KLTS: A Rigorous Method to Compute the Confidence Intervals for the Three-Cornered Hat and for Groslambert Covariance

Éric Lantz, Claudio Eligio Calosso, Enrico Rubiola, Vincent Giordano, Christophe Fluhr, Benoît Dubois, and François Vernotte

***Abstract*— The three-cornered hat/Groslambert Covariance (GCov) methods are widely used to estimate the stability of each individual clock in a set of three, but no method gives reliable confidence intervals for large integration times. We propose a new KLTS (Karhunen-Loève Tansform using Sufficient statistics) method which uses these estimators to consider the statistics of all the measurements between the pairs of clocks in a Bayesian way. The resulting cumulative density function (CDF) yields confidence intervals for each clock Allan variance (AVAR). This CDF provides also a stability estimator that is always positive. Checked by massive Monte Carlo simulations, KLTS proves to be perfectly reliable even for one degree of freedom. An example of experimental measurement is given.**

***Index Terms*— Allan variance (AVAR), Bayesian analysis, clock stability, confidence interval, covariances, three-cornered hat.**

I. Introduction

ALTHOUGH the three-cornered hat [1] and the Groslambert Covariance (GCov) [2] methods are widely used to measure the stability of each individual clock in a set of three, the only methods that exist to compute error bars are limited to the smallest integration times, i.e., when the number of equivalent degrees of freedom (EDF) is high [3]–[5]. However, there is no reliable method to assess confidence intervals over the estimates if their number of EDF is low, typically 5 or below. However, since this case occurs for the largest integration times, it is an important issue for all applications dealing with long-term stability (e.g., time keeping).

Likewise, another problem frequently arises when using the three-cornered hat or the GCov method: negative variance estimates can occur. Although this issue was already considered by Premoli and Tavella [6], it would be useful to get a method that could provide simultaneously a positive estimate as well as its confidence interval.

In a previous article, we performed a first Bayesian attempt to estimate confidence intervals from the statistics of the three-cornered estimates, but we observed that this method was only valid beyond an EDF of 5 [5]. We propose here a new method, also based on Bayesian inversion, which uses the statistics of the data themselves, instead of the statistics of the estimates. Nevertheless, these statistics of the data can be computed from the values of the estimates, since these estimates form "sufficient statistics" [7] for variance estimation. The resulting cumulative density functions (CDFs) yield the lower and upper bounds of the 95% confidence interval. On the other hand, the CDF at 50% provides a useful stability estimator of the stability of each clock, i.e., the median value, which has the advantage of always being positive.

The performances of this method have been checked by using massive Monte Carlo simulations. The principle of these simulations is described in this article and the comparisons with the theoretical confidence intervals given by our new method are discussed. Finally, this method is applied to the measurement of three cryogenic sapphire oscillators (CSOs).

Manuscript received May 2, 2019; accepted July 23, 2019. Date of publication July 30, 2019; date of current version November 21, 2019. This work was supported in part by the Agence Nationale de la Recherche (ANR) Programmes d'Investissement d'Avenir (PIA) Oscillator IMP under Project 11-EQPX-0033, in part by FIRST-TF under Project 10-LABX-0048), and in part by Conseil Régional de Bourgogne Franche-Comté. *(Corresponding authors: Éric Lantz; François Vernotte.)*

E. Lantz is with the Département d'Optique P.M. Duffieux, FEMTO-ST, UMR 6174 CNRS, Université Bourgogne Franche-Comté, 25030 Besançon, France (e-mail: eric.lantz@univ-fcomte.fr).

C. E. Calosso is with the Physics Metrology Division, Istituto Nazionale di Ricerca Metrologica (INRiM), 10135 Turin, Italy.

E. Rubiola is with the Physics Metrology Division, Istituto Nazionale di Ricerca Metrologica (INRiM), 10135 Turin, Italy, and also with the Department of Time and Frequency, FEMTO-ST, Observatory THETA, UMR 6174 CNRS, Université Bourgogne Franche-Comté, 25030 Besançon, France.

V. Giordano and F. Vernotte are with the Department of Time and Frequency, FEMTO-ST, Observatory THETA, UMR 6174 CNRS, Université Bourgogne Franche-Comté, 25030 Besançon, France (e-mail: francois.vernotte@obs-besancon.fr).

C. Fluhr and B. Dubois are with the FEMTO Engineering, 25030 Besançon, France.

Digital Object Identifier 10.1109/TUFFC.2019.2931911

II. Statement of the Problem

A. Clock Comparisons

Let us consider three independent clocks: A, B, and C. It is possible to compare these clocks by pairs by using three time-interval counters (TICs), as shown in Fig. 1, and to estimate the corresponding Allan variances (AVARs). Let us denote $\bar{\mathsf{y}}_{ABk}$ the kth frequency deviation sample between A and B integrated without dead time over a duration τ, $\bar{\mathsf{y}}_{ABk} = (1/\tau)\int_{t_k}^{t_k+\tau} \mathsf{y}(t)dt$, and $\mathsf{z}_{ABk} = (\bar{\mathsf{y}}_{ABk+1} - \bar{\mathsf{y}}_{ABk})/\sqrt{2}$. For clocks A and B, the AVAR is

$$\sigma_{AB}^2(\tau) = \mathbb{E}\left[\mathsf{z}_{ABk}^2\right] \tag{1}$$

where $\mathbb{E}[\cdot]$ stands for the mathematical expectation of the quantity between the brackets. For the sake of concision of the

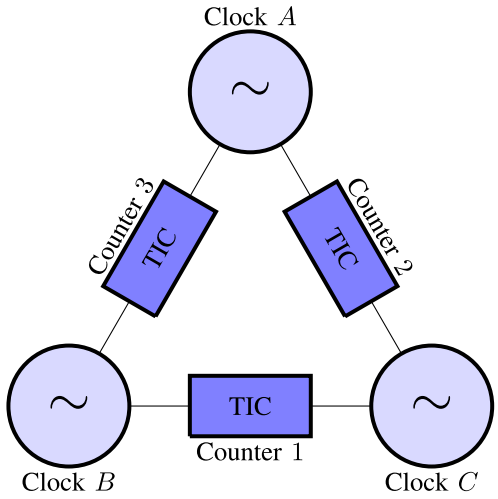

Fig. 1. Layout of the clock measuring device.

notations, we will drop the dependence on τ in the following. Since $\sigma_{AB}^2 = \sigma_A^2 + \sigma_B^2$, the three-cornered hat is based on the following property:

$$\mathrm{TCH}_A = \frac{1}{2}\left(\sigma_{AB}^2 - \sigma_{BC}^2 + \sigma_{CA}^2\right) = \sigma_A^2. \tag{2}$$

On the other hand, the GCov is based on this other property

$$\mathrm{GCov}_A = \mathbb{E}[\mathsf{z}_{ABk} \cdot \mathsf{z}_{ACk}] = \sigma_A^2. \tag{3}$$

B. Measurement Noise Influence

Let us consider the layout of Fig. 1:

1) clocks B and C are connected to counter 1, which is affected by a measurement noise $\{\mathsf{z}_{N1k}\}$;
2) clocks C and A are connected to counter 2, which is affected by a measurement noise $\{\mathsf{z}_{N2k}\}$;
3) clocks A and B are connected to counter 3, which is affected by a measurement noise $\{\mathsf{z}_{N3k}\}$.

We assume that these three measurement noises are uncorrelated.

The measurement given by counter 3 is then the sum of the clock noises plus the corresponding measurement noise: $\mathsf{z}_{3k} = \mathsf{z}_{ABk} + \mathsf{z}_{N3k}$. Its variance is

$$\sigma_3^2 = \mathbb{E}[(\mathsf{z}_{ABk} + \mathsf{z}_{N3k})^2] = \sigma_{AB}^2 + \sigma_{N3}^2 = \sigma_A^2 + \sigma_B^2 + \sigma_{N3}^2. \tag{4}$$

In the following, we will call z_{1k}, z_{2k}, and z_{3k} the *measurements*.

The three-cornered hat becomes $\mathrm{TCH}_A = (1/2)(\sigma_3^2 - \sigma_1^2 + \sigma_2^2) = \sigma_A^2 + (1/2)(\sigma_{N3}^2 - \sigma_{N1}^2 + \sigma_{N2}^2)$. Let us assume that the three counter measurement noises are of the same level: $\sigma_{N1}^2 \approx \sigma_{N2}^2 \approx \sigma_{N3}^2 = \sigma_N^2$. Therefore, $\mathrm{TCH}_A = \sigma_A^2 + (1/2)\sigma_N^2$.

By contrast, the GCov remains, because the measurement noises are uncorrelated

$$\mathrm{GCov}_A = \mathbb{E}[(\mathsf{z}_{ABk} + \mathsf{z}_{N3k}) \cdot (\mathsf{z}_{ACk} + \mathsf{z}_{N2k})] = \sigma_A^2. \tag{5}$$

Therefore, the only difference between these two approaches concerns the measurement noise due to the counters, since the expectation of the GCov estimates is not sensitive to this noise [4]. Note, however, that the variance of the GCov estimates does include a measurement noise term.

C. Model Parameters and Estimates

The previously defined quantities σ_A^2, σ_{AB}^2, TCH_A, GCov_A, and $\sigma_1^2, \ldots$ are unknown real values. However, they can be estimated by estimates that are random variables. We can then define the following estimates:

$$\begin{aligned}\hat{\sigma}_3^2 &= \frac{1}{M}\sum_{k=1}^{M} \mathsf{z}_{3k}^2 = \frac{1}{M}\sum_{k=1}^{M} (\mathsf{z}_{ABk} + \mathsf{z}_{N3k})^2 \\ &= \hat{\sigma}_A^2 + \hat{\sigma}_B^2 + \hat{\sigma}_N^2\end{aligned} \tag{6}$$

where the hat $\hat{(\cdot)}$ stands for *estimate* and M is the number of available consecutive z_{ABk}. Let us call the $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, and $\hat{\sigma}_3^2$ the *elementary estimates*.

Similarly, $\widehat{\mathrm{TCH}}_A = \hat{\sigma}_A^2 + (1/2)\hat{\sigma}_N^2$ and

$$\begin{aligned}\widehat{\mathrm{GCov}}_A &= \frac{1}{M}\sum_{k=1}^{M} \mathsf{z}_{3k}\mathsf{z}_{2k} \\ &= \frac{1}{M}\sum_{k=1}^{M} (\mathsf{z}_{ABk} + \mathsf{z}_{N3k})(\mathsf{z}_{ACk} + \mathsf{z}_{N2k}) \\ &= \hat{\sigma}_A^2.\end{aligned} \tag{7}$$

Let us call the $\hat{\sigma}_A^2$, $\hat{\sigma}_B^2$, and $\hat{\sigma}_C^2$ the *final estimates*.

Meanwhile, let us call the σ_A^2, σ_B^2, and σ_C^2 the *model parameters*.

The aim of this article consists in calculating a confidence interval over each model parameter σ_A^2, σ_B^2, and σ_C^2, from the final and elementary estimates.

D. Estimation of the Measurement Noise by the Closure Relationship

The closure relationship is obtained by computing the sum of the measurements of all counters for a given k

$$\begin{aligned}\mathsf{z}_{cls,k} &= \mathsf{z}_{1k} + \mathsf{z}_{2k} + \mathsf{z}_{3k} \\ &= \mathsf{z}_{BCk} + \mathsf{z}_{N1k} + \mathsf{z}_{CAk} + \mathsf{z}_{N2k} + \mathsf{z}_{ABk} + \mathsf{z}_{N3k} \\ &= \mathsf{z}_{N1k} + \mathsf{z}_{N2k} + \mathsf{z}_{N3k}\end{aligned} \tag{8}$$

since $\mathsf{z}_{ABk} = \mathsf{z}_{Bk} - \mathsf{z}_{Ak}$. These three measurement noises being uncorrelated and of equal level, the variance of the closure is $\sigma_{cls}^2 = \mathbb{E}[(\mathsf{z}_{N1k} + \mathsf{z}_{N2k} + \mathsf{z}_{N3k})^2] = 3\sigma_N^2$.

This gives an efficient way to estimate the variance of the measurement noise

$$\hat{\sigma}_N^2 = \frac{1}{3M}\sum_{k=1}^{M} (\mathsf{z}_{N1k} + \mathsf{z}_{N2k} + \mathsf{z}_{N3k})^2 = \frac{1}{3}\hat{\sigma}_{cls}^2. \tag{9}$$

E. Bayesian Inference

In Bayesian analysis, we have to consider the model parameters $\vec{\Theta} = (\theta_1, \ldots, \theta_m)^T$ that have, in the model world, m-definite but unknown values, and the measurements $\vec{X} = (x_1, \ldots, x_n)^T$ that are n random variables. In our case, the parameters are the three true σ_A^2, σ_B^2, and σ_C^2 AVAR values of the three clocks and the measurements are either the elementary estimates $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, and $\hat{\sigma}_3^2$ (see the previous method described in [5]) or directly the $3M$ measurements $\{\mathsf{z}_{ABk}, \mathsf{z}_{BCk}, \mathsf{z}_{ACk}\}$. In the present method, we will compute

the probability density function (PDF) of these $3M$ measurements, by using only the measured values of the final and elementary estimates, because these estimates form a sufficient statistics for the measurements [7], [8]: the precise knowledge of the set of measurements does not bring any new information beyond the estimates.

While using the same estimates as [5], the present method computes only the Gaussian PDF of the measurements themselves, while [5] lays on the approximation of the PDF of the estimators by a Gaussian law, inducing errors for a low number of EDF.

The Bayesian inversion lays on the distinction between two issues.

1) The *direct problem*, which consists, in the model world, in calculating the PDF of the estimates knowing the model parameters $p(\vec{X}|\vec{\Theta})$.
2) The *inverse problem*, which consists, in the experimenter world, in calculating the PDF of the model parameters knowing the estimates $p(\vec{\Theta}|\vec{X})$. This is the most precise knowledge we can gain on these parameters after a measurement.

The direct problem has been solved in [5] and the inverse problem may be solved thanks to the Bayes theorem

$$\begin{cases} p(\vec{\Theta}|\vec{X}) \propto \pi(\vec{\Theta}) \cdot p(\vec{X}|\vec{\Theta}) \\ \int p(\vec{\Theta}|\vec{X}) d\vec{X} = 1 \end{cases} \tag{10}$$

where $\pi(\vec{\Theta})$, called the prior, is the *a priori* probability of the parameter $\vec{\Theta}$ before any measurement.

III. The KLTS Method

In the following, to distinguish our previous method described in [5] and the present method, we will call the former the KLTG method, for "Karhunen–Loève transform with Gaussian approximation" method, and the latter the KLTS method, for "Karhunen–Loève transform using sufficient statistics" method.

A. Using the Measurements z

The KLTS method relies on the use of the z_{1k}, z_{2k}, and z_{3k} measurements, which are Gaussian random variables (r.v.), instead of the $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$, and $\hat{\sigma}_3^2$ elementary estimates, which are a linear combination of χ^2 random variable. The main advantage of this approach lays in the property of the Gaussian r.v., which remain Gaussian when they are linearly combined.

However, these measurements are strongly correlated for two reasons.

1) The z_{ABk} and z_{ABk+1} are not independent (except in the case of White FM and AVAR without overlapping).
2) For a given k, the z_{1k}, z_{2k}, and z_{3k} are not independent (e.g., their sum is null if the measurement noise is neglected).

We will postpone the treatment of the correlation between successive measurements. We first treat the correlations among the three measurements at a given time. Hence, let us assume that the $3M$ measurements form M independent triplets, successive realizations of three correlated Gaussian r.v. We aim to determine three linear combinations of these r.v. that are independent of each other

$$\mathsf{w}_{l,k} = \alpha_l \mathsf{z}_{1,k} + \beta_l \mathsf{z}_{2,k} + \gamma_l \mathsf{z}_{3,k}, \quad l = 1, 2, 3. \tag{11}$$

The solution of this problem is given by the Karhunen–Loève (K.L.) transform: the nine coefficients $\alpha_l, \beta_l, \gamma_l$ form the eigenvectors of the rotation matrix that diagonalizes the covariance matrix of the measurements, whose diagonal elements are given by (4) and nondiagonal elements by (5). Note that this matrix is singular in the absence of measurement noise because $\mathsf{z}_{BCk} = \mathsf{z}_{ACk} - \mathsf{z}_{ABk}$: if the measurement noise is negligible, a 2×2 matrix must be used instead of a 3×3 matrix.

Now, since the $\mathsf{w}_{l,k}$ are all independent, the PDF is easy to calculate in the model world, by assuming definite values for the true variances: $p(\vec{W}|\vec{\Theta}) = \prod_{k=1}^{M} \prod_{l=1}^{3} p(\mathsf{w}_{l,k}|\vec{\Theta})$. To render more explicit this expression, let us introduce the three K.L. variances V_l obtained by diagonalization of the covariance matrix. All PDFs are Gaussian and their product can be written as

$$p(\vec{W}|\vec{\Theta}) = \prod_{l=1}^{3} \frac{1}{V_l^{M/2}} \exp\left(-\frac{\sum_{k=1}^{M} \mathsf{w}_{l,k}^2}{2V_l} \right). \tag{12}$$

B. Using the Sufficient Statistic Properties of the Estimates

In this equation, it is quite interesting to develop, using (11), the numerator of the exponential argument, which is the only term that depends on the actual measurements

$$\sum_{k=1}^{M} \mathsf{w}_{l,k}^2 = M \left[\alpha_l^2 \hat{\sigma}_1^2 + \beta_l^2 \hat{\sigma}_2^2 + \gamma^2 \hat{\sigma}_3^2 + 2\alpha\beta \widehat{\mathrm{GCov}}_A - 2\alpha\gamma \widehat{\mathrm{GCov}}_B + 2\beta\gamma \widehat{\mathrm{GCov}}_C \right]. \tag{13}$$

Equation (13) means that the only knowledge of the six elementary or final estimates is sufficient to compute the PDF of the actual set of $3M$ measurements. The number of estimates reduces to two in the absence of measurement noise (in this case, the final estimates can be computed from the elementary ones).

This result was expected: indeed, the variance estimates form a sufficient statistics for the variance estimation, meaning that the precise knowledge of the set of measurements $\vec{X} = (x_1, \ldots, x_n)^T$ does not bring any new information beyond the estimates. More precisely, the vector of estimates $\vec{E}$ forms a sufficient statistics for $\vec{\Theta}$ if $p(\vec{X}|\vec{\Theta}) = f(\vec{E}, \vec{\Theta}) \cdot g(\vec{X})$, where f and g are two functions [8]. Indeed, we obtain after Bayesian inversion: $p(\vec{\Theta}|\vec{X}) \propto \pi(\vec{\Theta}) \cdot p(\vec{X}|\vec{\Theta}) \propto \pi(\vec{\Theta}) \cdot f(\vec{E}, \vec{\Theta}) \cdot g(\vec{X})$. In the experimenter world, $\vec{X}$ is formed by actual measurements, with known values. Hence, $g(\vec{X})$ appears as a constant. On the other hand, $p(\vec{\Theta}|\vec{X})$ is a function of the random variable $\vec{\Theta}$, whose integral is unity. Because of this normalization, the multiplication by the constant $g(\vec{X})$ does not change the value of the function that is entirely determined by $f(\vec{E}, \vec{\Theta})$ and the prior. However, it does not mean that only the PDF of the estimates makes sense. We prove in this

article that the best way is to use these estimates to compute the Gaussian PDF of the data themselves that remain Gaussian after applying the K.L. transform.

Because we only use the estimates, the correlation between successive data does not lead to more complexity. Indeed, let N be the number of degrees of freedom corresponding to a set of M measurements, with $N < M$. The estimation of N has been reported in [9]. We have to simply write (12) for N independent measurements by using estimates computed on M correlated measurements, giving

$$p(\vec{W}|\vec{\Theta}) = \prod_{l=1}^{3} \frac{1}{V_l^{N/2}} \exp\left(-\frac{N\sum_{k=1}^{M} \mathsf{w}_{l,k}^2}{2MV_l}\right). \qquad (14)$$

C. KLTS Algorithm

To compute $p(\vec{\Theta}|\vec{W})$, we apply the same approach as in [5]. To consider all the *a priori* values of $\vec{\Theta}$, we use a Monte Carlo scheme with random sampling. This sampling ensures the observance of the total ignorance *a priori* law: the samples are chosen at random on a logarithmic scale, independently of each variance. We work in the experimenter point of view: we assume that two triplets of estimates, final and elementary, have been calculated from the $3M$ elementary measurements. These six numbers have six definite values that will be used in the calculations detailed below. The different steps of the calculation are performed in the same order as in [5] (the common steps of the two methods are recalled here, for the sake of completeness).

1) Choose at random a triplet of true variances, with a uniform probability on a logarithmic scale for each variance and independence between the three variances.
2) Calculate for this triplet the covariance matrix of the measurements given by (1) and (3).
3) Calculate the eigenvectors and eigenvalues of this covariance matrix.
4) Use these values to calculate the PDF given in (14), using (13).
5) For an exhaustive exploration of the PDF, repeat 10^7 times the entire process.
6) For each of the three variables, normalize the probability densities by dividing by their sum (sum of 10^7 values).
7) Also for each of the three variables, sort the true variance values and calculate the CDF by a partial sum of the associated normalized probability densities.
8) Determine a confidence interval at 95% on each true variance from the corresponding CDFs.
9) Verify that the low limit of the confidence interval is meaningful. For a Gaussian distribution, 99.7% of data are included in a confidence interval at $\pm 3\sigma$. If the low limit of this $\pm 3\sigma$ confidence interval (in logarithmic scale) is smaller than the low limit of the *a priori* range (here 10^{-5}), we suspect (and have verified) that the low limit of the $\pm 2\sigma$ confidence interval calculated in the preceding step will depend on the low limit of the *a priori* range. If it occurs, we replace the low limit of the confidence interval by 0.

Finally, we calculate the median value, i.e., the argument giving the CDF equal to 0.5. This value, always positive, may be an alternative estimate of the parameters.

IV. Validation of the KLTS Method by Monte Carlo Simulations

A. Principle of the Simulation

In order to validate the KLTS method, we have compared its results with Monte Carlo simulations.

The algorithm is as follows.

1) Select a target set of three final estimates $(\hat{\sigma}_A^2, \hat{\sigma}_B^2, \hat{\sigma}_C^2) = (A_0, B_0, C_0)$. We call it "reference estimate set."
2) Draw at random a parameter triplets $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$.
3) Randomly draw $3N$ measurements $\{\mathsf{z}_{ABk}, \mathsf{z}_{BCk}, \mathsf{z}_{CAk}\}$ according to $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$.
4) Compute the final estimates $(\widehat{\mathrm{GCov}}_A, \widehat{\mathrm{GCov}}_B, \widehat{\mathrm{GCov}}_C)$ from these measurements.
5) If this final estimate set is close to the reference estimate set within 10%,[1] the corresponding parameter triplet $(\sigma_A^2, \sigma_B^2, \sigma_C^2)$ is kept; otherwise, it is thrown.
6) Go to Step 2).

Each simulation run stops when 10 000 achievements have been obtained to provide a meaningful knowledge of the parameter statistical distributions.

This ensemble of 10 000 parameter triplets giving $(\hat{\sigma}_A^2, \hat{\sigma}_B^2, \hat{\sigma}_C^2) = (A_0, B_0, C_0)$ is then compared with the confidence interval obtained by the KLTS method.

Thanks to the sufficient statistic properties of the estimates, it turns out that any $3N$ measurement set providing the reference estimate set leads to the same statistical distribution of the parameter triplet.

B. Results and Discussion

We will focus this study only on the efficiency of the method, i.e., its ability to fit the true confidence intervals, and not on the behavior of the confidence intervals versus different parameters. More details on this latter subject may be found in [5].

This validation way checks the accuracy of the KLTS method with respect to several physical variables.

1) Influence of the EDF: In order to compare the confidence interval given by several other methods, we chose to set the final estimate values to $\hat{\sigma}_A^2 = \hat{\sigma}_B^2 = \hat{\sigma}_C^2 = 1$ and to vary the number of EDF. The results given by the different methods, a simple Gaussian approximation [4], the Ekstrom–Koppang (EK) method [3], and the KLTG method [5], are shown in Fig. 2(a). All these confidence intervals fit pretty well the empirical error bars obtained by 10 000 Monte Carlo simulations above 200 EDF. Between 50 and 200 EDF, EK and KLTG remain usable, but only KLTG fits quite well between 5 and 50 EDF. Finally, below 5 EDF, none of these methods is satisfactory.

[1]This 10% interval was selected by a compromise ensuring that the final estimate set is sufficiently close on a log scale to the reference estimate set, while the calculation time is reasonable (the computation time increases as the cube of the inverse of the interval width).

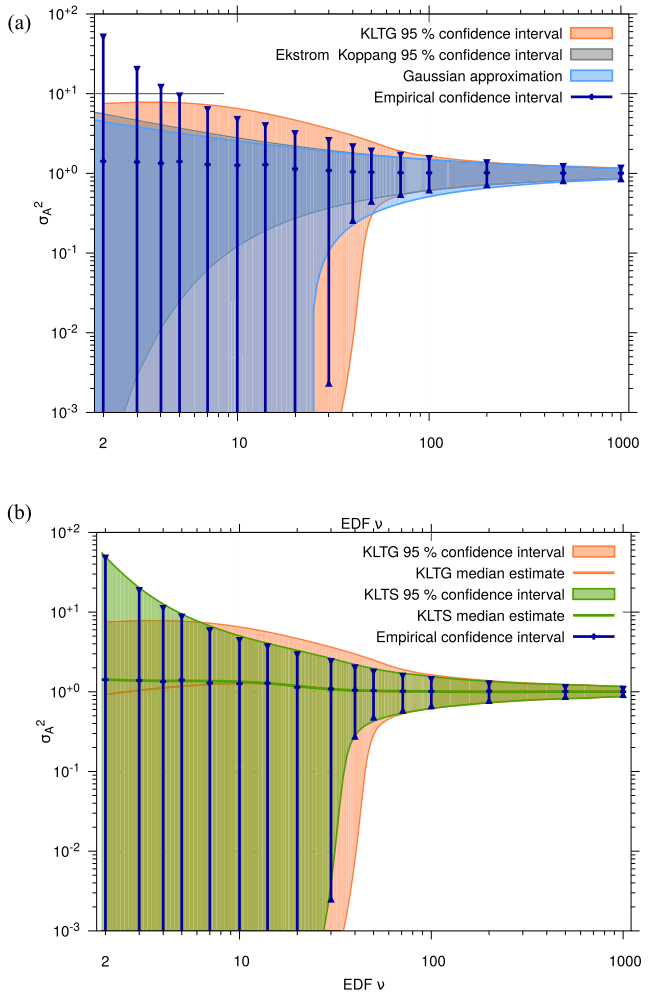

Fig. 2. Estimation of the 95% confidence interval for a set of clocks with equal final estimates ($\hat{\sigma}_A^2 = \hat{\sigma}_B^2 = \hat{\sigma}_C^2 = 1$) versus the EDF number. (a) Results obtained from the previous methods and (b) results of the KLTS method (green area) compared with the KLTG (red area). The reference is given by the blue error bars that were obtained by massive Monte Carlo simulations (see algorithm Section IV-A).

The new KLTS method is compared with KLTG in Fig. 2(b). The KLTG confidence intervals fit pretty well the empirical error bars even for 2 EDF. The same is true for the median estimates that correspond perfectly to the empirical median values at the center of the error bars.

However, this comparison is limited to 2 EDF, because the set $\hat{\sigma}_A^2 = \hat{\sigma}_B^2 = \hat{\sigma}_C^2 = 1$ is impossible for 1 EDF without measurement noise. It was demonstrated that, if the measurement noise is negligible and this is always the case when there is only 1 EDF, the third final estimate is totally determined by the other two ones [5]. For this reason, we set $\hat{\sigma}_B^2 = \hat{\sigma}_C^2 = 1$, which leads to $\hat{\sigma}_A^2 = -0.5$. Table I compares the confidence interval obtained by KLTG and KLTS with the empirical bounds given by 10 000 Monte Carlo simulations.

Unlike the results of KLTG, the KLTS bounds, as well as the median estimate, show good agreement with the empirical bounds and median. However, because of the very low level of the 2.5% bounds, they should be considered as equal to 0. In these conditions, KLTS gives fully reliable 95% upper limits (95.7% for σ_A^2, 95.6% for σ_B^2 and σ_C^2).

TABLE I

COMPARISON FOR 1 EDF OF THE CONFIDENCE INTERVALS AS WELL AS THE MEDIAN ESTIMATES (50%) OBTAINED BY THE MONTE CARLO SIMULATIONS (EMPIRICAL), BY THE KLTG METHOD OF [5], AND BY THE NEW KLTS METHOD. THE (A_0, B_0, C_0) ESTIMATE TRIPLET IS $(-0.5, 1, 1)$

	Bounds	$\hat{\sigma}_A^2 = -0.5$		$\hat{\sigma}_B^2 = \hat{\sigma}_C^2 = 1$	
Emp.	2.5 %	$1.78 \cdot 10^{-5}$	(2.5 %)	$3.0 \cdot 10^{-5}$	(2.5 %)
	50 %	0.192	(50.0 %)	0.90	(50.0 %)
	95 %	27.5	(95.0 %)	76	(95.0 %)
	97.5 %	92	(97.5 %)	182	(97.5 %)
KLTG	2.5 %	$1.36 \cdot 10^{-5}$	(1.52 %)	$1.23 \cdot 10^{-4}$	(5.31 %)
	50 %	$5.5 \cdot 10^{-3}$	(28.7 %)	0.42	(37.0 %)
	95 %	0.83	(68.8 %)	3.23	(70.4 %)
	97.5 %	1.38	(75.0 %)	5.3	(76.3 %)
KLTS	2.5 %	$1.67 \cdot 10^{-5}$	(2.59 %)	$2.86 \cdot 10^{-5}$	(2.37 %)
	50 %	0.200	(50.6 %)	0.90	(49.8 %)
	95 %	35	(95.5 %)	90	(95.3 %)
	97.5 %	98	(97.8 %)	208	(97.6 %)

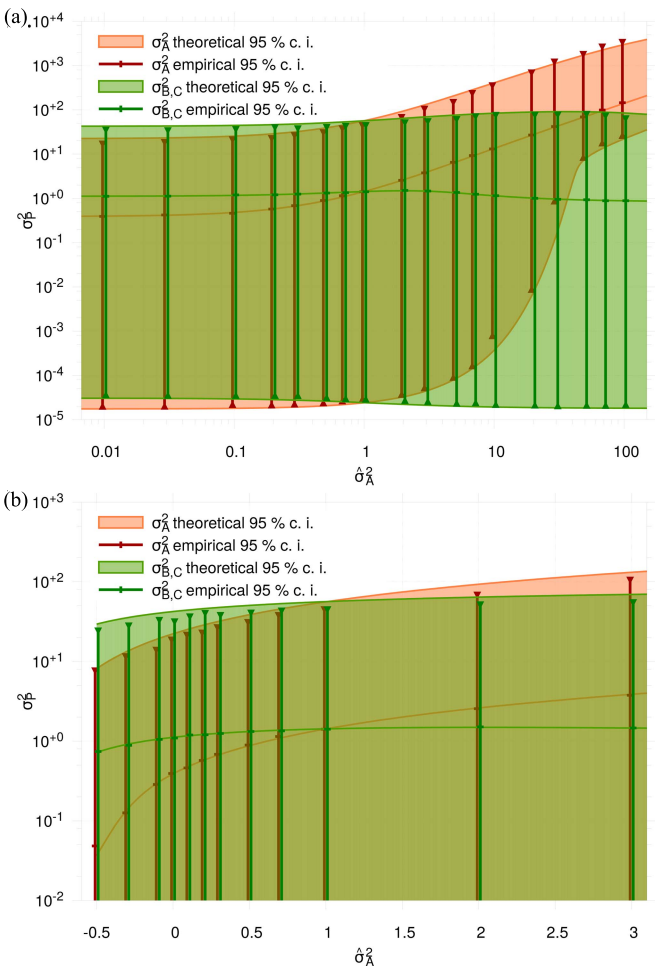

Fig. 3. Estimation of the 95% confidence intervals for two clocks with equal final estimates ($\hat{\sigma}_B^2 = \hat{\sigma}_C^2 = 1$) versus the final estimate of the other clock ($-0.5 \leq \hat{\sigma}_A^2 \leq 100$). The number of EDF is 2. (a) Results obtained from positive final estimates on a log–log plot. (b) Results obtained from negative final estimates using a linear scale. The error bars, red for σ_A^2 and green for σ_B^2 and σ_C^2, were obtained by massive Monte Carlo simulations.

2) Influence of the Disparity of the Three Final Estimates: Fig. 3 shows for 2 EDF the evolution of the confidence interval of all parameters versus the final estimates $\hat{\sigma}_A^2$, which vary

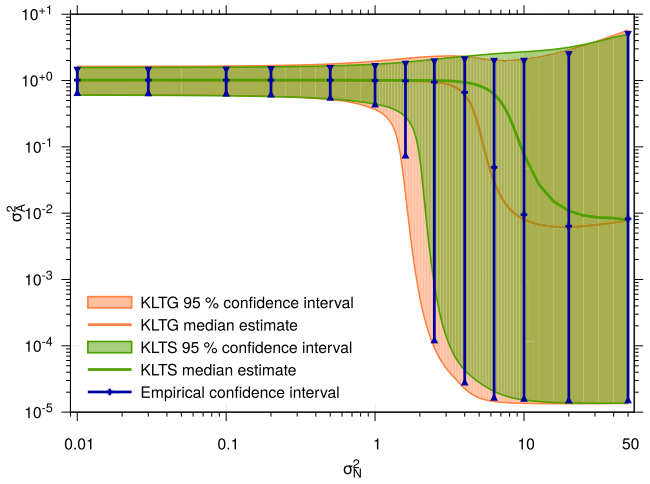

Fig. 4. Estimation of the 95% confidence intervals for a set of clocks with equal noise measurements ($\hat{\sigma}_A^2 = \hat{\sigma}_B^2 = \hat{\sigma}_C^2 = 1$) versus the measurement noise level ($0.01 \leq \hat{\sigma}_N^2 \leq 50$). The number of EDF is 100. The results obtained by the KLTS method (green area) may be compared with the ones obtained by KLTG (red area). The blue error bars were obtained by massive Monte Carlo simulations.

from 0.01 to 100, whereas the two other ones are set to $\hat{\sigma}_B^2 = \hat{\sigma}_C^2 = 1$. On the log–log plot as well as on the linear plot, the agreement between the error bars and the KLTS confidence intervals is excellent.

3) Influence of the Measurement Noise: The main advantage of GCov over the three-cornered hat relies on its rejection of the measurement noise. Since the measurement noise is directly addressed by KLTS, e.g., to solve (13), it is of importance to study its influence on the confidence intervals. This influence is shown in Fig. 4, where the final estimates are set to ($\hat{\sigma}_A^2 = \hat{\sigma}_B^2 = \hat{\sigma}_C^2 = 1$), the number of EDF is 100, and the variance of the measurement noise σ_N^2 varies from 0.01 to 50. Despite slight discrepancies between the confidence intervals and the error bars obtained from the Monte Carlo simulations, the agreement is quite good. The main differences appear for $1 < \sigma_N^2 < 10$ when the 2.5% bound decreases drastically. This effect is also visible for the median estimate that seems to decrease at a different rate for the theoretical and empirical median estimates. A very small discrepancy is also visible for the upper bound at $\sigma_N^2 = 7$, 10, and 20. These discrepancies can be explained as follows: in the PDF computation of KLTS, the product (14) of the elementary probabilities of each data becomes extremely small when the number of data becomes large, and the result of (14) can be wrongly truncated to 0 even in double precision. As a consequence, the use of KLTS must be restricted for cases where the number of EDFs does not exceed 100. For 100 EDF as shown in Fig. 4, the truncation errors already explain some discrepancies for a high level of measurement noise, which do not, however, affect much the upper bound of the confidence interval, undoubtedly the most important result.

4) Discussion: KLTS Versus KLTG: KLTS is undoubtedly a rigorous approach that does not rely on approximations or simplifying assumptions. As a consequence, it provides very relevant confidence intervals even for very low EDF, including the limit case of 1 EDF (see Table I), as proved by the almost perfect agreement between the confidence intervals given by KLTS and obtained by the Monte Carlo simulations in Figs. 2–4. The only slight differences can be attributed to the way of computing the PDF by numerical integration.

However, we have seen that truncation errors can occur with KLTS when the number of data becomes large: the use of KLTS must be restricted for cases where the number of EDFs does not exceed 100.

On the other hand, KLTG relies on a Gaussian approximation of the estimates [5], which can be assumed only for high EDF. Unlike KLTS, KLTG must be restricted for high EDFs. Fig. 4 shows that KLTG and KLTS provide almost the same results for 100 EDFs and these results are in perfect agreement with the Monte Carlo simulations. KLTG is then a very good substitute to KLTS above 100 EDFs.

C. Application to a Set of Real Clocks

These methods were applied to assess the confidence intervals for the GCov measurements of a set of three CSOs designed and made in FEMTO-ST. These measurements were carried out by the Tracking Direct Digital Synthesizer (DDS) designed and made in INRiM (for more details, see [10]). It may be noticed that, since the clocks are in the same room, we cannot fully assume their total independence. Fig. 5 shows the same measurements as the ones shown in [10, Fig. 11 (bottom)], but we have added the 95% confidence intervals obtained by using KLTS (below 100 EDF, i.e., $\tau > 5000$ s) and KLTG (above 100 EDF, i.e., $\tau < 5000$ s).

As expected, the confidence intervals are pretty tight around the estimates for low τ values, whereas they extend downward for high τ values. The lower bounds tend toward 0 above $\tau = 2000$ s for CSO A, above $\tau = 20\,000$ s for CSO C, and above $\tau = 100\,000$ s for CSO B. In these cases, only an upper limit of the stability of the clock may be assessed. As expected, the less stable clocks have the smallest confidence intervals and we can see these intervals increasing or decreasing regarding the relative positions of the other CSOs. For instance, A is the most stable clock between 500 and 30 000 s but becomes the less stable clock around 10^5 s; as a consequence, the lower bound of its confidence interval increases from ~ 0 at 10^4 s to $\sim 2 \cdot 10^{-16}$ at 10^5 s.

The median estimates remain generally close to the GCov estimates. However, when the lower bound of a clock tends toward 0, the median estimate decreases significantly and becomes far lower than the GCov estimate when this latter exists, i.e., when it is positive. Nevertheless, there is still a positive median estimate even when the GCov estimate is negative (e.g., C above $\tau = 30\,000$ s). In such a case, the upper limit of the confidence interval is the relevant information, while the estimate is of little value. It may be noticed that the relevance of the median estimates could be improved by using a more stringent prior $\pi(\vec{\Theta})$ in (10), i.e., a prior based on a perfectly objective *a priori* knowledge of the range in which the parameters can vary. To stay in the most general case, we have preferred to

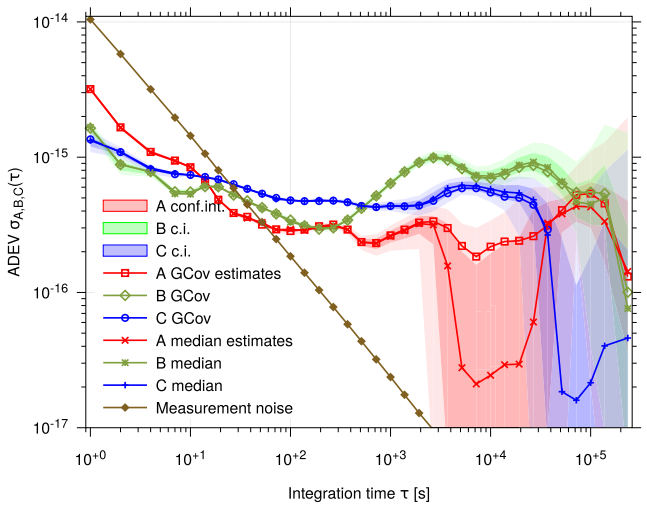

Fig. 5. GCov measurements of a set of three CSOs of FEMTO-ST. The confidence intervals are represented by colored areas: bright colors for 68% c.i. (1σ) and pale colors for 95% c.i. (2σ). Brown line: measurement noise obtained from the closure. The linear-frequency drifts of the CSO have been removed.

stick to a "total ignorance" prior (see Section III-C) in this article.

We have also added the Allan Deviation (ADEV), i.e., the square root of the Allan Variance, of the measurement noise in Fig. 5. It is prevailing below 10 s and we can see that the confidence intervals of the most stable clock in this range, B and C, have a larger confidence interval (e.g., C at 1 s) than for $\tau = 100$ s for instance. At 1 s, the measurement noise is approximately five times higher than the stability of C, and despite its huge EDF number (EDF = 370 000 at 1 s), the effect of the measurement noise affects clearly its confidence interval as in Fig. 4. In other words, the measurement noise is too high to be fully rejected by GCov.

V. Conclusion

The method we propose, KLTS, provides the CDFs of the clock noise variances and the corresponding confidence intervals. Unlike the direct GCov estimates, the medians of these intervals cannot be negative.

KLTS is a rigorous method since it involves neither approximations nor simplifying assumptions. It is based on the property of sufficient statistics that form the estimates.

KLTS is valid even for a very low EDF of 1. Massive Monte Carlo simulations have perfectly validated KLTS in different contexts: confidence interval versus the EDF, versus the stability disparity of the clocks, and versus the level of the measurement noise.

However, KLTS suffers from the drawback of not being easily computable for high EDF (> 100). However, it has been proven in this article that the KLTG method, presented in a previous article [5], is perfectly reliable above 100 EDF even if the measurement noise is strong. The combination of these two methods provides then a powerful tool to assess the confidence intervals for the clock noise variances, whatever the EDF and the measurement noise level (see [11] for a software solution).

Appendix
Glossary of Symbols

A. Measurements

$\bar{y}_{ABk}$	kth frequency deviation sample between clock A and clock B.
$\bar{y}_{1k}$	kth frequency deviation sample at the output of counter 1.
$\bar{y}_{N1k}$	kth frequency deviation sample of the intrinsic measurement noise of counter 1.
z_{ABk}	Difference between the two consecutive clock frequency deviation samples $\bar{y}_{ABk+1}$ and $\bar{y}_{ABk}$ divided by $\sqrt{2}$.
z_{1k}	Difference between the two consecutive counter frequency deviation samples $\bar{y}_{1k+1}$ and $\bar{y}_{1k}$ divided by $\sqrt{2}$.
z_{N1k}	Difference between the two consecutive noise frequency deviation samples $\bar{y}_{N1k+1}$ and $\bar{y}_{N1k}$ divided by $\sqrt{2}$.
$z_{cls,k}$	Sum of the kth z measurements of all counters (closure relationship).

B. Parameters and Estimates

The list below gives the parameters, i.e., the mathematical expectation we want to assess. The same symbols with a hat $\hat{\cdot}$ stand for the estimates of these parameters, i.e., the values computed from a finite number of measurements.

σ^2_{AB}	AVAR of the comparison between clock A and clock B.
σ^2_1	AVAR of the output of counter 1 (*elementary estimate*).
σ^2_{N1}	AVAR of the intrinsic measurement noise of counter 1. Since the noise level is assumed to be the same for all counters, the subscript 1, 2 or 3 may be omitted.
σ^2_{cls}	AVAR of the closure.
σ^2_A	AVAR of clock A (σ^2_A is called *model parameter* whereas $\hat{\sigma}^2_A$ is called *final estimate*).
TCH_A	Three-cornered hat computation of AVAR for clock A.
GCov_A	Groslambert covariance computation of AVAR for clock A.

C. Bayesian Formalism

$p\left(\vec{X}\middle\vert\vec{\Theta}\right)$	Conditional PDF of the multidimensional measurement $\vec{X}$ knowing the multidimensional parameter $\vec{\Theta}$ (model world).
$p\left(\vec{\Theta}\middle\vert\vec{X}\right)$	Conditional PDF of the multidimensional parameter $\vec{\Theta}$ knowing the multidimensional measurement $\vec{X}$ (experimenter world).
$\pi\left(\vec{\Theta}\right)$	*A priori* PDF of the multidimensional parameter $\vec{\Theta}$ before any measurement (prior).

D. KLTS Method

$\begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix}$	Eigenvectors of the rotation matrix which diagonalizes the covariance matrix of the measurements (K.L. transform).

$\vec{\Theta} = (V_1, V_2, V_3)^T$	KLT variances, i.e., eigenvalues of the covariance matrix of the measurements.
$\vec{W} = (\mathsf{w}_{1k}, \mathsf{w}_{2k}, \mathsf{w}_{3k})^T$	Image of the measurements $(\mathsf{z}_{1k}, \mathsf{z}_{2k}, \mathsf{z}_{3k})^T$ by the KLT.
M	Number of measurements.
N	Number of Equivalent degrees of freedom.

REFERENCES

[1] J. E. Gray and D. W. Allan, "A method for estimating the frequency stability of an individual oscillator," in *Proc. IEEE 28th Annu. Symp. Freq. Control*, May 1974, pp. 243–246.

[2] J. Groslambert, D. Fest, M. Olivier, and J. J. Gagnepain, "Characterization of frequency fluctuations by crosscorrelations and by using three or more oscillators," in *Proc. IEEE IFCS*, May 1981, pp. 458–463.

[3] C. R. Ekstrom and P. A. Koppang, "Error bars for three-cornered hats," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 53, no. 5, pp. 876–879, May 2006.

[4] F. Vernotte, C. E. Calosso, and E. Rubiola, "Three-cornered hat versus Allan covariance," in *Proc. IEEE IFCS*, May 2016, pp. 1–6.

[5] F. Vernotte and E. Lantz, "Three-cornered hat and Groslambert covariance: A first attempt to assess the uncertainty domains," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 3, pp. 643–653, Mar. 2019.

[6] A. Premoli and P. Tavella, "A revisited three-cornered hat method for estimating frequency standard instability," *IEEE Trans. Instrum. Meas.*, vol. 42, no. 1, pp. 7–13, Feb. 1993.

[7] A. S. Kholevo, *Encyclopedia of Mathematics*, B. V. Kluwer. New York, NY, USA: Academic, 2001. [Online]. Available: https://www.encyclopediaofmath.org/index.php/Sufficient_statistic

[8] G. Saporta, *Probabilités Analyse des Données et Statistiques*. Paris, France: Editions Technip, 1990.

[9] C. Greenhall and W. Riley, "Uncertainty of stability variances based on finite differences," in *Proc. 35th Annu. Precise Time Time Interval Meeting*, San Diego, CA, USA, Dec. 2003, pp. 267–280.

[10] C. E. Calosso, F. Vernotte, V. Giordano, C. Fluhr, B. Dubois, and E. Rubiola, "Frequency stability measurement of cryogenic sapphire oscillators with a multichannel tracking DDS and the two-sample covariance," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 3, pp. 616–623, Mar. 2019.

[11] F. Vernotte, A. Kinali, F. Meyer, and C. Plantard. (2019). *Sigmatheta Software*. [Online]. Available: https://sourcesup.renater.fr/www/sigmatheta/

Éric Lantz worked in the field of solar energy and heat pumps from 1979 to 1985. In 1986, he joined the Optics Laboratory of Besançon, now a part of the FEMTO ST Institute, Besançon, France. He is currently a Professor with the Université de Franche-Comté, Besançon. He initiated the work about travelling-wave parametric image amplification in crystals that has evolved in recent years to the study of spatial fluctuations of quantum origin. His current research interests also include the study of oscillator stability.

Claudio Eligio Calosso was born Asti, Italy, in 1973. He received the Ph.D. degree in communication and electronic engineering from the Polytechnic of Turin, Turin, Italy, in 2002.

In 2002, he joined IEN. He is currently a permanent Researcher at INRIM, Turin, where he develops low noise digital electronics for time and frequency applications. His activities include primary frequency standards, vapor cells clocks, frequency dissemination over fiber links, phasemeters, frequency division and synthesis, and recently, real-time time scale generation. He is also interested in signal analysis, with particular attention to the role of aliasing in time interval counters and two-sample variances.

Enrico Rubiola, photograph and biography not available at the time of publication.

Vincent Giordano was born in Besançon, France, in 1962. He received the Engineer degree (five years degree) in mechanics from the Ecole Superieure de Mécanique et des Microtechniques, Besançon, France, in 1984, and the Ph.D. degree in physical sciences from the Paris XI University, Orsay, France, in 1987.

From 1984 to 1993, he was a Researcher of the permament staff with the Laboratoire de l'Horloge Atomique, Orsay, where he worked on a laser diode optically pumped cesium beam frequency standard. In 1993, he joined the Laboratoire de Physique et de Métrologie des Oscillateurs (LPMO), Besançon. In January 2004, FEMTO-ST was founded from the merger of five different laboratories active in different fields of engineering science: mechanics, optics and telecommunications, electronics, time-frequency, energetics, and fluidics. In 2006, the Department Time and Frequency was created, grouping all the activities related to time and frequency metrology and to micro-acoustics components and systems. His current research interests include ultra-high stability microwave oscillators based on sapphire resonators, microwave and optical atomic clocks, and time and frequency metrology.

Christophe Fluhr was born in Strasbourg, France, in 1988. He received the master's degree in microsystem engineering with embedded electronics specialization from Franche-Comté University, Besançon, France, in 2012, and the Ph.D. degree in engineering science from Bourgogne Franche-Comté University, Besançon, in 2017.

During his thesis, he developed a new generation of cryogenic sapphire oscillator (CSO) based on a low power consumption cryostat. He is currently an Electrical Engineer at Femto Engineering with the Time and Frequency Department, Femto-ST Institute. His main task is manufacturing and commercialization of CSO.

Benoît Dubois, photograph and biography not available at the time of publication.

François Vernotte received the Ph.D. degree in engineering sciences from the University of Franche-Comté (UFC), Besançon, France, in 1991.

He has been with the Time and Frequency Team, Observatory THETA/UTINAM, UFC, since 1989. His favorite tools are statistical data processing (time series and spectral analysis), parameter estimation (inverse problem and Bayesian statistics), and simulation (Monte Carlo). His current research interests include long-term stability of oscillators, such as atomic clocks and millisecond pulsars.