \alpha, m) = \sum_{r=0}^{m-1} \frac{(m+r-1)!|x-\mu|^{m-r-1}}{r!(m-r-1)!(2\alpha)^r}$$

F. Validation of the Theoretical Probability Laws by Monte Carlo Simulations

1) *Algorithm Description*: According to § III-D2 the probability density of \hat{Z} , equal to the difference of two independent χ^2 rv, can now be calculated using the function $p(x)$ of the Eq. (20) by assigning the values to the parameters in Eq. (13) and (18). In order to perform this comparison we use two algorithms, one for Monte Carlo (MC) simulation and the other one for computing Eq. (20).

– MC simulation algorithm

The simulation algorithm follows these 6 steps

- S1: Assignment of the 2 noise levels σ_A^2 , σ_B^2 , signal level σ_C^2 and the number of averaging spectra m .
- S2: Drawing of \hat{A} , \hat{B} , \hat{C} , following a normal centered distribution with respectively σ_A , σ_B , σ_C as standard deviation.
- S3: Computation of $\hat{Z} = (\hat{A} + \hat{C})(\hat{B} + \hat{C})$.
- S4: Repetition $2m$ times of the steps S2 to S3 and sum all \hat{Z} values.
- S5: Repetition $N = 10^7$ times of the steps S2 to S4 of this sequence.
- S6: Drawing the histogram of \hat{Z} .

In all simulations, we chose a number of dof $\nu = 2m$ in order to have a real and imaginary part in agreement with the experiment shown in Fig. (1).

– Modeling algorithm

The modeling algorithm follows also 6 steps:

- S1: Assignment of the 2 noise levels σ_A^2 , σ_B^2 , signal level σ_C^2 and the number of averaging spectra m .
- S2: **Independent basis**

- Computation of coefficients v_1^2 , v_2^2 according to Eq. (10)
- if $\sigma_A^2 = \sigma_B^2$ go to step S5 else perform steps S3 and S4

S3: Orthogonalization of the basis

- Computation of coefficients k_{AC1} , k_{AC2n} , k_{BC1} , k_{BC2n} to determine the new basis according to Eq. (15)
- Normalization of the basis by determining coefficients k'_{11} , k'_{22} , k'_{12} according to Eq. (16)

S4: Vector rotation

- Diagonalization of the matrix Q according to Eq. (17)
- Computation of its roots l_1 and l_2

S5: Compute the coefficients α , β , and $\lambda = m$ according to Eq. (13) and (18).

S6: Plotting the probability density with Eq. (20).

2) *When can the Instrument Noises be Assumed to be "About the Same"?*: Although the problem is quite simple when the instrument noises σ_A^2 and σ_B^2 are the same (see § III-C), it becomes more complex when $\sigma_A^2 \neq \sigma_B^2$. The question is then how far can we assume that $\sigma_A^2 \approx \sigma_B^2$ and then use the particular case formalism of § III-C? In order to answer this question, we use Monte-Carlo simulations which were performed according to § III-F1.

Afterwards we perform a histogram of these realizations and compare it first with the PDF obtained from the model without rotation, i.e. by using the VT parameters of Eq. (13), and next with the PDF obtained from the model with rotation, i.e. by using the VT parameters of Eq. (18). Fig. 3 shows an example of such a comparison. In this case ($\sigma_A^2 = 2$, $\sigma_B^2 = 1/2$, $\sigma_C^2 = 0$), the PDF of the model with rotation is in perfect agreement with the histogram whereas there are large discrepancies with the PDF of the model without rotation. We have thus a first result: the model without rotation should not be used when the ratio $\sigma_A^2/\sigma_B^2 \geq 4$.

To improve the efficiency of the test, we compute the theoretical quantiles by using the model without rotation and then deduce from them the theoretical confidence intervals which are often used (68 %, 90 %, 95 % and 99 %). These quantiles and intervals are compared to the ones obtained from the simulation histogram. In one example of Table I, which corresponds to the case plotted in Fig. 3, the confidence intervals are strongly overestimated. For instance, the expected 68 % confidence interval is significantly too large since it encompasses an interval of 76 %. Similarly, the expected 90 % interval is actually a 94 % interval. This reinforces our decision of using the model with rotation for a noise variance ratio ≥ 4 .

We use these 2 approaches, i.e. PDF curve as well as confidence intervals, for many different parameter sets (see Table I). In any case, the agreement between the model with rotation and the Monte-Carlo simulation histograms were perfect, since the residual deviations can be largely assumed to be due to the finite sample number of the simulation (less than 0.05 % of the CDF). However, this test is very interesting for the model without rotation since it allows us to answer to the question which is the title of this section: when can the

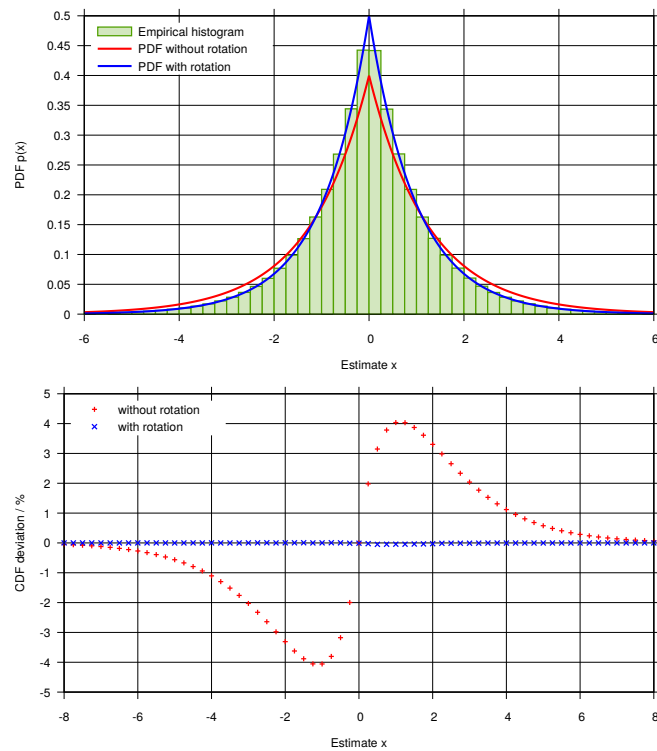


Fig. 3. Comparison of the empirical and theoretical PDF (above) with and without rotation of the basis vectors. The deviations between the empirical and the theoretical CDF are given in the bottom plot. The variances are: $\sigma_C^2 = 0, \sigma_A^2 = 2, \sigma_B^2 = 1/2$ and there are 2 dof.

TABLE I
COMPARISON OF THE EXPECTED QUANTILES AND INTERVALS

Expected probabilities (%)	True probabilities (%)					
	Degrees of freedom: 2				dof: 8	
	$\sigma_C^2 = 0$	$\sigma_C^2 = 0$	$\sigma_C^2 = 0$	$\sigma_C^2 = 0$	0	0.5
	$\sigma_B^2 = 2$	1	2/3	1/2	1	1
Quantiles						
0.5	0.50	0.39	0.25	0.16	0.35	0.39
2.5	2.50	2.10	1.58	1.19	1.98	2.13
5.0	5.00	4.36	3.51	2.82	4.18	4.44
16.0	16.00	14.95	13.43	12.05	14.68	15.18
50.0	50.00	50.01	50.00	50.00	50.00	49.99
84.0	84.00	85.07	86.59	87.96	85.32	84.58
95.0	95.00	95.65	96.50	97.19	95.82	95.37
97.5	97.50	97.91	98.43	98.82	98.02	97.73
99.5	99.50	99.62	99.75	99.84	99.65	99.57
Intervals						
68.0	68.00	70.12	73.16	75.91	70.64	69.41
90.0	90.00	91.29	92.98	94.37	91.64	90.93
95.0	95.00	95.82	96.84	97.63	96.04	95.60
99.0	99.00	99.23	99.50	99.68	99.30	99.18

The expected quantiles (above) and intervals (below) are computed by using the parameters from Eq. (13) with empirical probabilities. For all realizations $\sigma_A^2 = 2$.

instrument noises be assumed to be “about the same”? Table I is very useful in this connection. In a first step, let us study the case where the number of dof is 2 and there is no signal since it is the case which is the most sensitive to the difference between the noise levels. We can see on this table that the model without rotation is perfect when the 2 noise levels are equal ($\sigma_B^2 = 2$), fair when the ratio of the noise levels is equal

to 2 ($\sigma_B^2 = 1$), at the limit of acceptance when the ratio is 3 but not suitable for a ratio ≥ 4 . The other columns of Table I, obtained with 8 dof and with $\sigma_C^2 = \sigma_A^2/4$, confirm that the model without rotation is acceptable when the ratio of the noise variances is equal to 2.

Then we recommend to use the vector rotation process if the ratio of the noise variance greater than 2.

IV. INVERSE PROBLEM

A. Principle of the Method

The bayesian statistician has to solve the inverse problem in order to define a confidence interval for the true variance σ_C^2 , given a set of measurements and a priori information. Thereby the cross-spectrum measurement \hat{Z} is now fixed as well as the instrument noise levels σ_A^2 and σ_B^2 , whereas the signal true variance σ_C^2 appears as a random variable. According to the Bayes theorem the a posteriori density of an unknown true value θ given the measurements, here the cross-spectrum \hat{Z} , is

$$\begin{cases} p(\theta|\hat{Z}) \propto p(\hat{Z}|\theta) \cdot \pi(\theta) \\ \int_0^\infty p(\theta|\hat{Z})d\theta = 1 \end{cases} \quad (21)$$

where $\pi(\theta)$ is the a priori density, named prior and $p(\hat{Z}|\theta)$ is the PDF which corresponds to Eq. (12) determined in the direct problem. It remains to determine the prior $\pi(\theta)$ (i.e. the PDF before any measurement) to compute the a posteriori density.

One of the main issue of Bayesian analysis concerns the choice of this prior. We have no a priori knowledge about the behavior of the parameter θ . A total ignorance of knowledge leads to a prior equal to θ^{-1} which means all order of magnitudes have the same probability. The choice of θ is subject to discussion and the reader should refer to [36, Appendix B].

The quantity that can be actually measured is the sum of the signal and the measurement noise. Hence the prior should be accordingly given as a function of this sum. In other words, it is not possible to have any information on a signal with a level much smaller than the measurement noise. Hence choosing a prior function of $\sigma_N^2 + \sigma_C^2$ ensures that the corresponding magnitude order of σ_C^2 do not dominate the a posteriori probability distribution. The measurement noise level decreases as m^{-1} , according to [5, Eq. 11], when averaging over different spectra realizations m . So it should depend to the number of dof $\nu = 2m$ (i.e. taking in account the real and imaginary part). From these considerations, we choose the following prior according to Fig. 4:

$$\pi(\theta) = \frac{1}{\theta} = \frac{1}{\sigma_N^2/\nu + \sigma_C^2}, \quad (22)$$

where $\sigma_N^2 = (\sigma_A^2 + \sigma_B^2)/2$ is the known, “not random” averaged noise level. Thus small level of σ_C^2 are distributed roughly uniformly on a linear scale and large values are distributed with equal probability for equal logarithmic intervals.

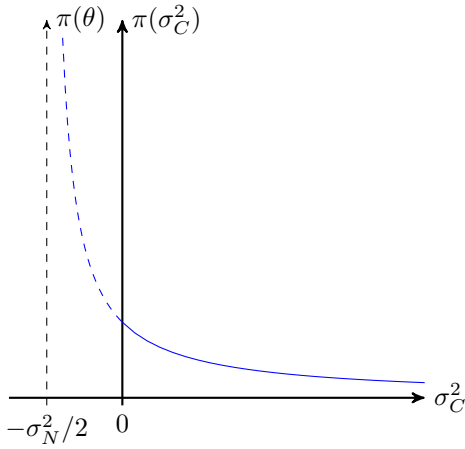


Fig. 4. Prior of the sum of the noise σ_N^2 and signal σ_C^2 levels for the case when there is no averaging spectra (i.e. $\nu = 2$).

B. Check of the Posterior Probability Density Function

According to Eq. (20), for 2 dof or $m = 1$ spectrum average and the particular case $\sigma_A^2 = \sigma_B^2 = \sigma_N^2$, we know that

$$p(\hat{Z}|\sigma_C^2) = \frac{e^{\hat{Z}/\sigma_N^2}}{2(\sigma_N^2 + \sigma_C^2)}.$$

Therefore, the posterior PDF of the cross-spectrum estimator is

$$\begin{cases} p(\sigma_C^2|\hat{Z}) \propto \frac{e^{\hat{Z}/\sigma_N^2}}{2(\sigma_N^2 + \sigma_C^2)(\sigma_N^2 + 2\sigma_C^2)} & \text{if } \hat{Z} \leq 0 \\ p(\sigma_C^2|\hat{Z}) \propto \frac{e^{-\hat{Z}/(\sigma_N^2 + 2\sigma_C^2)}}{2(\sigma_N^2 + \sigma_C^2)(\sigma_N^2 + 2\sigma_C^2)} & \text{if } \hat{Z} \geq 0. \end{cases} \quad (23)$$

We have checked this posterior PDF by using the inverse problem Monte-Carlo algorithm we already used in [31, § IV.B.1)] and [28, § IV.A.]. The principle is the following:

- S1: Select a target estimate $\hat{Z} = Z_0$.
- S2: Draw at random the signal level σ_C^2 according to

$$\sigma_C^2 = 10^{\lceil \eta + u_{[0,1]}(e_{max} - \eta) \rceil} - \frac{\sigma_N^2}{2}$$

where $\eta = \log_{10}(\sigma_N^2/2)$ and $u_{[0,1]}$ is a pseudo-random function which is uniform within $[0, 1]$. This draw ensures the parameter follows the prior of Eq. (22) up to $10^{e_{max}}$. We have chosen $e_{max} = 4$ which is in accordance with Fig. 4

- S3: Draw at random (Gaussian) the noise and signal estimates $\hat{A}, \hat{B}, \hat{C}$ and compute the measurements \hat{X}, \hat{Y} according to Eq. (8).
- S4: Compute the estimate \hat{Z} .
- S5: Compare the estimate \hat{Z} with the target Z_0 : if $\hat{Z} = Z_0 \pm p$, store the current σ_C^2 value as it is able to generate an estimate equal to the target; otherwise throw this σ_C^2 value. We have chosen $p = (Z_0 + \sigma_N^2/2)/50$ when $Z_0 > 0$ and $p = \sigma_N^2/100$ when $Z_0 \leq 0$.
- S6: Go to step 2.
- S7: Stop when a set of 10 000 σ_C^2 values is reached.

TABLE II
COMPARISON OF THE QUANTILES 95 % OBTAINED BY MONTE-CARLO SIMULATION AND BY THE POSTERIOR CDF

Target Z_0 (a.u.)	95 % bound		True prob. (%)
	Emp	Theo	
-1.00	14.04	13.65	94.90
0.00	15.11	13.65	94.53
0.10	14.52	14.27	94.91
0.20	14.98	14.90	94.96
0.32	15.87	15.66	94.94
0.50	17.37	16.90	94.87
1.00	20.40	20.14	94.93
2.00	27.61	28.65	95.18
3.16	38.19	39.08	95.08
5.00	57.15	56.55	94.91
10.00	109.66	104.82	94.78

The quantiles 95 % are computed for a noise level $\sigma_N^2 = 1$ a.u. The theoretical bounds (denoted ‘‘Theo’’) are obtained by numerical integration and then correspond to the true probabilities (denoted ‘‘True prob.’’).

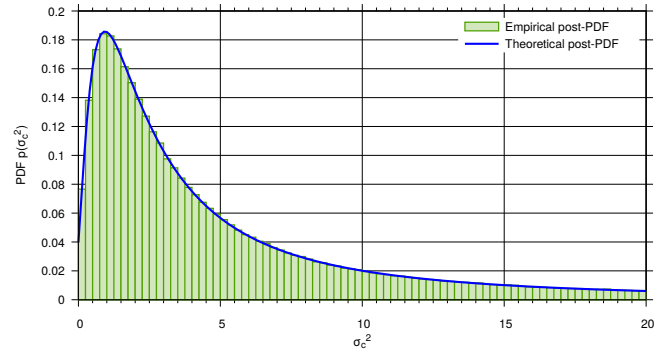


Fig. 5. Comparison of the empirical and theoretical posterior PDF for a noise level $\sigma_N^2 = 1$ a.u. and a target estimate $Z_0 = 1$ a.u.

It must be noticed that such an algorithm is obviously not able to justify the choice of the prior since this prior is included in the algorithm. It will only ensure that no mistake has been done in the expression of the posterior PDF.

Fig. 5 shows the comparison of the posterior PDF computed according to Eq. (23) (blue curve) and the histogram obtained thanks to the inverse problem Monte-Carlo algorithm (green boxes) with a noise level $\sigma_N^2 = 1$ a.u. and a target estimate $Z_0 = 1$ a.u. We can verify that the agreement is excellent.

Table II compares the 95% quantiles obtained by the inverse problem Monte-Carlo algorithm (denoted ‘‘Emp’’ for empirical) and by the integration of the posterior PDF (denoted ‘‘Theo’’ for theoretical), i.e. the posterior CDF, for different values of target and for a noise level $\sigma_N^2 = 1$. Here also the agreement is very good whether for the 95% bounds or for the true probabilities of the theoretical bounds. Moreover, the fluctuations of the empirical bounds prove that the slight differences between empirical and theoretical values are due to the fluctuations of the empirical bounds because of the limited number of realizations (10 000) of the inverse problem Monte-Carlo algorithm.

C. KLT Method

The KLT method stands for ‘‘Karhunen-Loève Transform’’ and was developed in our previous paper [28]. In that paper,

KLT has proved to be as efficient as well as rigorous method, making the most of the property of “sufficient statistics”. However the difference with [28] is that we don’t have the “sufficient statistics” property (see [37]). It means that KLT method will not give the same result as the cross-spectrum method whereas it should have in the case of “sufficient statistics”. First let us remind the theory. Then in a second time, we will explain what can bring the KLT method in addition to the cross-spectrum one.

1) *A Posteriori Distribution*: The KLT method relies on the use of \hat{X} , and \hat{Y} measurements according to Eq. (2), which are Gaussian rv instead of the product of \widehat{AB} , \widehat{AC} , \widehat{BC} and \widehat{C}^2 , which are linear combination of Bessel of the second kind functions and χ^2 random variables. The main advantage of this approach lays in the property of the Gaussian rv which remain Gaussian when they are linearly combined. However, these measurements are not independent. That is why we aim to determine two linear combinations of these rv that are independent one of each other. Hence we define the covariance matrix between \hat{X} and \hat{Y} given by

$$M = \begin{pmatrix} \sigma_A^2 + \sigma_C^2 & \sigma_C^2 \\ \sigma_C^2 & \sigma_B^2 + \sigma_C^2 \end{pmatrix}. \quad (24)$$

The KLT consists in using the rv corresponding to the diagonalization of this matrix. In order to simplify the equations we study solely the case where $\sigma_A^2 = \sigma_B^2 = \sigma_N^2$. The eigenvalues of M are

$$\begin{aligned} \lambda_1 &= \sigma_N^2 + 2\sigma_C^2 \\ \lambda_2 &= \sigma_N^2 \end{aligned} \quad (25)$$

with the following normalized eigenvectors,

$$V_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad V_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (26)$$

The likelihood function is then given by

$$p_{KLT}(\hat{Z}|\sigma_C^2) = \prod_{i=1}^2 \frac{1}{\lambda_i^{\nu/2}} \exp\left(-\frac{\sum_{j=1}^{\nu} \hat{w}_{i,j}^2}{2\lambda_i}\right) \quad (27)$$

The numerator of the exponential argument is then the only term that depends on the actual measurements:

$$\hat{w}_{i,j}^2 = V_{i,1}^2 \hat{X}_j^2 + V_{i,2}^2 \hat{Y}_j^2 + 2V_{i,1}V_{i,2} \hat{Z}_j \quad (28)$$

where $\|V_i\|^2 = \sum_j V_{i,j}^2$. So the KLT method involve the spectral density \hat{X}^2 , \hat{Y}^2 in addition to the cross-spectrum.

Keeping the same prior defined in Eq. (22) we have the following a posteriori density,

$$\begin{cases} p_{KLT}(\sigma_C^2|\hat{Z}) \propto \frac{1}{\sigma_N^2/2 + \sigma_C^2} \cdot p_{KLT}(\hat{Z}|\sigma_C^2) \\ \int_{\mathbb{R}} p_{KLT}(\sigma_C^2|\hat{Z}) d\sigma_C^2 = 1. \end{cases} \quad (29)$$

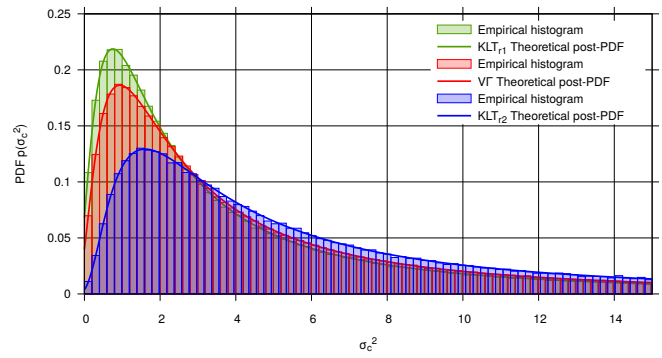


Fig. 6. Comparison of the empirical and theoretical posterior PDF for VF and KLT methods with a noise level $\sigma_N^2 = 1$ a.u and a target estimate $Z_0 = 5$ a.u. KLT_{r1} and KLT_{r2} are the same method but differ by their combination of spectral density \hat{X}^2 and \hat{Y}^2 which are fixed, see Table III, whereas they are rv for the VF method.

2) *Validation of the Method by Monte Carlo Simulation*: In order to validate the KLT method, we have compared its results to Monte Carlo simulations. The algorithm is as follows:

S1: Select a noise level $\sigma_N^2 = \sigma_A^2 = \sigma_B^2$, a target $\hat{Z} = Z_0$ and a combination $\hat{X} = X_0$, $\hat{Y} = Y_0 = Z_0/X_0$ for all the dof.

S2: Draw at random the signal level σ_C^2 according to

$$\sigma_C^2 = 10^{\lceil \eta + u_{[0,1]}(e_{max} - \eta) \rceil} - \frac{\sigma_N^2}{2}$$

where $\eta = \log_{10}(\sigma_N^2/2)$ and $u_{[0,1]}$ is a pseudo-random function which is uniform within $[0, 1]$. This draw ensures the parameter follows the prior of Eq. (22) up to $10^{e_{max}}$. We have chosen $e_{max} = 4$.

S3: Draw at random (Gaussian) the noise and signal estimates \hat{A} , \hat{B} , \hat{C} and compute the measurements \hat{X} , \hat{Y} according to Eq. (8).

S4: Compute the estimates \hat{X} and \hat{Y} .

S5: Compare the estimates \hat{X} , \hat{Y} with the targets X_0 , Y_0 for all the dof: if $\hat{X} = X_0 \pm p$, $\hat{Y} = Y_0 \pm q$, store the current σ_C^2 value as it is able to generate an estimate equal to the target; otherwise throw this σ_C^2 value. We have chosen a precision p , q of tenths of respectively X_0 and Y_0 .

S6: Go to step 2.

S7: Stop when a set of n σ_C^2 values is reached. The number of values n depending on the computation time.

3) *Results and Discussion*: Fig. 6 shows the comparison between the PDF of VF method developed in § III and the PDF of KLT method for two different realizations. The theoretical post-PDF fits very well the empirical histogram for each method. The “sufficient statistics” property being not valid, different combinations of the spectral density \hat{X} and \hat{Y} were tested and are given in Table III. Indeed KLT_{r1} and KLT_{r2} realizations do not give the same PDF unlike the VF method. KLT_{r1} has then a peak which is higher than the VF method whereas KLT_{r2} has a smaller one. This is explained by a more stringent confidence interval for KLT_{r1} than VF, and a less stringent for KLT_{r2} as referred in Table III. The 95 % quantiles

TABLE III

COMPARISON OF THE 95 % QUANTILES OBTAINED BY MONTE-CARLO SIMULATION (EMPIRICAL) AND BY THE POSTERIOR CDF

Method	Measurement				95 % bound	
	\hat{X}'	\hat{X}''	\hat{Y}'	\hat{Y}''	Emp	Theo
VI	rv	rv	rv	rv	56.4	56.6
KLT _{r1}	1.6	1.6	1.6	1.6	48.4	48.3
KLT _{r2}	4.0	0.6	2.5	1.0	82.3	80.8

The 95 % quantiles are computed for a noise level $\sigma_N^2 = 1$ and a target estimate $Z_0 = 5$ a.u.

obtained with MC simulations are in a good agreement with the theoretical ones, especially for KLT_{r1} and VI methods. It is explained by the number of data which is not the same for all of these simulations. VI, KLT_{r1} and KLT_{r2} have respectively 1 000 000; 500 000 and 245 000 data. VI MC simulations takes only 2 minutes whereas it needs respectively 54 hours and 35 days using 17 cores, for KLT_{r1} and KLT_{r2}. KLT_{r1} is chosen to have the spectral density combination which leads to the most stringent confidence interval. Whereas KLT_{r2} is chosen to be more defavourable than the general case VI using only the knowledge of the cross-spectrum measurement.

The KLT method can then have a slightly more stringent confidence interval than the cross-spectrum method using VI for certain case. However it requires to have the knowledge of both spectral density of each channel. It then uses more information, the “sufficient statistics” property being not valid. So the KLT method is preferred when the spectral densities are known.

V. CONCLUSION

The method developed, VI, provides the Probability Density Function of the signal level studied when using the cross-spectrum method. It allows the determination of confidence intervals through numerical integration, where only the high bound has a physical meaning. It is especially relevant for one or several measurements of the cross-spectrum as the PDF will tend to a Gaussian distribution for many dof.

VI is a rigorous method since it is the exact density solution of the cross-spectrum real part statistics, with no approximation. We shall notice that the noise level of each measurement instruments has to be known. If these noise level are the same except at a factor of 4 and higher, we can assume that all the theoretical part of orthogonalizing and the rotation of the basis is not necessary. This method works whatever the number of measurement spectra and noise level.

However using KLT method to compute the confidence interval is a more rigorous method because it uses the knowledge of the spectral density in addition to the cross-spectrum. That is why we recommend to use the KLT method which turns out to be a slightly better estimator than VI.

VI. 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. B. Blackman and J. W. Tuckey. *The Measurement of Power Spectra*. Dover, 1959.
- [2] Gwilym M. Jenkins and Donald G. Watts. *Spectral Analysis and its Applications*. Holden Day, San Francisco, CA, 1968.
- [3] O. E. Brigham. *The Fast Fourier Transform and its Applications*. Prentice-Hall, 1988.
- [4] D. B. Percival and A. T. Walden. *Spectral Analysis for Physical Applications*. Cambridge, Cambridge, UK, 1993.
- [5] E. Rubiola and F. Verlotte. The cross-spectrum experimental method. arXiv:1003.0113v1, March 2010.
- [6] R. H. Dicke. The measurement of thermal radiation at microwave frequencies. *Rev. Sci. Instrum.*, 17(7):268–275, July 1946.
- [7] M. Sampietro, L. Fasoli, and G. Ferrari. Spectrum analyzer with noise reduction by cross-correlation technique on two channels. *Rev. Sci. Instrum.*, 70(5):2520–2525, May 1999.
- [8] A. Hati, C. W. Nelson, and D. A. Howe. Cross-spectrum measurement of thermal-noise limited oscillators. *Rev. Sci. Instrum.*, 87:034708, March 2016.
- [9] Y. Gruson, V. Giordano, U. L. Rohde, A. K. Poddar, and E. Rubiola. Cross-spectrum pm noise measurement, thermal energy, and metamaterial filters. *IEEE Trans. Ultras. Ferroelec. Freq. Contr.*, 64(3):634–642, March 2017.
- [10] A. C. Cárdenas-Olaya, E. Rubiola, Friedt J.-M., P.-Y. Bourgeois, M. Ortolano, S. Micalizio, and C. E. Calosso. Noise characterization of analog to digital converters for amplitude and phase noise measurements. *Rev. Sci. Instrum.*, 88:065108 1–9, June 2017.
- [11] G. Feldhaus and A. Roth. A 1 MHz to 50 GHz direct down-conversion phase noise analyzer with cross-correlation. In *Proc. Europ. Freq. Time Forum*, York, UK, April 4-7 2016.
- [12] T. M. Fortier, C. W. Nelson, A. Hati, F. Quinlan, J. Taylor, H. Jiang, C. W. Chou, T. Rosenband, A. Lemke, N. Ludlow, D. Howe, C. W. Oates, and S. A. Diddams. Sub-femtosecond absolute timing jitter with a 10 ghz hybrid photonic-microwave oscillator. *Appl. Phys. Lett.*, 100:231111 1–3, June 7 2012.
- [13] E. Rubiola. The measurement of am noise of oscillators. arXiv:physics/0512082, December 2005.
- [14] A. H. Verbruggen, H. Stoll, K. Heeck, and R. H. Koch. A novel technique for measuring resistance fluctuations independently of background noise. *Appl. Phys. A*, 48:233–236, March 1989.
- [15] C. M. Allred. A precision noise spectral density comparator. *J. Res. NBS*, 66C:323–330, October-December 1962.
- [16] J. A. Nanzer and R. L. Rogers. Applying millimeter-wave correlation radiometry to the detection of self-luminous objects at close range. *IEEE Trans. Microw. Theory Tech.*, 56(9):2054–2061, September 2008.
- [17] D. R. White, R. Galleano, A. Actis, H. Brixy, M. De Groot, J. Dubbel-dam, A. L. Reesink, F. Edler, H. Sakurai, R. L. Shepard, and Gallop J. C. The status of Johnson noise thermometry. *Metrologia*, 33(4):325–335, August 1996.
- [18] J. P. W. Verbiest, M. Bailes, W. A. Coles, G. B. Hobbs, W. Van Straten, D. J. Champion, F. A. Jenet, R. N. Manchester, N. D. R. Bhat, J. M. Sarkissian, et al. Timing stability of millisecond pulsars and prospects for gravitational-wave detection. *Month. Not. Roy. Astronom. Soc.*, 400(2):951–968, 2009.
- [19] S. Detweiler. Pulsar timing measurements and the search for gravitational waves. *Am. Astronom. Soc.*, 234:1100–1104, December 1979.
- [20] R. W. Hellings and G. S. Downs. Upper limits on the isotopic gravitational radiation background from pulsar timing analysis. *Am. Astronom. Soc.*, 265:39–42, February 1983.
- [21] S. R. Taylor, M. Vallisneri, J. A. Ellis, C. M. F. Mingarelli, T. J. W. Lazio, and R. van Haasteren. Are we there yet? time to detection of nanohertz gravitational waves based on pulsar-timing array limits. *Astrophys. J. Lett.*, 819(1):6, 2016.
- [22] D. Perrodin and A. Sesana. *Radio Pulsars: Testing Gravity and Detecting Gravitational Waves*, volume 457 of *Astrophysics and Space Science Library*. Springer, Cham, 2018.
- [23] S. Drasco and E. E. Flanagan. Detection methods for non-gaussian gravitational wave stochastic backgrounds. *Phys. Rev. D*, 67(8):082003, 2003.
- [24] N. J. Cornish and J. D. Romano. Towards a unified treatment of gravitational-wave data analysis. *Phys. Rev. D*, 87(12):122003, 2013.
- [25] C. G. Bassa, G. H. Janssen, R. Karuppusamy, M. Kramer, K. J. Lee, K. Liu, J. McKee, D. Perrodin, M. Purver, S. Sanidas, R. Smits, and B. W. Stappers. Leap: the large european array for pulsars. *Month. Not. Roy. Astronom. Soc.*, 456(2):2196–2209, December 2015.

- [26] S. Chen, F. Vernotte, and E. Rubiola. Applying clock comparison methods to pulsar timing observations. *MNRAS*, 2020.
- [27] International vocabulary of basic and general terms in metrology (VIM). International Organization for Standardization (ISO), 2004.
- [28] E. Lantz, C. E. Calosso, E. Rubiola, V. Giordano, C. Fluhr, B. Dubois, and F. Vernotte. KLTS: A rigorous method to compute the confidence intervals for the three-cornered hat and for Gros Lambert covariance. *IEEE Trans. Ultras. Ferroelec. Freq. Contr.*, 66(12):1942–1949, December 2019.
- [29] N. J. Kasdin. Discrete simulation of colored noise and stochastic processes and 1/f power law noise generation. *Proceedings of the IEEE*, 83(5):802–827, May 1995.
- [30] F. Vernotte, G. Zalamansky, and E. Lantz. Time stability characterization and spectral aliasing part i: a time-domain approach. *Metrologia*, 35:723–730, 1998.
- [31] F. Vernotte and E. Lantz. Three-cornered hat and Gros Lambert covariance: A first attempt to assess the uncertainty domains. *IEEE Trans. Ultras. Ferroelec. Freq. Contr.*, 66(3):643–653, March 2019.
- [32] M. Haas and C. Pigorsch. *Financial Economics, Fat-Tailed Distributions*, pages 3404–3435. Springer New York, New York, NY, 2009.
- [33] B. Sorin and P. Thionet. Lois de probabilités de Bessel. *Rev. Stat. App.*, 16(4):65–72, 1968.
- [34] D. Madan and E. Seneta. The variance gamma (v.g.) model for share market returns. *J. Business*, 63(4):511–524, 1990.
- [35] G. N. Watson. *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, 1922.
- [36] M. P. McHugh, G. Zalamansky, F. Vernotte, and E. Lantz. Pulsar timing and the upper limits on a gravitational wave background : a Bayesian approach. *Phys. Rev. D*, 54(10):5993–6000, November 1996.
- [37] G. Saporta. *Probabilités, analyse des données et statistique*. Technip, 1990.