

Detection and Identification of Phase and Frequency Drifts in Clock Ensembles

Christian Trainotti and Gabriele Giorgi

Institute of Communications and Navigation, German Aerospace Center (DLR), Germany

Christian Trainotti received a M.Sc. in mechanical engineering from the Technical University of Munich, Germany and a M.Sc. in mechatronic engineering from the University of Trento, Italy. Since 2018 he works at the Institute of Communications and Navigation at the German Aerospace Center (DLR) on precise and robust time generation for GNSS applications.

Gabriele Giorgi is a researcher at the Institute of Communications and Navigation, German Aerospace Center (DLR). He obtained his doctorate following his work on Global Navigation Satellite System (GNSS) for aerospace applications from the Delft Institute of Earth Observation and Space Systems (DEOS), Delft University of Technology, the Netherlands. His main research focuses on satellite navigation, visual navigation and multi-sensor fusion.

Abstract

A stable system time scale is a fundamental part of Global Navigation Satellite Systems (GNSSs). The composite clock approach provides a method for the generation of a stable and robust time scale. However, it can still be impacted by failures in single clocks. This paper describes two dedicated statistical tests aiming to detect and identify slow deviations from the clocks' nominal behavior. First, the batch test is described, which computes the overlapping Allan deviation (OADEV) of differential clock measurements over a sliding window and compare it to a nominal model. Second, the envelope test is developed. It compares the measured phase deviations with the expected value at each time step. The tests are derived for an arbitrary size of the ensemble, while their performance are evaluated with simulations and on a real testbed using three cesium frequency standards. Two fault scenarios are analyzed. The first case involves a linearly increasing frequency bias, while the second includes an oscillating frequency behavior, which intends to simulate a bias with period signature. The fault detection capabilities of the statistical tests are given in terms of time-to-detect and minimum detectable biases.

1 INTRODUCTION

A stable system time scale is a fundamental part of Global Navigation Satellite Systems (GNSSs). A degradation of the time scale directly impacts the quality of positioning and time dissemination services offered by the system. The continuity of these services is vital, making the robustness of time generation a paramount aspect that has to be ensured and monitored. The composite clock approach provides a method for the generation of a time scale that is inherently stable and robust. Instead of relying on a single clock signal, the measurements of a set of clocks are used. A Kalman filter processes these measurements, allowing the generation of a weighted average, the so-called Implicit Ensemble Mean (IEM) [1].

While being inherently robust, the IEM realization can still be impacted by failures in single clocks. By employing Fault Detection and Identification (FDI) techniques, dedicated statistical tests can be constructed, aiming to detect and identify these faults. For instance, it is possible to detect phase jumps in the signal of one clock by means of a snapshot test, which checks for blunders in the predicted-minus-observed phase residuals of the Kalman filter. Slow deviations from the clocks' nominal behavior such as phase drifts can be revealed by a batch test. This makes use of a sliding window to compute the overlapping Allan deviation (OADEV) [2] of the signals and compare it to a nominal model. These methods were described and tested in a previous work by the authors [3]. The aim of this work is to extend the analyses therein by focusing on the fault detection and identification of different types of clock drifts. Furthermore, an additional method for drift detection is implemented, which compares the measured phase deviation with the expected value at each time step.

In order to test the detection algorithm, two specific drift scenarios are examined. The first case includes a linear frequency drift (quadratic phase drift) followed by a constant frequency bias (linear phase drift) in one clock. Second, an oscillating frequency behavior is assumed. The latter case intends to simulate periodic influences in the clocks, such as orbital effects on satellites' clocks or effects that correlate with daily temperature fluctuations. By using the aforementioned two scenarios as examples, the performance of the proposed FDI methods is analyzed as function of various settings and parameters, such as the length of the sliding window and the sampling interval at which the OADEV is computed. The Minimum Detectable Biases (MDB) are then derived for each of the tests developed. The MDBs describe the smallest fault magnitude that the detector is able to reveal, given the integrity requirements set on a particular scenario. In this way, it is possible to estimate the relation between the number of clocks in the ensemble, the MDBs, and the integrity demands for different fault cases.

The fault scenarios are reproduced in simulation and in hardware in the DLR's TimeLab with a set of three cesium frequency references. A dedicated ensembling algorithm combines the measurements of the clocks and it includes the FDI tests. A given phase profile can be injected in the signal of one of the clocks by means of a High Resolution Offset Generator (HROG), so that the test algorithms can be assessed with a controlled fault pattern. In parallel, the same ensemble setup and fault scenario are simulated in software, providing baseline results of a tightly controlled environment to compare with hardware results.

In [4] the fault detection in clock ensemble was initiated. Other works include [5,6] using Kalman filter and [7,8] without filters. The Generalized Likelihood Ratio Test (GLRT) was approached in [9,10].

2 CLOCK ENSEMBLE AND CLOCK MODELS

The system under analysis is an ensemble of $n + 1$ clocks of the same type. The relative phase differences between the units are measured by comparing the signals of the clocks. These measurements can be used in a Kalman filter to estimate the states of each clock. The IEM [1] is a weighted average of the estimation error of each clock, which is more stable than any clock in the ensemble. However, this is only a paper clock, since the original clocks' states are not observable. It is possible to have a physical realization of the IEM by using a second Kalman filter and a regulator. By means of a dedicated control strategy, the signal of a frequency reference can be steered to the IEM. For more details regarding the steering and realization of the IEM, the reader can refer to [11, 12, 13, 14, 15, 16, 17, 18] and references therein.

The ensembling algorithm runs the Kalman filters and computes the necessary steering command. The latter is sent to an HROG that acts as steering device. The HROG applies the phase offset to the signal of an Oven Controlled Oscillator (OCXO). The output signal of the HROG represents the realized IEM of the ensemble. In the experimental setup, every signal in the system is additionally monitored by a measurement device referenced to an Active Hydrogen Maser (AHM), in order to obtain a controlled set of reference measurements.

The algorithm is based on a 2-state clock model with constant frequency drift [11,13]. It describes the clock behaviour with a set of three parameters: σ_1 refers to the phase noise, σ_2 to the frequency noise and d describes the clock's constant frequency drift. A number of Markov processes can be added to the basic clock model in order to improve it, each one being determined by two parameters, U_i and R_i . If K Markov processes are added, the expected Overlapping Allan variance can be expressed as:

$$\sigma_A^2(\tau) = \frac{\sigma_1^2}{\tau} + \frac{\sigma_1^2}{3}\tau + \frac{\tau^2}{2}d^2 + \sum_{i=1}^K U_i \frac{-3 + 4 \exp(-R_i\tau) - \exp(-2R_i\tau) - 2R_i\tau}{R_i^2\tau^2}. \quad (1)$$

Two sets of parameters for the cesium frequency references are used in this paper: the first is an overbounding model of the OADEV of the cesium clocks and it used in the Kalman filter and for the snapshot test (not described here, see [3]). The second is a best-fit description of the cesium clocks' OADEV and it is used in the tests for drift detection. The parameters are the following:

$$\begin{array}{l} \text{Overbounded : } \sigma_1^2 = 8 \cdot 10^{-25} \text{ s} \quad \sigma_2^2 = 0 \quad d = 0 \quad U_1 = 8.05 \cdot 10^{-23} \quad R_1 = 1.113 \text{ s}^{-1} \\ \text{Best-fit : } \sigma_1^2 = 8 \cdot 10^{-25} \text{ s} \quad \sigma_2^2 = 0 \quad d = 0 \quad U_1 = 5.05 \cdot 10^{-23} \quad R_1 = 0.262 \text{ s}^{-1} \end{array} \quad (2)$$

3 DETECTION OF CLOCK DRIFTS IN AN HOMOGENEOUS ENSEMBLE

The phase and frequency drifts that are analyzed here are slow-growing biases that remain undetected by snapshot statistical tests. Since the magnitudes of these biases remain below the set threshold for detection of phase or frequency

jumps when inspecting the update step of the Kalman filter, it becomes necessary to accumulate measurements over time. Statistical tests on data batches can be derived in different ways, depending on the metric used to monitor the ensemble. One approach is to investigate the OADEV of the accumulated measurements over time, a technique that also reveals the type of drift detected but it is insensitive to linear phase drifts (constant frequency offsets). A second option is to impose an upper bound on the accumulated relative clock measurements, which should remain below a predefined threshold set by pre-analyzing the wandering behavior of the clocks over time. These two approaches are based on the same concept: if the time-behavior model of the clocks, i.e. their stability over different time samples, is known a-priori, any departure from said model can be statistically assessed when a sufficient number of data is accumulated. The next sections separately discuss these two approaches.

3.1 Detection And Identification of Clock Drifts Through Inspection of OADEV - Batch Test

Let us consider a batch of differential phase observations between an ensemble of $n + 1$ clocks:

$$\mathbf{Z}_{(k-s):k} = [\mathbf{z}_{k-s}, \mathbf{z}_{k-s+1}, \dots, \mathbf{z}_{k-1}, \mathbf{z}_k] . \quad (3)$$

Matrix $\mathbf{Z}_{(k-s):k}$, of dimensions $n \times (s + 1)$, contains the $s + 1$ vectors of differential phase measurements taken between the $n + 1$ clocks in the ensemble from epoch $k - s$ to epoch k . Each row of matrix $\mathbf{Z}_{(k-s):k}$ contains the time series of the differential phase measurements between two clocks in the ensemble over the window of $s + 1$ epochs. Computing the OADEV for a sampling interval τ of each row, we obtain the n -element vector of OADEV values for each series:

$$\boldsymbol{\sigma}(\tau) = (\sigma_1(\tau), \dots, \sigma_n(\tau))^\top . \quad (4)$$

Note that, in order for the test to be reliable, the maximum sample interval that can be investigated should not exceed $\tau_{\max} < \frac{s\Delta_t}{10}$, with Δ_t the time interval between measurements (in seconds).

If a model of the clock stability is available, its nominal OADEV at each sampling interval τ is known. The functional and stochastic models for the measured OADEV are

$$\mathcal{E}(\boldsymbol{\sigma}(\tau)) = \mathbf{u}_n \sigma_M(\tau) \quad ; \quad \mathcal{D}(\boldsymbol{\sigma}(\tau)) = \mathbf{Q}_M(\tau) = \mathbf{H}^\top \mathbf{H} \xi(\tau) , \quad (5)$$

where:

- $\mathcal{E}(\cdot)$ and $\mathcal{D}(\cdot)$ denote the expectation and dispersion operators, respectively;
- \mathbf{u}_n is a n -element vector of ones;
- $\sigma_M(\tau)$ is the nominal OADEV of the measurements between clocks for a sampling interval τ . If the nominal OADEV for the type of clock used is denoted with $\sigma_A(\tau)$, the OADEV of measurements between any two clocks of the same type is obtained as $\sigma_M(\tau) = \sqrt{2} \sigma_A(\tau)$, when assuming no additional noise is introduced by the measuring instruments. $\sigma_A(\tau)$ can be obtained from (1), once the clock model parameters are known;
- $\xi(\tau)$ is the variance of the measured OADEV of a single clock. This is obtained by analyzing the distribution of the OADEV obtained over several different runs for the same sampling interval τ (see Appendix);
- \mathbf{H} is the measurement matrix, describing how the clock are measured in respect to each other. In this case, since the first clock is used as a reference, the matrix is

$$\mathbf{H} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} .$$

The correct determination of the stochastic model in (5) is the key to build a reliable test. A direct approach may attempt to determine the real-time OADEV of all clocks in the ensemble by processing the whole amount of data up to the current epoch, and comparing the result with the reference value. This causes both detectability and latency problems: the computed OADEV would start departing from the nominal values after a time proportional to the ratio between the epochs affected by the drift and the epochs in which the drift was not present. For example, it was shown in [3] that a frequency drift of magnitude $3 \cdot 10^{-14} \text{ Hz s}^{-1}$ would only trigger the test after $6 \cdot 10^4 \text{ s}$ from its onset, occurring after an error-free interval of 10^5 seconds has elapsed. In order to reduce the latency in detecting a clock drift, multiple statistical tests are applied on sliding windows that only include a portion of past data points. Model

(5) is then adapted by introducing an additional parameter: the width of the sliding temporal window applied to the data in run-time, denoted with T_w :

$$\mathcal{E}(\boldsymbol{\sigma}(\tau, T_w)) = \mathbf{u}_n \sigma_M(\tau) \quad ; \quad \mathcal{D}(\boldsymbol{\sigma}(\tau, T_w)) = \mathbf{Q}_M(\tau, T_w) = \mathbf{H}^T \mathbf{H} \xi(\tau, T_w) . \quad (6)$$

While the expected value of the OADEV does not change, the nominal covariance of the measured OADEV, $\xi(\tau, T_w)$, is now dependent on the length of the data batch being processed. This reference variance value needs then to be obtained for varying lengths of the sliding window T_w , in order to apply a statistical test to the correct expected distribution of the OADEV measurement error. Appendix A details how the reference values $\xi(\tau, T_w)$ are obtained. Following from model (6), we are now capable of testing at current epoch k the residual vector

$$\boldsymbol{\rho}_k(\tau, T_w) = \boldsymbol{\sigma}(\tau, T_w) - \mathbf{u}_n \sigma_M(\tau) . \quad (7)$$

Vector $\boldsymbol{\rho}_k$ distributes normally with a covariance matrix equal to $\mathbf{Q}_M(\tau, T_w)$. The expected value is null for error-free data, and non-null for data departing from the nominal model. Two working hypotheses can now be outlined: a null hypothesis \mathcal{H}_0 which postulate drift-free data, and an alternative hypothesis \mathcal{H}_a in which the presence of a certain drift is assumed. These read

$$\begin{aligned} \mathcal{H}_0 : \quad & \boldsymbol{\rho}_k(\tau, T_w) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_M(\tau, T_w)) \\ \mathcal{H}_a : \quad & \boldsymbol{\rho}_k(\tau, T_w) \sim \mathcal{N}(\mathbf{C}\boldsymbol{\nabla}, \mathbf{Q}_M(\tau, T_w)) \end{aligned} . \quad (8)$$

Under the null hypothesis, the residuals vector distributes as a zero-mean Gaussian with covariance matrix $\mathbf{Q}_M(\tau, T_w)$: this is the nominal expected behavior. Under the alternative hypothesis, the expected mean of the residual vector is non-null, and it takes the value $\mathbf{C}\boldsymbol{\nabla}$: a q -vector of model errors $\boldsymbol{\nabla}$ maps through a $n \times q$ matrix \mathbf{C} into the residual vector $\boldsymbol{\rho}_k(\tau, T_w)$. The product $\mathbf{C}\boldsymbol{\nabla}$ is used to capture changes in the measured quantities – the OADEV differential measurements between clocks in the ensemble. Matrix \mathbf{C} is assumed to have full rank.

The ratio between the probability density functions (PDFs) associated to the null- and alternative hypotheses defines the Generalized Likelihood Ratio (GLR) test [19]:

$$\text{Reject } \mathcal{H}_0 \text{ if } e^{-\frac{1}{2}(\boldsymbol{\rho}_k^T(\tau, T_w) \mathbf{Q}_M^{-1}(\tau, T_w) \boldsymbol{\rho}_k(\tau, T_w) - (\boldsymbol{\rho}_k(\tau, T_w) - \mathbf{C}\hat{\boldsymbol{\nabla}})^T \mathbf{Q}_M^{-1}(\tau, T_w) (\boldsymbol{\rho}_k(\tau, T_w) - \mathbf{C}\hat{\boldsymbol{\nabla}}))} < \alpha , \quad (9)$$

where

$$\hat{\boldsymbol{\nabla}} = (\mathbf{C}^T \mathbf{Q}_M^{-1}(\tau, T_w) \mathbf{C})^{-1} \mathbf{C}^T \mathbf{Q}_M^{-1}(\tau, T_w) \boldsymbol{\rho}_k(\tau, T_w) . \quad (10)$$

For later use, the covariance matrix of $\hat{\boldsymbol{\nabla}}$ is

$$\mathbf{Q}_{\hat{\boldsymbol{\nabla}}\hat{\boldsymbol{\nabla}}}(\tau, T_w) = (\mathbf{C}^T \mathbf{Q}_M^{-1}(\tau, T_w) \mathbf{C})^{-1} . \quad (11)$$

Using $\hat{\mathbf{e}}_0 = \boldsymbol{\rho}$, $\hat{\mathbf{e}}_a = \boldsymbol{\rho} - \mathbf{C}\hat{\boldsymbol{\nabla}}$, and $k_a = 2\ln\frac{1}{\alpha}$, the test can be rewritten in compact form as

$$\text{Reject } \mathcal{H}_0 \text{ if } T = \hat{\mathbf{e}}_0^T \mathbf{Q}_M^{-1}(\tau, T_w) \hat{\mathbf{e}}_0 - \hat{\mathbf{e}}_a^T \mathbf{Q}_M^{-1}(\tau, T_w) \hat{\mathbf{e}}_a > k_\alpha \quad (12)$$

Scalar k_α is the selected threshold, and its value defines the statistical properties of the test. In order to investigate these properties, an equivalent expression for test T can be derived. Using (10)-(11) and the related normal equation $\mathbf{C}^T \mathbf{Q}_M^{-1}(\tau, T_w) \mathbf{C}\hat{\boldsymbol{\nabla}} = \mathbf{C}^T \mathbf{Q}_M^{-1}(\tau, T_w) \boldsymbol{\rho}$, the test (12) can be reworked as

$$\text{Reject } \mathcal{H}_0 \text{ if } T = \hat{\boldsymbol{\nabla}}^T \mathbf{Q}_{\hat{\boldsymbol{\nabla}}\hat{\boldsymbol{\nabla}}}^{-1}(\tau, T_w) \hat{\boldsymbol{\nabla}} > k_\alpha . \quad (13)$$

The test T is then χ^2 -distributed with q degrees of freedom (the dimension of the error vector $\boldsymbol{\nabla}$) and central under \mathcal{H}_0 , non-central under \mathcal{H}_a , with non-centrality parameter $\lambda = \boldsymbol{\nabla}^T \mathbf{Q}_{\hat{\boldsymbol{\nabla}}\hat{\boldsymbol{\nabla}}}^{-1}(\tau, T_w) \boldsymbol{\nabla}$:

$$\begin{aligned} \mathcal{H}_0 : \quad & T \sim \chi^2(q, 0) \\ \mathcal{H}_a : \quad & T \sim \chi^2(q, \lambda) \end{aligned} . \quad (14)$$

The test can then be designed to detect whether the set of measurements is affected by a bias (detection), or to identify which clock in the ensemble is drifting (identification).

3.1.1 Detection of clock drift(s)

By applying test (13) with $q = n$, no redundancy is left in the alternative hypothesis. The estimate $\widehat{\mathbf{V}} = \mathbf{C}^{-1}\boldsymbol{\rho}$ absorbs all residuals, and by using $\mathbf{C} = \mathbf{I}_n$ the Overall Model (OM) test is obtained:

$$\text{Reject } \mathcal{H}_0 \text{ if } T = \boldsymbol{\rho}^T \mathbf{Q}_M^{-1}(\tau, T_w) \boldsymbol{\rho} > k_\alpha . \quad (15)$$

The following estimates for the error vector, its covariance matrix and the non-centrality parameter are derived:

$$\begin{aligned} \widehat{\mathbf{V}} &= \boldsymbol{\rho} \\ \mathbf{Q}_{\widehat{\mathbf{V}}\widehat{\mathbf{V}}} &= \mathbf{Q}_M(\tau, T_w) \\ \lambda &= \nabla^T \mathbf{Q}_M^{-1}(\tau, T_w) \nabla . \end{aligned} \quad (16)$$

For a given threshold k_α , the probability of false alarm can be assigned immediately, and viceversa. The probability of missed detection is a function of the expected bias.

3.1.2 Identification of clock affected by a drift

The matrix \mathbf{C} is set to $\mathbf{C} = \mathbf{c}_i$, with \mathbf{c}_i a n -element vector with i -th entry equal to one, and zero elsewhere. By varying i from 1 to n , all elements of the residual vector $\boldsymbol{\rho}_k(\tau, T_w)$ are tested as

$$\text{Reject } \mathcal{H}_0 \text{ if } T = \frac{\widehat{\mathbf{V}}^2}{\sigma_{\widehat{\mathbf{V}}}^2} > k_\alpha . \quad (17)$$

This is the w-test used to ‘data-snoop’ for biases in the residual vector. Note that both $\widehat{\mathbf{V}} = \frac{\mathbf{c}_i^T \mathbf{Q}_M^{-1}(\tau, T_w) \boldsymbol{\rho}_k(\tau, T_w)}{\mathbf{c}_i^T \mathbf{Q}_M^{-1}(\tau, T_w) \mathbf{c}_i}$ and $\sigma_{\widehat{\mathbf{V}}}^{-2} = \mathbf{c}_i^T \mathbf{Q}_M^{-1}(\tau, T_w) \mathbf{c}_i$ are now scalars. The test is χ^2 -distributed with 1 degree of freedom, and central under \mathcal{H}_0 , non-central under \mathcal{H}_a , with non-centrality parameter $\lambda = \nabla^2 \mathbf{c}_i^T \mathbf{Q}_M^{-1}(\tau, T_w) \mathbf{c}_i$. The minimum detectable bias (MDB) associated to a given value of probability of missed detection can be then obtained:

$$|\nabla_{\text{MDB}}| = \sqrt{\frac{\lambda}{\mathbf{c}_i^T \mathbf{Q}_M^{-1}(\tau, T_w) \mathbf{c}_i}} . \quad (18)$$

Fig. 1 shows the computed values of the MDBs as function of different parameters. The MDBs decrease when the ensemble gets larger, since more redundancy is introduced into the system. The MDBs increase if the probability of missed detection P_{md} is decreased or when a smaller probability of false alarm P_{fa} is set.

3.2 Detection And Identification of Clock Drifts Through Inspection of Differential Phase Measurements - Envelope Test

The tests outlined in the previous section enable detecting and identifying clocks drifts that originate from non-constant frequency drifts. The OADEV is invariant with respect to a constant clock frequency offset. This section discusses a second set of statistical tests designed to detect any remaining phase drift that exceeds the expected nominal model. Under the reference clock models given in Section 2, the differential phase measurement between any two clocks has a random-walk behavior, with a slow-growing envelope with a slope proportional to the clock stability. Fig. 2 shows an example of 10^3 s long measurements generated between 1000 simulated cesium clocks whose phases are modeled according to (1) and (2). By sampling the distribution of these measurements at given time instances, the histograms below are obtained, which can be approximated to Gaussian distributions with same null mean and different variance. Plotted with the dashed red line, the standard deviation of these distributions vary as function of time: $\sigma_d(t_k)$. Using the clock model previously introduced, this function is well approximated by the following polynomial [20, 21, 22]:

$$\sigma_d^2(t) = 2 \left[\sigma_1^2 t + \sigma_2^2 \frac{t^3}{2} + t^3 \sigma_M^2 a(R_1 t) \right] , \quad (19)$$

where

$$\sigma_M^2 = 2U_1 R_1 \quad , \quad a(x) = \frac{-\frac{3}{2} + x + 2 \exp(-x) - \frac{\exp(-2x)}{2}}{x^3} .$$

The factor 2 is included since the measurements consist of differences between the same type of clocks.

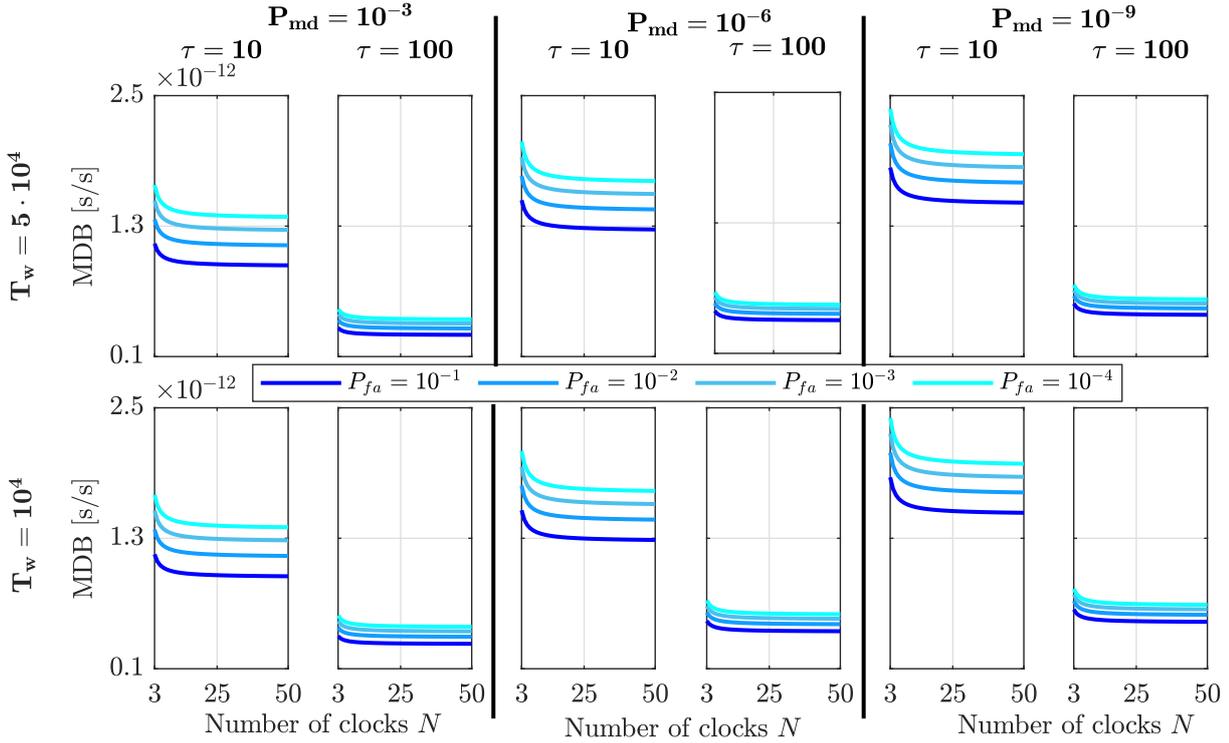


Figure 1: Minimum detectable biases (MDBs) for the batch test (15) as function of the number of clocks in the ensemble N , for different values of probability of false alarm P_{fa} , probability of missed detection P_{md} , length of sliding window T_w and sampling interval τ . The scales of all subplots' axes are the same.

Based on the clock statistical behavior, the following functional and stochastic models capturing the expected temporal behavior of the n differential measurements between clocks' phases $\varphi(t_k)$ can be formulated:

$$\mathcal{E}(\varphi(t_k)) = \mathbf{0} \quad , \quad \mathcal{D}(\varphi(t_k)) = \sigma_d^2(t_k) \mathbf{I}_n . \quad (20)$$

Following the same reasoning of the previous section, the following two hypotheses are defined:

$$\begin{aligned} \mathcal{H}_0 : \quad & \varphi(t_k) \sim \mathcal{N}(\mathbf{0}, \sigma_d^2(t_k) \mathbf{I}_n) \\ \mathcal{H}_a : \quad & \varphi(t_k) \sim \mathcal{N}(\mathbf{C}\nabla, \sigma_d^2(t_k) \mathbf{I}_n) . \end{aligned} \quad (21)$$

The measurement vector $\varphi(t_k)$ distributes as a zero-mean Gaussian with covariance matrix $\sigma_d(t_k) \mathbf{I}_n$ under the null hypothesis, whereas the expected values of $\varphi(t_k)$ are $\mathbf{C}\nabla$ under the alternative hypothesis. The GLR test is then

$$\text{Reject } \mathcal{H}_0 \text{ if } \quad T = \frac{1}{\sigma_d^2(t_k)} \varphi^T(t_k) \mathbf{C} (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \varphi(t_k) > k_\alpha . \quad (22)$$

The test distributes as in (14): a χ^2 distribution with q degrees of freedom, central under \mathcal{H}_0 , non-central under \mathcal{H}_a , with non-centrality parameter $\lambda = \frac{1}{\sigma_d^2(t_k)} \nabla^T (\mathbf{C}^T \mathbf{C}) \nabla$:

$$\begin{aligned} \mathcal{H}_0 : \quad & T \sim \chi^2(q, 0) \\ \mathcal{H}_a : \quad & T \sim \chi^2(q, \lambda) . \end{aligned} \quad (23)$$

3.2.1 Detection of clock drift(s)

By applying test (22) with $q = n$ and $\mathbf{C} = \mathbf{I}_n$, the Overall Model (OM) test is obtained:

$$\text{Reject } \mathcal{H}_0 \text{ if } \quad T = \frac{1}{\sigma_d^2(t_k)} \varphi^T(t_k) \varphi(t_k) > k_\alpha . \quad (24)$$

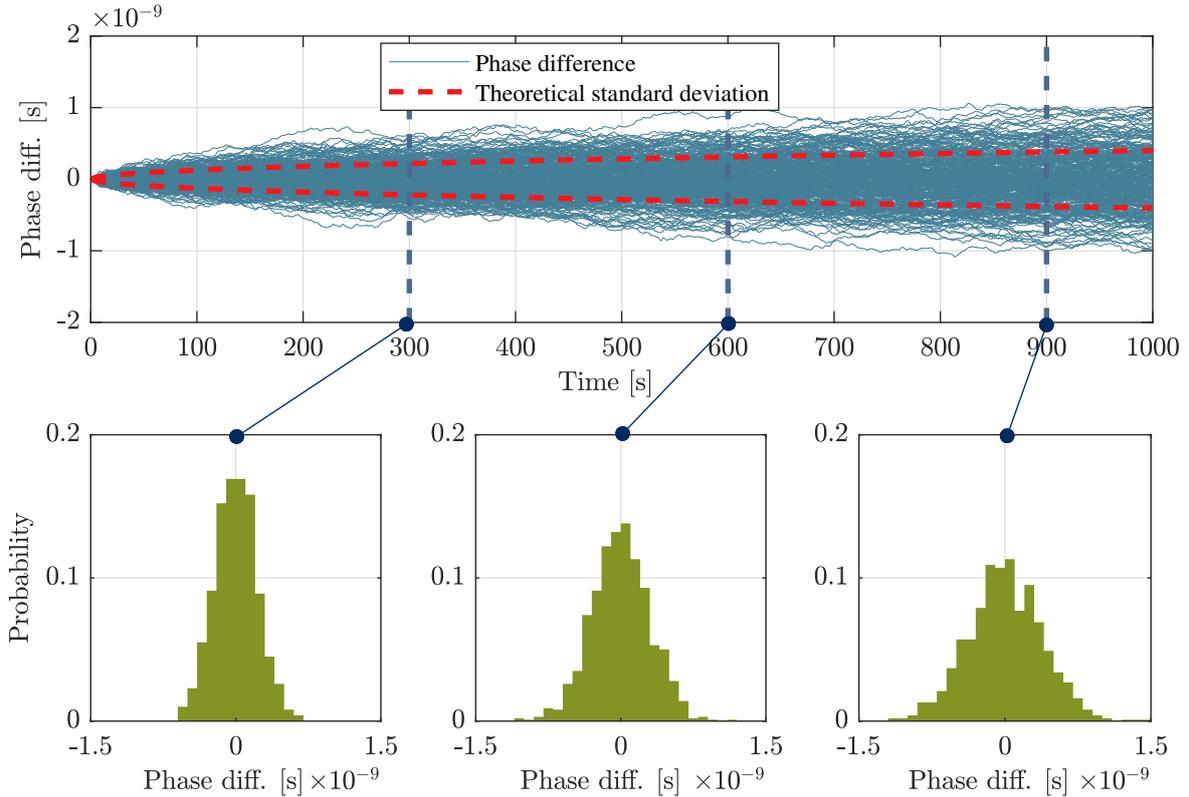


Figure 2: The distribution of the phase differences between clocks changes over time.

This is a simple test on the magnitude of the residuals. Note that the magnitude of T is scaled by the expected measurement covariance $\sigma_d^2(t_k)$, which absorbs the inherent ‘spread’ of the measurements resulting from the random walk process characterizing any differential clock phase signal. The non-centrality parameter is in this case $\lambda = \frac{1}{\sigma_d^2(t_k)} \nabla^T \nabla$. For a given threshold k_α , the probability of false alarm can be assigned immediately, and viceversa. The probability of missed detection is a function of the expected bias.

3.2.2 Identification of clock affected by a drift

All elements of the measurement vector $\varphi_i(t_k)$ are tested by setting $\mathbf{C} = \mathbf{c}_i$, resulting in the following w-test:

$$\text{Reject } \mathcal{H}_0 \text{ if } T = \frac{\varphi_i^2(t_k)}{\sigma_d^2(t_k)} > k_\alpha . \quad (25)$$

The minimum detectable bias (MDB) associated to a given value of probability of missed detection is now

$$|\nabla_{\text{MDB}}| = \sigma_d(t_k) \sqrt{\lambda} . \quad (26)$$

Fig. 3 show the computed values of the MDBs as function of time. The MDBs increase for decreasing probability of missed detection.

4 EXAMPLES OF APPLICATION AND RESULTS

The detection and identification algorithms are tested by means of two experiments, in which an additional drift is injected in the signal of the second cesium clock (labelled as ‘Cesium 1’). The first case involves the injection of a linearly increasing frequency drift, while the second introduces an oscillatory phase component.

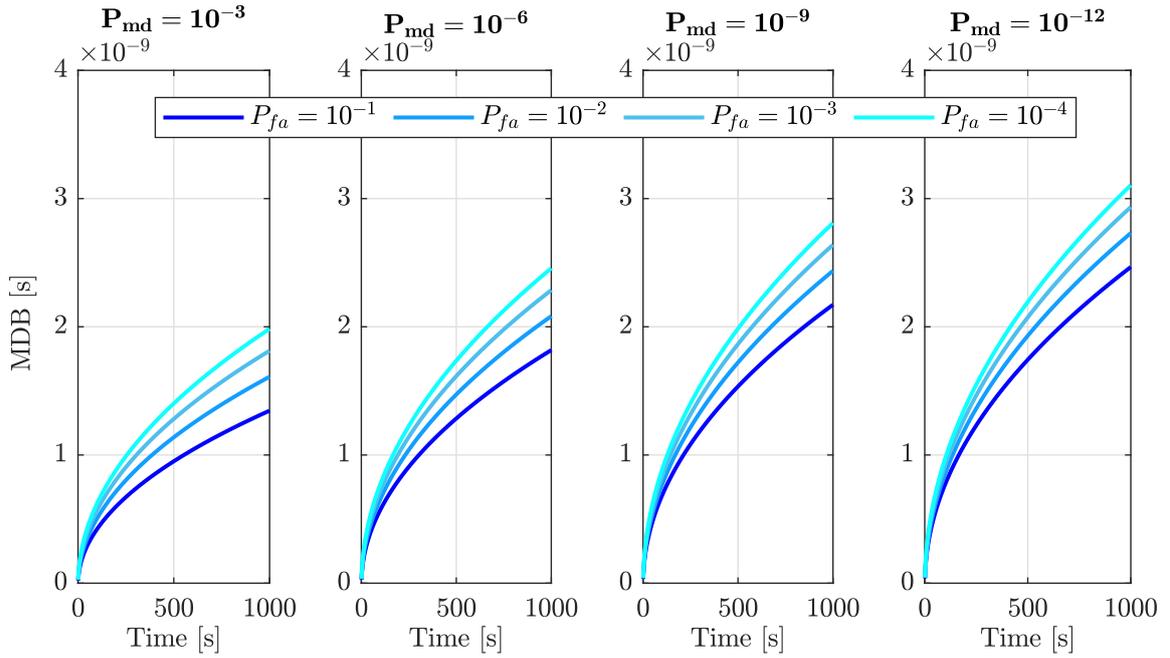


Figure 3: Minimum detectable biases (MDBs) for the envelope test as function of time, for different values of probability of false alarm P_{fa} and probability of missed detection P_{md} .

4.1 Example 1: Linear Frequency Drift

The first scenario is depicted in Fig. 4 and can be divided in three parts: in the first 10^5 s, no drift is injected in order to evaluate the tests under nominal conditions. Then, the frequency deviation of the Cesium 1 is slowly increased within the following 10^5 s to reach the final value of 10^{-10} , corresponding to 1 mHz with a nominal frequency of 10 MHz (Fig. 4b). The phase deviation in Fig. 4a shows a quadratic phase deviation, followed by a linear behavior once the frequency drift stops increasing. The effects of the injected drift can be also seen by plotting the OADEV computed on the entire length (Fig. 5): the drift has an effect on the long term behavior and it also influences the realized IEM. This example includes both linear and quadratic phase drift: while both behaviors can be detected by the envelope test of Section 3.2, the batch test is not triggered by linear phase drifts. The expected result of the batch test is therefore a test value increasing during the quadratic phase and then gradually decreasing when a constant frequency drift (linear phase drift) is reached. The drift is expected to be detected for sample intervals τ larger than about 10^3 s, as can be deduced from Fig. 5.

The results of the overall model batch test applied in this scenario are given in Fig. 6. The batch test was run with increasing window's length and various values of τ have been checked. With a longer window it is possible to test longer values of sampling intervals τ . For $\tau = 10$ s the test is triggered only when $T_w = 10^3$ s, however these detections are false alarms, due to the uncertainty in the computation of the OADEV over short batch windows. The activation of the test is compatible with the value of the false alarm probability ($P_{fa} = 10^{-3}$). For $\tau = 100$ s, the test also shows some triggerings at short length windows, but no effects due to the added drift can be identified yet. The drift pattern is clearly detected with $\tau = 10^3$ s and $\tau = 10^4$ s. The simulated results show generally lower values of the test T , probably due to the presence of a higher measurement noise in the experimental environment. In the real measurements, the detection delay is about 2 hours for $T_w = 5 \cdot 10^4$ s, but only 30 min for $T_w = 5 \cdot 10^3$ s.

For the case of $T_w = 5 \cdot 10^4$ s and $\tau = 10^4$ s, the corresponding w-test is shown in Fig. 7, left. The purpose is to identify which clock is showing the detected drift, as described in Section 3.1.2. The highest value of the test is found for $\mathbf{c}_1 = [1, 0]^T$, meaning that the first measurement is the most probable candidate as faulty clock. However, also the other case activates the test, due to the correlation between the differential measurements given by the common reference clock. In order to resolve this ambiguity, the measurement showing the highest test value is excluded and the overall model test is repeated with the reduced measurement set. The result is shown in Fig. 7, right. Since now the test is not activated, it can be correctly concluded that the drifting clock is Cesium 1.

By applying the envelope test on the same measurements (Section 3.2), the results in Figure 8 are obtained. The simulated values show the expected behavior: the drift is detected about $5 \cdot 10^3$ s after the injection. The drifting

clock is correctly identified since only the test on φ_1 is triggered (plots on the right). The hardware measurements show however a different behavior. The value of the tests grows linearly in time, and the overall model test is triggered even before the onset of the injected drift. This linear growth can be clearly seen also in the w-test in the lower right plot, which should not be influenced by the additional injected drift. This kind of residual drift is probably due to a constant frequency bias in the measurements that is not correctly modelled. This bias can not be seen in the batch test, since the OADEV is insensitive to constant frequency biases. If the bias is removed from the phase measurements by linear detrending, the test in hardware is now detecting only the injected drift (dashed blue line). This example shows the ability of the envelope test to detect linear phase drifts that are not modelled or foreseen, which instead can not be detected by the batch test on the OADEV.

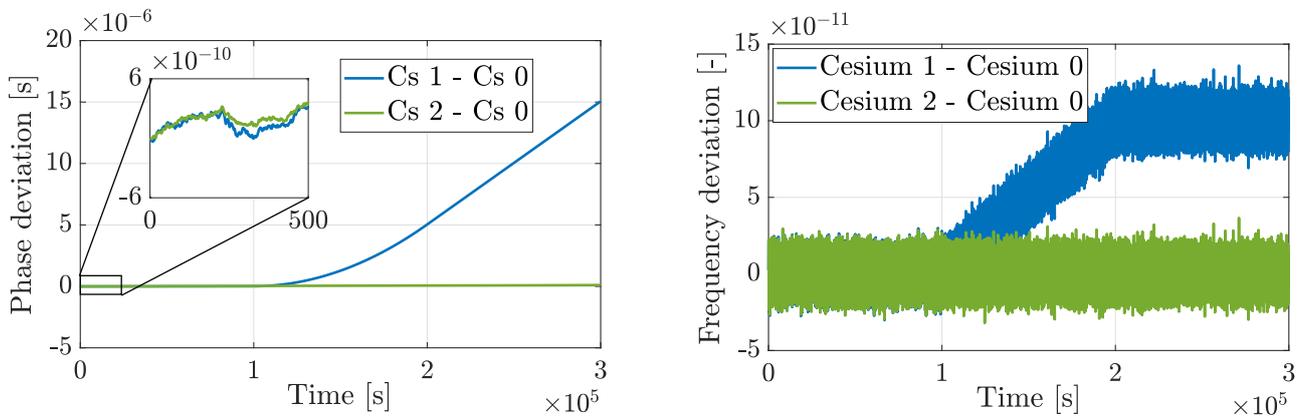
4.2 Example 2: Phase Oscillation

This scenario aims to study the effects of an oscillatory component in one of the ensemble's clocks. This effect can be caused for instance by periodic environmental changes during an orbit or daily effects in a ground lab. An oscillation with period of 90 minutes is introduced after 10^5 s with a linearly increasing amplitude. Once the amplitude reaches a value of $5^\circ \approx 1.4 \cdot 10^{-9}$ s after 10^5 s, the magnitude increase is stopped and the amplitude is kept constant. The resulting phase difference measurements and the applied offset are shown in Fig. 9a and 9b, respectively. In the OADEV plot of Fig. 10, the bumps caused by the oscillation can be seen both in the clock signal and in the IEM realization.

The batch test for different window lengths results in the values shown in Fig. 11. For short sampling intervals the results are very similar to the previous example: for $\tau = 10$ s the hardware test shows a high noise level which decreases with longer τ and T_w values. In the case $\tau = 10^3$ s the drift is correctly detected. A clear increase in the value can be seen about $5 \cdot 10^4$ s \simeq 13 hours after the start of the drift injection. For $\tau = 10^4$ s the test as expected does not show a clear drift detection, since the bump due to the oscillation is located at shorter sampling intervals.

For the test at $T_w = 5 \cdot 10^4$ s and $\tau = 5 \cdot 10^3$ s, the corresponding w-test for fault identification is shown in Fig. 12, left. After the drift injection, the test value for $\mathbf{c}_1 = [1, 0]$ is clearly the highest, even if the test is triggered also in the other cases. In order to uniquely identify the source of the drift, the measurement most likely to be its source is excluded and an overall model test is repeated with the reduced set of measurements (Fig. 12, right). By excluding the measurement of Cesium 1, the test is no longer triggered, and the faulty unit can be correctly identified.

Regarding the envelope test on the phase differences, the results are shown in Fig. 13. Also in this example the test is triggered due to an unmodelled constant frequency bias. By detrending the measurements the test is not triggered any longer, suggesting that the injected oscillation remains undetected. The injected behaviour has a zero-mean long-term influence on the corresponding phase, and the envelope test would not be apt to detect such fault scenario.



(a) Phase deviation measurements of the Cesium 1 and 2 with respect to Cesium 0. (b) Fractional frequency deviation measurements of the Cesium 1 and 2 with respect to Cesium 0.

Figure 4: Example 1 - time evolution of the phase and frequency difference measurements of the clocks.

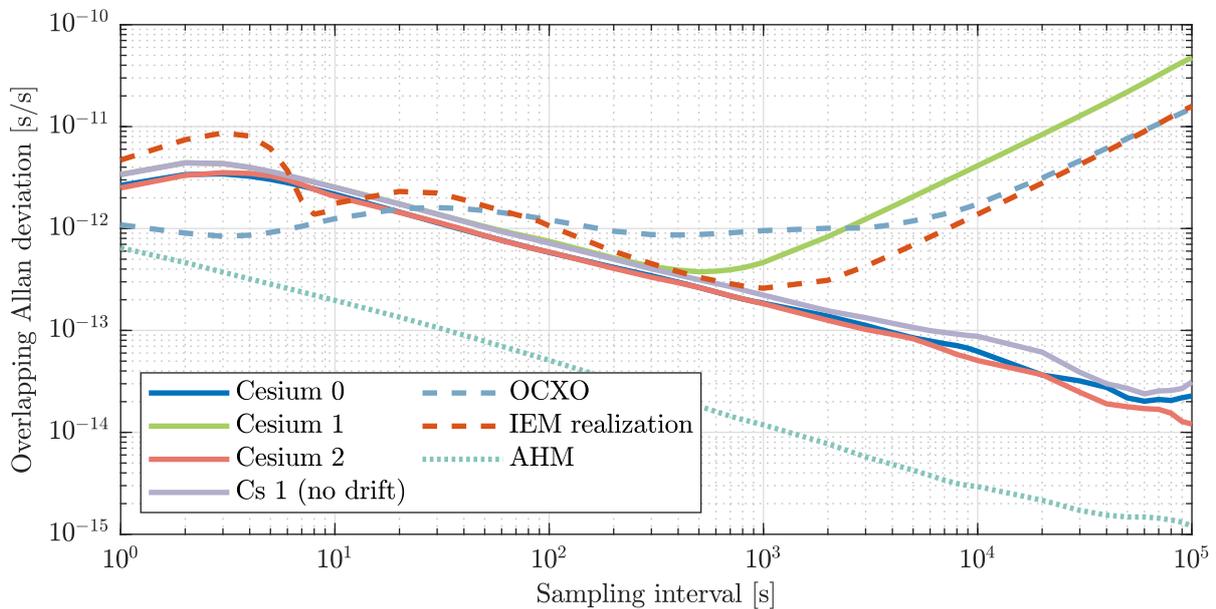


Figure 5: Example 1 - overlapping Allan deviation of the phase of the clocks, measured with respect to an active hydrogen maser (AHM), computed on the entire measurement length. The steering for IEM realization is applied every 20 s. The gain is computed with the pole placement method ($\lambda = 0.6$).

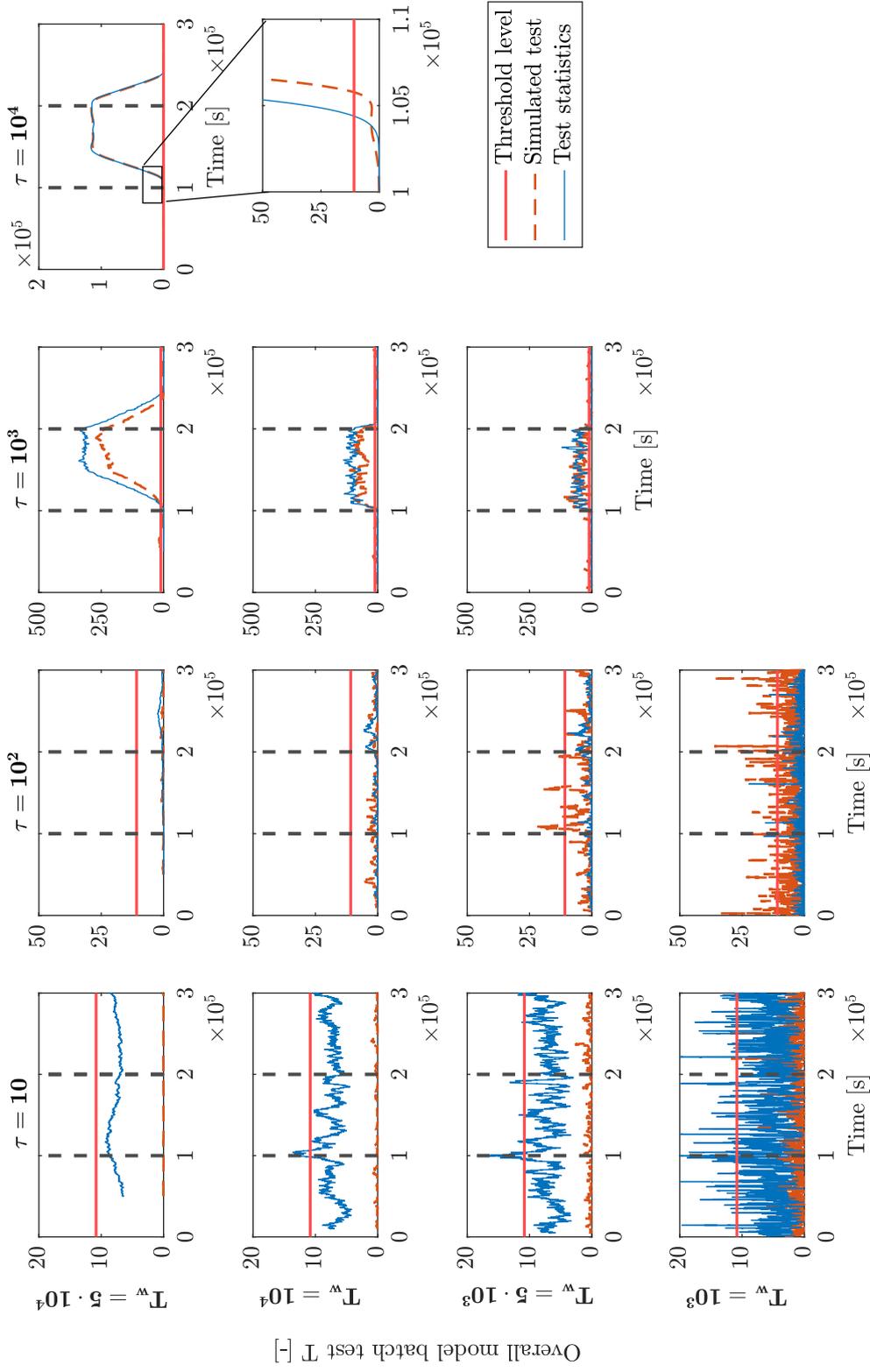


Figure 6: Example 1 - resulting test value for the overall model batch test, for different window lengths T_w and computed at different values of sampling interval τ . The blue solid lines show the test computed on real differential clock measurements, while the dashed orange ones a corresponding simulation.

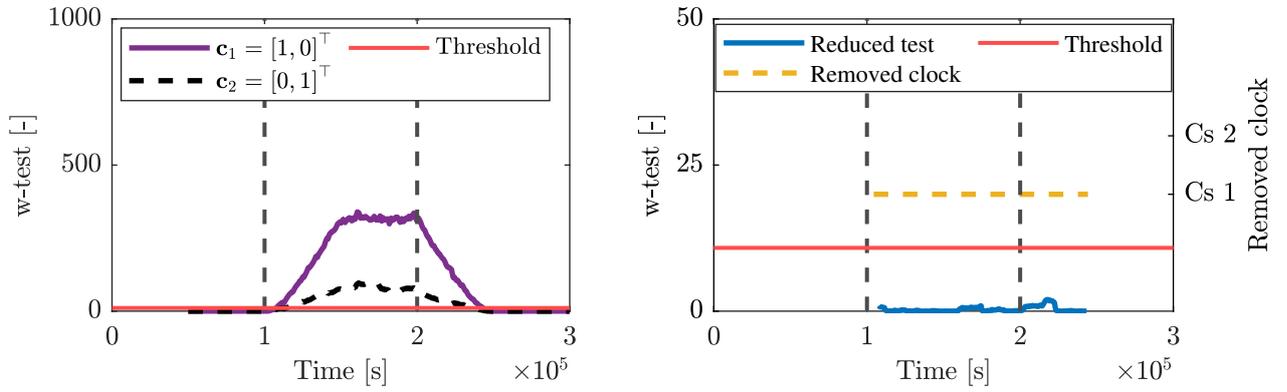


Figure 7: Example 1 - batch analysis on the hardware measurements, for $T_w = 5 \cdot 10^4$ s and $\tau = 10^4$ s. The left plot shows the w-test for the two different vectors c_i . The right plot shows the resulting overall model test, if the first measurement is removed.

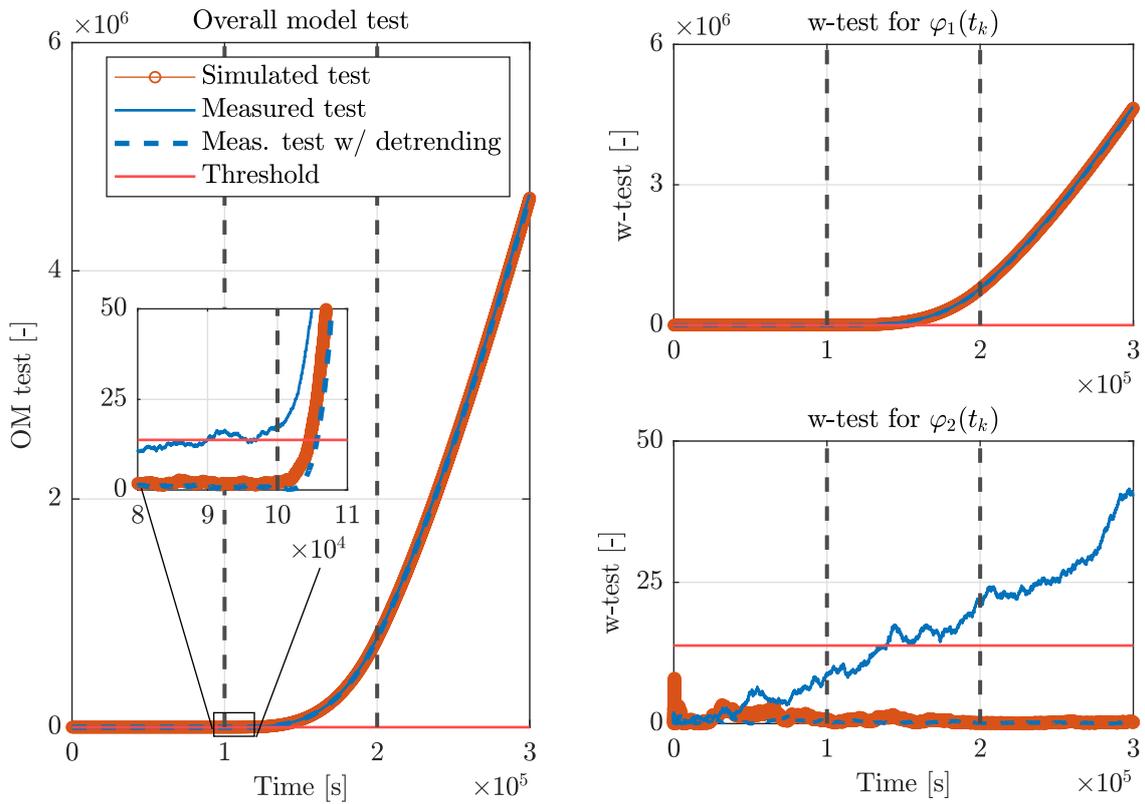
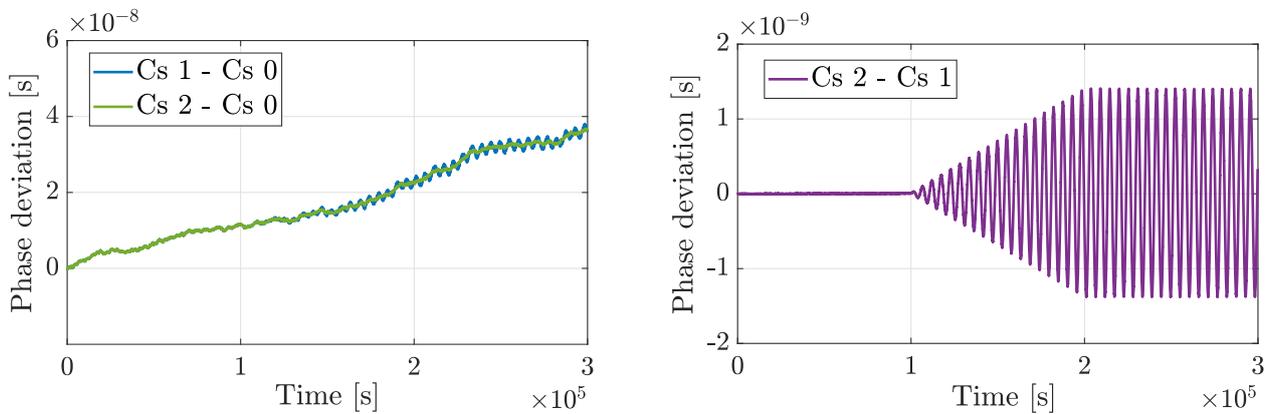


Figure 8: Example 1 - envelope test for the hardware measurement, detrended measurement and simulation. The left plot shows the overall model test, while the right plots show the two tests for fault identification.



(a) Phase deviation measurements of the Cesium 1 and 2 with respect to Cesium 0. (b) Additional phase deviation injected to Cesium 1.

Figure 9: Example 2 - phase difference measurements of the Cesium signal and injected additional phase component.

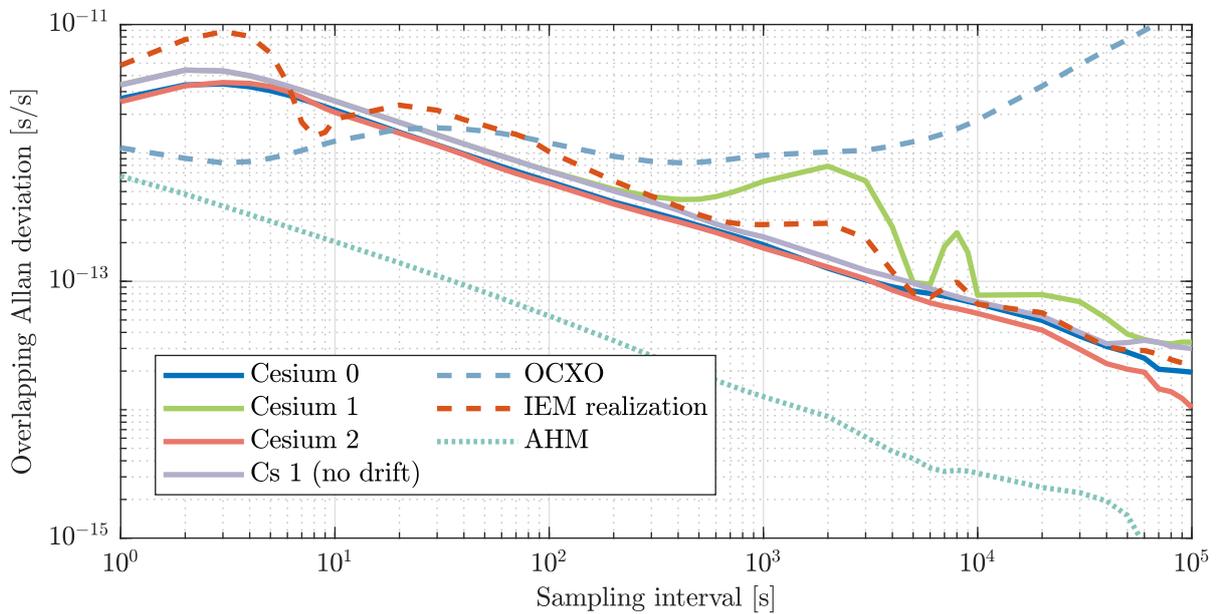


Figure 10: Example 2 - overlapping Allan deviation of the phase of the clocks, measured with respect to an active hydrogen maser (AHM), computed on the entire measurement length. The steering for IEM realization is applied every 20s. The gain is computed with the pole placement method ($\lambda = 0.6$).

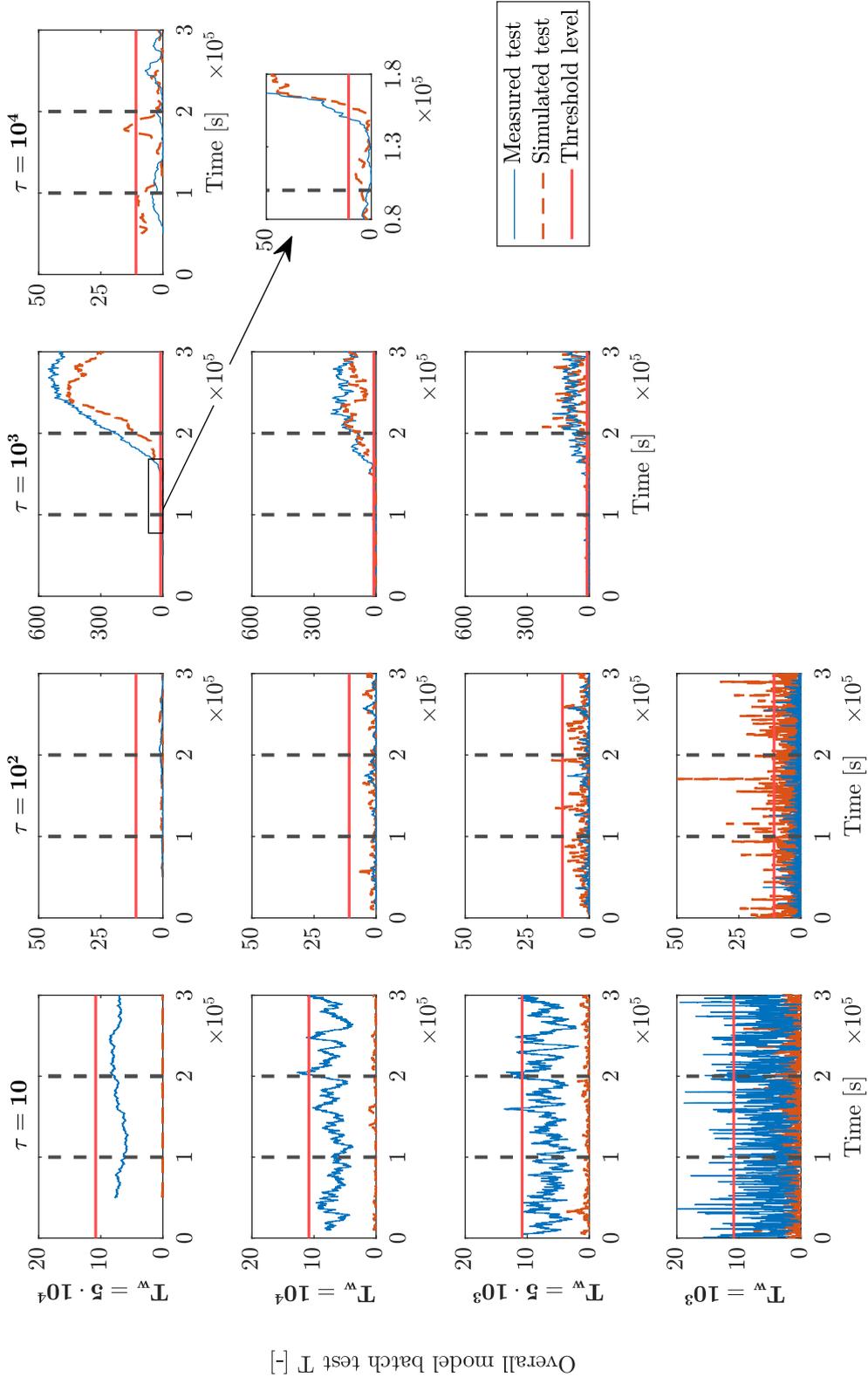


Figure 11: Example 2 - resulting test value for the overall model batch test, for different window lengths T_w and computed at different values of sampling interval τ . The blue solid lines show the test computed on real differential clock measurements, while the dashed orange ones a corresponding simulation.

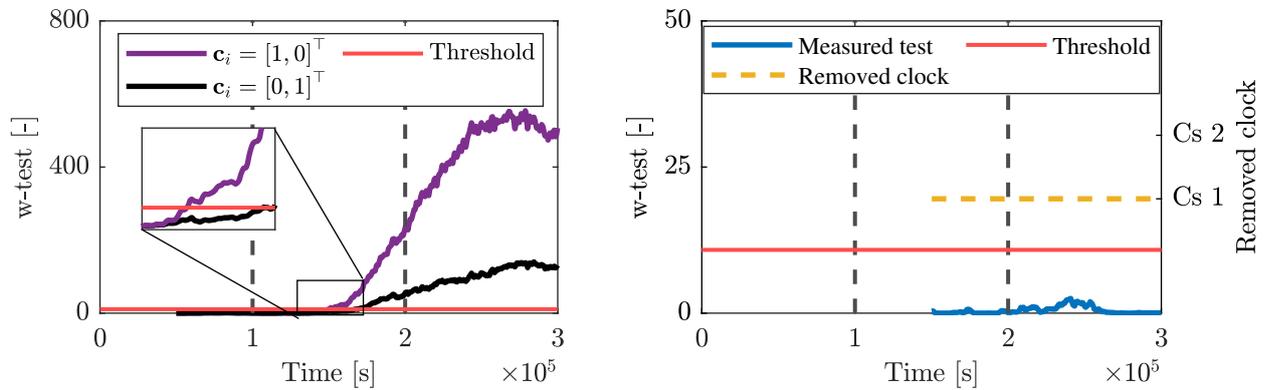


Figure 12: Example 2 - batch analysis on the hardware measurements, for $T_w = 5 \cdot 10^4$ s and $\tau = 10^3$ s. The left plot shows the w-test for the two different vectors \mathbf{c}_i . The right plot shows the resulting overall model test, if the first measurement is excluded.

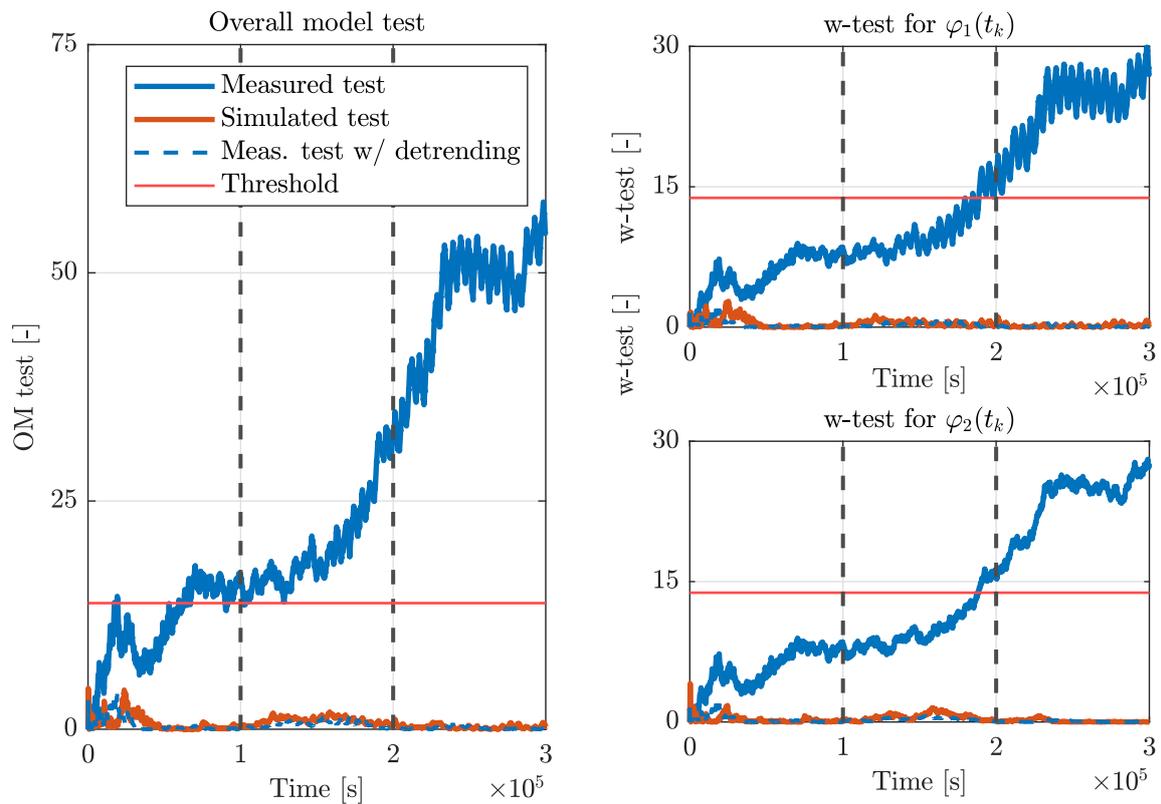


Figure 13: Example 2 - envelope test for the hardware measurement and simulation. The left plot shows the overall model test, while the right plots show the two tests for fault identification.

5 CONCLUSION AND OUTLOOK

This paper describes two techniques for the detection and identification of unmodelled, slowly changing behaviors in ensembles of clocks of the same type. The batch test makes use of a sliding window over which the OADEV of the clocks' measurements is computed. The envelope test instead monitors the distribution of the phase measurements at every time step. The two tests are described in a general way for an arbitrary number of clocks, while experimental and simulated results are obtained for a set of three cesium frequency references. For both test, the steps of detection and identification are derived, defined as the overall model test and w-test, respectively.

Two fault scenarios are used to verify the tests' performance. The results show the expected tests outcome, both in simulation and in hardware measurements. In the first case, a linear frequency drift is detected by both the batch test and envelope test, while a constant frequency bias remains undetected in the batch test. In the second case, the oscillatory component triggers the batch test, while being too small for a detection with the envelope test. From this point of view, the two tests are complementary. The unexpected linear behavior found in the experimental results of the envelope test revealed that an unmodelled frequency bias is present in the measurements. This shows the power of this test and also that an appropriate modelling and characterization of the clocks and measurements is paramount to ensure the best performances of the ensembling algorithms and of the statistical tests.

The tests described in this paper aim to detect slowly-growing faults. They complement the snapshot test described in [3], which is designed to detect abrupt changes such as jumps and steps in the clocks' signals. Altogether, this set of tests monitors the correct functioning of the clock ensemble against a number of fault modes. Future work is going to adapt these approaches to consider different measurement topologies (for instance when clocks are measured in a ring structure) and ensembles including clocks of different types (mixed ensembles).

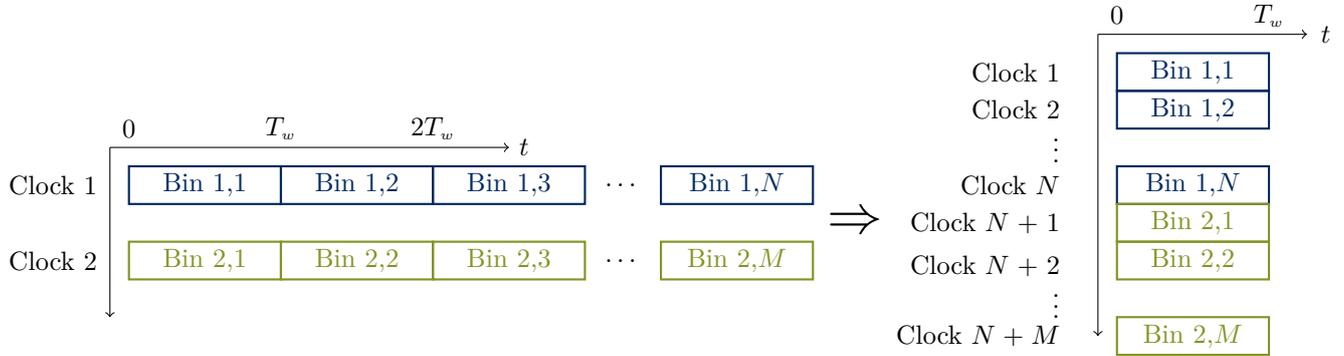
A APPENDIX

The analysis of a set of clock measurements allows the extraction of a nominal value for $\xi(\tau, T_w)$, the standard deviation of the measured OADEV for a sampling interval τ , computed over a time window of T_w seconds. Such analysis would require a large statistical population, but in DLR's TimeLab only three cesium references are present. Therefore, the population size is artificially increased by splitting the measurements of these clocks in bins and considering each bin as a measurement of an additional independent frequency reference. Figure 14a shows a scheme of the process. The measurements are split into bins of length T_w , corresponding to the batch length under investigation. The OADEV over each bin is computed and the values for all the bins are used to compute the standard deviation of the OADEV, as function of the batch length T_w and the sample interval τ .

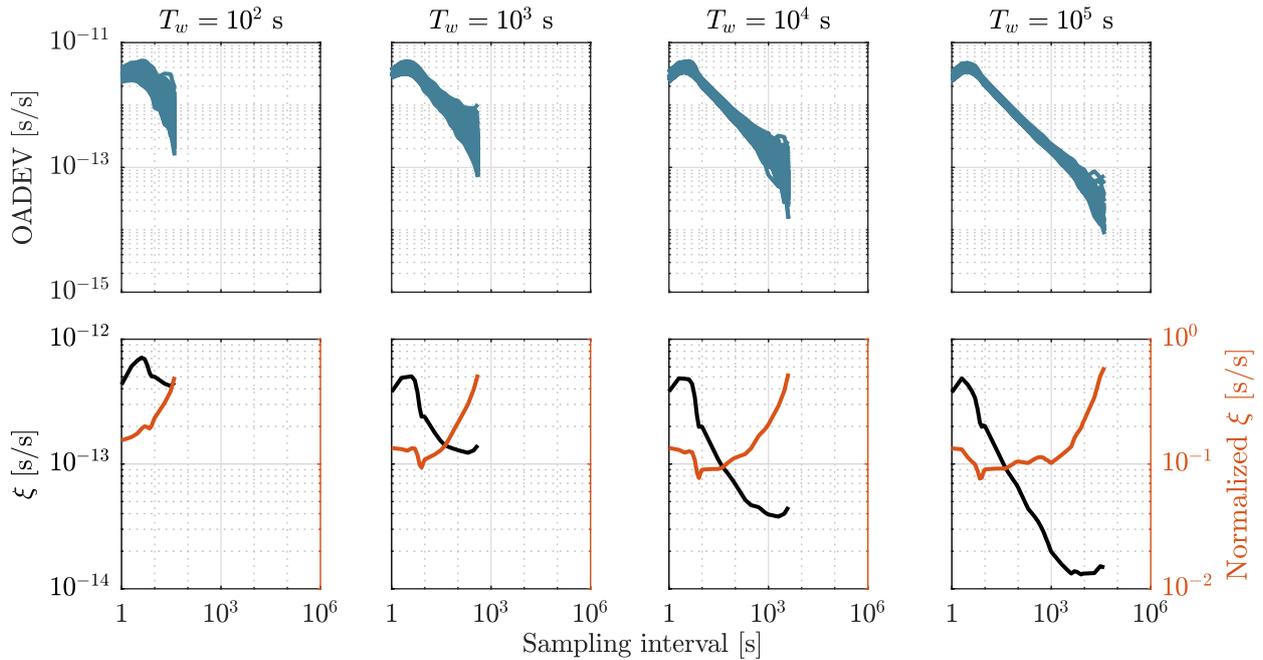
Starting from the measurements of three cesium clocks, with a joint measurement length of $6.4 \cdot 10^6$ s, the measurements are split in bins of length $T_w = 1 \cdot 10^{\{1,2,3,4,5\}}$ and $T_w = 5 \cdot 10^{\{1,2,3,4\}}$. The corresponding resulting values of $\xi(\tau, T_w)$ for each selection of T_w are stored in a file which is then read to retrieve the desired value when performing the test. Figure 14b shows a selection of results: in the upper plots the OADEV of every bin, in the lower ones the corresponding deviation.

ACKNOWLEDGMENTS

This work is supported by the Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V. under grant number ZT-0007 (ADVANTAGE, Advanced Technologies for Navigation and Geodesy).



(a) Estimation of the reference values for variance of the OADEV. Real clocks' measurements are split into bins, in order to increase the number of measurements available.



(b) OADEV of every measurement bin and corresponding standard deviation, for four different values of bin length T_w . The normalized OADEV standard deviation ξ is computed by dividing ξ by the corresponding expected values. For graphical reasons, only a reduced number of bins is plotted.

Figure 14: Derivation of the nominal value of the overlapping Allan deviation and its standard deviation, as function of the batch window length.

References

- [1] K. R. Brown Jr., “The theory of the GPS composite clock,” in *Proceedings of the 4th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS)*, (Albuquerque, NM, USA), pp. 223–242, ION, ION, 1991.
- [2] D. W. Allan, “Time and frequency (time-domain) characterization, estimation, and prediction of precision clocks and oscillators,” in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 34, pp. 647–654, Nov. 1987.
- [3] C. Trainotti, G. Giorgi, and J. Furthner, “Detection and Identification of Faults in Clock Ensembles,” in *Proceedings of the 32nd International Technical Meeting of the Satellite Division of the Institute of Navigation 2019 (ION GNSS+ 2019)*, (Miami, FL, USA), pp. 3342–3358, Sept. 2019.
- [4] R. H. Jones and P. V. Tryon, “Estimating time from atomic clocks,” *Journal of research of the National Bureau of Standards*, vol. 88, no. 1, pp. 17–24, 1983.
- [5] L. Galleani and P. Tavella, “Using the kalman filter to detect frequency jumps in atomic clocks,” in *2011 Joint Conference of the IEEE International Frequency Control and the European Frequency and Time Forum (FCS) Proceedings*, (San Francisco, CA, USA), pp. 1–3, May 2011.
- [6] A. Zenzinger, T. Bartusch, C. Kuehl, S. Fischer, and A. Shrestha, “Failure detection and correction for clock ensemble in space,” in *2012 6th ESA Workshop on Satellite Navigation Technologies (Navitec 2012) European Workshop on GNSS Signals and Signal Processing*, (Noordwijk, Netherlands), pp. 1–8, Dec 2012.
- [7] Q. Wang and P. Rochat, “An anomaly clock detection algorithm for a robust clock ensemble,” in *Proceedings of the 41st Annual Precise Time and Time Interval Systems and Applications Meeting*, (Santa Ana Pueblo, NM, USA), pp. 121 – 130, Nov. 2009.
- [8] Q. Wang, F. Droz, and P. Rochat, “Robust clock ensemble for time and frequency reference system,” in *2015 Joint Conference of the IEEE International Frequency Control Symposium the European Frequency and Time Forum*, (Denver, CO, USA), pp. 374–378, April 2015.
- [9] E. Nunzi and P. Carbone, “Monitoring signal integrity of atomic clocks by means of the glrt,” *Metrologia*, vol. 45, no. 6, pp. S103–S107, 2008.
- [10] E. Nunzi, L. Galleani, P. Tavella, and P. Carbone, “Detection of anomalies in the behavior of atomic clocks,” *IEEE Transactions on Instrumentation and Measurement*, vol. 56, pp. 523–528, April 2007.
- [11] M. Gödel and J. Furthner, “Robust ensemble time onboard a satellite,” in *Proceedings of the 48th Annual Precise Time and Time Interval Systems and Applications Meeting*, (Monterey, CA, USA.), pp. 26–43, ION, ION, Jan. 2017.
- [12] T. D. Schmidt, M. Gödel, and J. Furthner, “Investigation of pole placement technique for clock steering,” *Proceedings of the 49th Annual Precise Time and Time Interval Systems and Applications Meeting*, pp. 22–29, 2018.
- [13] C. Trainotti, T. Schmidt, and J. Furthner, “Simulating the realization of a mixed clock ensemble,” in *Proceedings of the Joint Conference of the IEEE International Frequency Control Symposium (IFCS) and European Frequency and Time Forum (EFTF)*, (Orlando, FL, USA), 2019.
- [14] P. Koppang, D. Johns, and J. Skinner, “Application of control theory in the formation of a timescale,” tech. rep., NAVAL OBSERVATORY WASHINGTON DC, 2004.
- [15] L. A. Breakiron and P. Koppang, “Frequency steering of hydrogen masers,” in *Frequency Control Symposium, 1996. 50th., Proceedings of the 1996 IEEE International.*, pp. 1113–1122, IEEE, 1996.
- [16] P. Koppang and R. Leland, “Linear quadratic stochastic control of atomic hydrogen masers,” *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 46, no. 3, pp. 517–522, 1999.
- [17] D. Matsakis, “The unification of the pole placement and linear quadratic gaussian techniques,” in *Proceedings of the 50th Annual Precise Time and Time Interval Systems and Applications Meeting*, (Reston, VA, USA), pp. 299–335, Institute of Navigation, feb 2019.

- [18] D. Matsakis, “The effects of proportional steering strategies on the behavior of controlled clocks,” *Metrologia*, vol. 56, p. 025007, mar 2019.
- [19] P. J. Teunissen, *Testing theory, an introduction*. Vereniging voor Studie- en Studentenbelangen te Delft (VSSD), 2006.
- [20] L. Galleani, L. Sacerdote, P. Tavella, and C. Zucca, “A mathematical model for the atomic clock error,” *Metrologia*, vol. 40, no. 3, p. S257, 2003.
- [21] C. Zucca and P. Tavella, “The clock model and its relationship with the allan and related variances,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 52, pp. 289–296, Feb. 2005.
- [22] G. Panfilo, P. Tavella, and C. Zucca, “How long does a clock error remain inside two threshold barriers? an evaluation by means of stochastic processes,” in *18th European Frequency and Time Forum (EFTF 2004)*, (Guildford, United Kingdom), pp. 110–115, IEEE, Apr. 2004.