On the Use of Quickest Detection Theory for Signal Integrity Monitoring in Single-Antenna GNSS Receivers

Daniel Egea-Roca, Gonzalo Seco-Granados, José A. López-Salcedo
Department of Telecommunications and Systems Engineering
Universitat Autònoma de Barcelona (UAB)
Bellaterra 08193, Barcelona (Spain)

***Abstract*—Local threats such as radio frequency interference, multipath and spoofing have attracted the attention of many researchers in the past years thus leading to a myriad of contributions in the field of threat detection. As a matter of fact, the current state of the art relies on classical detection techniques, which are not well suited for threat detection. In this paper, we take a leap forward in local threat detection, presenting a quickest detection framework for threat detection. The aim is to detect the presence of local threats as soon as possible in order to improve the integrity of GNSS receivers.**

I. Introduction

With the widespread deployment of Global Navigation Satellite Systems (GNSS), one of the major challenges to be solved is the provision of integrity to users beyond the civil aviation community, where this feature is already a well-established performance criterion. *Integrity* refers to the ability of the user receiver to guarantee the quality and trust of the received signal, in such a way that critical or commercial applications can be safely operated.

Position integrity is typically provided in civil aviation by Receiver Autonomous Integrity Monitoring (RAIM) algorithms and Satellite Based Augmentation Systems (SBAS). However, such methods assume that local effects like multipath, Non-Line-Of-Sight (NLOS) propagation, radio frequency interference and spoofing have a controlled influence on the signal [1], which is not the case in terrestrial environments. This is the reason why the analysis of signal integrity greatly contributes to the capability to provide PVT integrity, which is currently a concern within the GNSS community due to the widespread deployment of terrestrial GNSS receivers and the emergence of new GNSS-based services [2].

In order to improve integrity in terrestrial environments we have to detect local effects as soon as posible in order to quickly alert the user. These effects are well known problems within the GNSS community, leading to a plethora of contributions in the existing literature for interference [3], multipath [4] and spoofing [5] detection. However, all the existing contributions have been addressed adopting a classical detection framework, which is not well suited to fulfil the requirements of safety-critical applications.

For the prompt detection of integrity threats, it is essential to formulate the problem under the framework of *statistical change detection*, which is also often known as quickest detection. This framework is aimed at minimising the detection delay, and then it is the right way for proceeding in the problem under consideration. The quickest detection framework has been extensively studied in the past decades, being applied to automatic control since a long time ago [6], and more recently, to multi-antenna receivers for signal detection [7], and to spectrum sensing in cognitive radio [8]. Nevertheless, to the best of the authors' knowledge, quickest detection has not been applied yet to GNSS for signal-level integrity.

Based on this observation, the goal of this work is to bridge the evident gap between quickest detection theory and GNSS threat detection. We already addressed this problem for the case of multi-antenna GNSS receivers (see [9], [10]), where threat detection is simpler due to the availability of spatial information. In this paper, though, we focus instead on mass-market, and thus, single-antenna receivers. In this case, the problem is more general and complicated due to lack of spatial information, which dramatically reduces the number of available degrees of freedom to work with. Our contribution in this work is twofold: on the one hand, to extend the concept of quickest detection to single-antenna GNSS receivers; on the other hand, to introduce and stimulate the use of quickest detection within the GNSS community.

The rest of this paper is organised as follows. Section II introduces the signal model and Section III introduces the fundamentals of quickest detection. In Section IV we present the application of quickest detection to interference and multipath detection. Finally, Section V concludes the paper.

II. Signal Model

Let us consider a sequence of K observations $\mathbf{x} \doteq [x(0), x(1), \ldots, x(v), \ldots, x(K-1)]^T$, where v is the time instant at which an integrity threat appears (i.e. unknown change time). Consequently, it is assumed that before v (i.e. at hypothesis $\mathcal{H}_0$) the observation $x(n)$ follows a given statistical distribution, whereas after the change (i.e. at hypothesis $\mathcal{H}_1$) it follows a different one:

$$\begin{aligned} \mathcal{H}_0 &: x(n) \sim f_0(x(n)), \quad n < v \\ \mathcal{H}_1 &: x(n) \sim f_1(x(n)), \quad n \geq v \end{aligned} \tag{1}$$

Based on these premises, quickest detection aims at finding the strategy that minimises the detection delay, while keeping the mean time between false alarms larger than a conveniently set value. For this purpose, the CUSUM algorithm was proposed by Page based on a very important concept in statistics, namely the logarithm of the likelihood ratio:

$$\mathrm{LLR}(n) \doteq \ln \frac{f_1(x)}{f_0(x)} \tag{2}$$

and referred to as the log-likelihood ratio (LLR). For the sake of clarity we have omitted the time index n from the independent random variables x, keeping in mind that each random variable x corresponds to a given time instant (i.e. $x(n)$).

The key statistical property of this ratio is as follows: Let E_0 and E_1 denote the expectations of the random variables under the two distributions f_0 and f_1, respectively. Hence,

$$\mathrm{E}_0\left(\mathrm{LLR}\right) < 0 \text{ and } \mathrm{E}_1\left(\mathrm{LLR}\right) > 0 \tag{3}$$

or equivalently, *a change in the distribution of the random variables $x(n)$ is reflected as a change in the sign of the mean value of the log-likelihood ratio.* This is depicted in Fig. 1, which shows the cumulative log-likelihood (lower plot)

$$S(n) = \sum_{i=1}^{n} \mathrm{LLR}(i) \tag{4}$$

for the case of a Gaussian mean change (upper plot). Nevertheless, this behaviour is independent of the probability distribution.

This work was partly supported by the grant TEC2014-53656-R.

978-1-4799-9858-6/15/$31.00

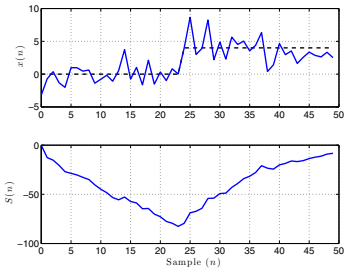

Fig. 1: Typical behaviour of the cumulative LLR $S(n)$ (bottom plot) corresponding to a mean change ($\mu_0 = 0$ and $\mu_1 = 4$) of a Gaussian sequence with constant variance ($\sigma^2 = 2.5$) (upper plot), for $v = 25$.

III. Fundamentals of Quickest Detection

A. CUSUM derivation

Next, we derive the CUSUM algorithm based upon a repeated used of the *sequential probability ratio test* (SPRT) [11]. The idea of Page was to repeatedly test the two simple hypotheses:

$$\begin{aligned} \mathcal{H}_0 &: f(x) = f_0(x) \\ \mathcal{H}_1 &: f(x) = f_1(x) \end{aligned} \tag{5}$$

with the aid of the SPRT. The SPRT is defined by the pair (d, T) where d is the decision rule and T is a stopping time defined as the time at which the final decision is taken. Formally, the SPRT is defined as

$$d = \begin{cases} 0 & \text{if } S(T) \leq -\epsilon \\ 1 & \text{if } S(T) \geq h \end{cases} \tag{6}$$

where T is the exit time:

$$T = T_{-\epsilon,h} = \min\left\{n : (S(n) \geq h) \cup (S(n) \leq -\epsilon)\right\} \tag{7}$$

with $S(n)$ defined in (4). The thresholds ϵ and h are conveniently chosen according to some probability of false alarm (α) and some probability of misdetection (γ). Then, the key idea of Page was to *restart the SPRT algorithm as soon as the previously taken decision was* $d = 0$. The first time at which $d = 1$, the SPRT is stopped, and this time becomes the *alarm time* at which the change is detected. To do so, Page suggested that the optimal value of the lower threshold ϵ should be zero (i.e. $\epsilon = 0$), resulting in the next decision rule:

$$g(n) = \begin{cases} g(n-1) + \mathrm{LLR}(n) & \text{if } g(n-1) + \mathrm{LLR}(n) > 0 \\ 0 & \text{if } g(n-1) + \mathrm{LLR}(n) \leq 0 \end{cases} \tag{8}$$

with $g(0) = 0$. This decision rule can be compacted into

$$g(n) = \left(g(n-1) + \mathrm{LLR}(n)\right)^+ \tag{9}$$

where $(x)^+ = \max(0, x)$. Finally, the stopping rule is defined by

$$t_a = \min\left\{n : g(n) \geq h\right\} \tag{10}$$

which is the well-known expression for the CUSUM algorithm.

B. Properties of the CUSUM algorithm

In this section, we describe several criteria for the performance evaluation of the CUUSM algorithm with the aim of providing a way for fixing the detection threshold h in order to fix a desired performance. For estimating the efficiency of the detection, it is convenient to use the mean delay for detection (i.e. $\overline{\tau}$) and the mean time between false alarms (i.e. $\overline{T}$). In fact, it would be interesting to have a specific function that contains all the information related to both values. This function is the average run length (ARL), which is defined as a function of the parameter θ, which represents the change of the distribution (i.e. $\mathcal{H}_0 : \theta = \theta_0$ and $\mathcal{H}_1 : \theta = \theta_1$):

$$L(\theta) \doteq \mathrm{E}_\theta[t_a] \tag{11}$$

with t_a the alarm time at which detection occurs. Thereby, the ARL function defines at θ_0 (i.e. at $\mathcal{H}_0$) the mean time between false alarms, and at θ_1 (i.e. at $\mathcal{H}_1$) the mean delay for detection.

The ARL function of the CUSUM can be obtained from the results of the SPRT performance analysis as follows [12]:

$$L(\theta) = \frac{\mathrm{E}_\theta\left[T_{-\epsilon,h}\right]}{1 - p(-\epsilon, h)} \tag{12}$$

with $\mathrm{E}_\theta\left[T_{-\epsilon,h}\right]$ and $p(-\epsilon, h)$ the average samples number (ASN) and operating characteristic (OC), respectively, of a SPRT with thresholds $-\epsilon$ and h. The ASN and OC are solutions of the Fredholm integral equation of the second kind, which has to be solved numerically. Thus, in order to avoid this numerical solution, and with the aim of making easier the design of the CUSUM algorithm, some approximations are available [12]. Usually, bounds on the ARL function are more desirable than approximations. This is because, in practice, it is important to fix a conveniently chosen performance of the change detection algorithm (given by the fixed threshold h) and then be sure that this performance will always be achieved within some limits. For this purpose, the use of bounds is mandatory to ensure that the desirable limits on the performance are always preserved.

A lower bound for the mean time between false alarms can be obtained from the Wald's inequalities [11]:

$$h \leq \ln\left(\frac{1-\gamma}{\alpha}\right) \quad \text{and} \quad -\epsilon \geq \ln\left(\frac{\gamma}{1-\alpha}\right) \tag{13}$$

For the CUSUM algorithm the lower threshold is equal to zero (i.e. $\epsilon = 0$). Regarding the misdetection probability (i.e. γ), *the CUSUM algorithm will always detect the presence of a change. It would take more or less time, but it will detect the change. Therefore the misdetection probability of the CUSUM detection is equal to zero (i.e.* $\gamma = 0$*).* With these considerations and taking into account the Wald inequality for the upper threshold we obtain the following lower bound for the mean time between false alarms, which can be defined as the inverse of the false alarm rate (i.e. $\overline{T} = 1/\alpha$, for $K \to \infty$):

$$\overline{T} = \mathrm{E}_0[t_a] \geq e^h. \tag{14}$$

On the other hand, we can obtain an upper bound for the mean detection delay from the result of the Wald's approximations [11]:

$$\overline{\tau} = \mathrm{E}_1[t_a] \approx \mathrm{K}\left(f_1, f_0\right)^{-1} \cdot \left((1-\gamma) \cdot h - \gamma \cdot \epsilon\right) \tag{15}$$

with $\mathrm{K}(f_1, f_0) \doteq \mathrm{E}_1\left[\mathrm{LLR}(n)\right]$ the Kullback divergence. Taking into account the previous considerations and the Wald's inequalities, we obtain the following bound:

$$\overline{\tau} = \mathrm{E}_1[t_a] \leq \frac{h}{\mathrm{K}(f_1, f_0)}. \tag{16}$$

Results in (14) and (16) provide us the asymptotic (i.e. when h goes to infinity) relationship between the mean time between false alarms and the detection delay with the fixed threshold. These results have been proved to be the optimal ones for the quickest detection problem. That is to say, that the optimality of the CUSUM algorithm is proved [11]. It is worth to mention that the optimality results are obtained in the case when the LLR is completely known. When the algorithm is used in situations where the actual distribution is different from the pre-assigned values, this optimality property cannot be guaranteed.

C. Unknown Log Likelihood Ratio

In this section, we discuss the case when the LLR is not completely known, that is the case when the parameters under $\mathcal{H}_1$ (i.e. after the change) are unknown. Without loss of generality, the parameters under $\mathcal{H}_0$ (i.e. before the change) are assumed to be known. With this interpretation, as it is explained in [12], one possibility is to replace the unknown parameters by its maximum likelihood estimate (i.e. GLR-CUSUM). However, in general, the GLR-CUSUM cannot be written in a recursive form since it depends upon the maximization over the unknown time change (i.e. we need all the collected samples). This gives rise to a big computational burden, thus making difficult to use the GLR-CUSUM in real-time applications. It is for this reason that we propose an alternative CUSUM-based approach in order to avoid the previous practical issue of the GLR-CUSUM.

In general, when the LLR is not completely known, it can be replaced by any other function of the observations $x(n)$. That is to say, $\rho(n) \doteq q(x(n))$, with some negative mean before the change and positive mean after the change (i.e. $\mathrm{E}_0\left[\rho(n)\right] < 0$ and $\mathrm{E}_1\left[\rho(n)\right] > 0$), in line with the key idea of the CUSUM (see (3)). That is,

$$g_{\mathrm{Offset}}(n) \doteq \left(g_{\mathrm{Offset}}(n-1) + \rho(n)\right)^+ . \tag{17}$$

In this case, the detection is no longer guaranteed to be optimal. Nevertheless, it is still a very good candidate, provided that an appropriate function is chosen, satisfying [10]:

$$\begin{array}{rcl} \overline{T} & \geq & e^{\omega_0 h} \\ \overline{\tau} & \leq & \frac{h}{\mathrm{E}_1\left[\rho(n)\right]} \end{array} \tag{18}$$

with $\omega_0 > 0$ the non-zero root of the equation $\mathrm{E}_0\left[e^{\omega\rho(n)}\right] = 1$. Usually, the mean before the change of $\rho(n)$ is not negative and then the idea proposed above is not applicable. In that case, we propose the following modified metric:

$$\rho(n) \doteq \rho_b(n) - b. \tag{19}$$

In this way, by selecting a proper offset b, the mean of $\rho(n)$ before the change will be negative, but it will become positive after the change. Specifically, the choice of the offset b should be large enough to ensure a negative mean before change and to provide a certain false alarm rate. But at the same time, b should be small enough to maintain a positive mean after change. From (18) and using expression of $\rho(n)$ in (19), we are able to adjust the false alarm rate through the nonzero root, ω_0, of the following equation:

$$\mathrm{E}_0\left[e^{\omega\cdot(\rho_b(n)-b)}\right] = \mathrm{E}_0\left[e^{\omega\cdot\rho_b(n)}\right] e^{-\omega\cdot b} = 1 \tag{20}$$

where $\mathrm{E}_0\left[e^{\omega\rho_b(n)}\right]$ is the characteristic function of $\rho_b(n)$ under $\mathcal{H}_0$.

IV. Quickest Detection for Integrity Monitoring

This section describes the application of the CUSUM algorithm to GNSS integrity monitoring. To do so, we use the CUSUM algorithm for interference and multipath detection, which are two of the most relevant and common problems in GNSS. We show different situations for applying the CUSUM algorithm with the aim of serving as a guideline for further works in GNSS signal integrity algorithms.

A. CASE 1: Completely known LLR

This section presents the case when both distributions before and after the change are completely known. This situation is often the case when addressing interference detection in GNSS. Interference detection must be carried out at the output of the GNSS front-end, since it is here where the interference is visible. In the absence of interference (i.e. under $\mathcal{H}_0$), the received signal will be dominated by noise, since the GNSS signal remains under the noise floor, whereas in the presence of interference (i.e. under $\mathcal{H}_1$), the received signal will be dominated by the interference. Then the detection problem becomes to a problem of measuring the degree of Gaussianity of the received signal samples.

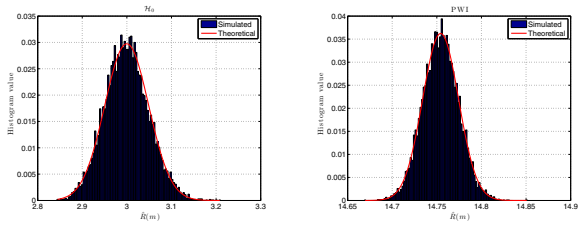

Fig. 2: Statistical characterisation of the kurtosis value. Comparison between simulated (i.e. histogram) and theoretical PDF under $\mathcal{H}_0$ (left) and $\mathcal{H}_1$ (right).

The kurtosis value is a good measurement for measuring the Gaussianity of the data. It is a statistical measurement equal to 3 if the data is Gaussian (i.e. under $\mathcal{H}_0$) , otherwise (i.e. under $\mathcal{H}_1$) it departs from 3. Let us define the kurtosis estimate as $\hat{R}(m) \doteq \frac{\hat{\zeta}_4^{(m)}}{\hat{\zeta}_2^{(m)^2}}$, with $\hat{\zeta}_i^{(m)}$ the N-samples estimates of the i-th central moment of the received samples. In turn, m stands for the snapshot index, where each snapshot includes N samples. In Chapter 9 of [13] it is stated that for large N (i.e. $N \gg 1000$), the estimate of the kurtosis of a Gaussian variable is another Gaussian variable with mean and variance below,

$$\begin{aligned} \mathrm{E}_0\left[\hat{R}(m)\right] &= \mu_0^{(\mathrm{kurt})} = 3\cdot\frac{N-1}{N+1}, \\ \mathrm{var}_0\left[\hat{R}(m)\right] &= \sigma_0^{2^{(\mathrm{kurt})}} = \frac{24}{N}. \end{aligned} \tag{21}$$

Moreover, in the presence of interference, the kurtosis estimates are still Gaussian distributed but suffering from a significant change in both mean and variance. From [14] we have the following expression for the mean under the presence of pulsed wave interferences (PW):

$$\mathrm{E}_1\left[\hat{R}(m)\right] = \mu_1^{(\mathrm{kurt})} = \mu_0^{(\mathrm{kurt})}\cdot\frac{1+2\mathrm{INR}+\frac{\mathrm{INR}^2}{2\mathrm{dc}}}{\left(1+\mathrm{INR}\right)^2} \tag{22}$$

where INR is the interference-to-noise ratio, and dc the duty cycle.

Thereby, we can formulate the kurtosis-based detection in a quickest interference detection framework as follows:

$$\begin{aligned} \mathcal{H}_0 &: \hat{R}(m) \sim \mathcal{N}\left(\mu_0^{(\mathrm{kurt})}, \sigma_0^{2^{(\mathrm{kurt})}}\right), \quad m < v \\ \mathcal{H}_1 &: \hat{R}(m) \sim \mathcal{N}\left(\mu_1^{(\mathrm{kurt})}, \sigma_1^{2^{(\mathrm{kurt})}}\right), \quad m \geq v \end{aligned} \tag{23}$$

This is shown in Fig. 2, where the distribution of the kurtosis is shown for the case of nominal conditions and under the presence of a PW. We see how under $\mathcal{H}_0$ (see left plot) the kurtosis distribution is Gaussian with a mean close to 3. On the other hand, under $\mathcal{H}_1$ (see right plot) the kurtosis distribution is still Gaussian but the mean departs from the baseline value equal to 3. These results have been obtained with $N =$ 1e4 samples, INR = 20dB, dc = 0.1, and 1e5 Monte-Carlo realisations. This value of INR is selected to show a distinguishable change in the mean of the kurtosis.

From (23) we have characterised the statistical behaviour of the kurtosis, with $\mu_0^{(\mathrm{kurt})}$, $\sigma_0^{2^{(\mathrm{kurt})}}$ and $\mu_1^{(\mathrm{kurt})}$ known. However, $\sigma_1^{2^{(\mathrm{kurt})}}$ is not known, and then we cannot fully characterise the distribution under $\mathcal{H}_1$. Nevertheless, we can assume that both variances before and after the change are equal (i.e. $\sigma_0^{2^{(\mathrm{kurt})}} = \sigma_1^{2^{(\mathrm{kurt})}}$), and then use the CUSUM algorithm as a Gaussian mean change detector. Doing so, we are able to fully characterise the LLR as follows:

$$\mathrm{LLR}_{\mathrm{kurt}}(m) = \frac{\mu_1^{(\mathrm{kurt})} - \mu_0^{(\mathrm{kurt})}}{\sigma_0^{2^{(\mathrm{kurt})}}}\left(\hat{R}(m) - \frac{\mu_1^{(\mathrm{kurt})} + \mu_0^{(\mathrm{kurt})}}{2}\right) \tag{24}$$

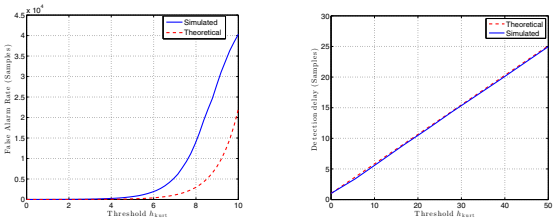

Fig. 3: CUSUM performance for the kurtosis method with respect to the detection threshold and $N = 1e4$. False alarm rate in samples (left) and detection delay for a PW with INR = -10dB.

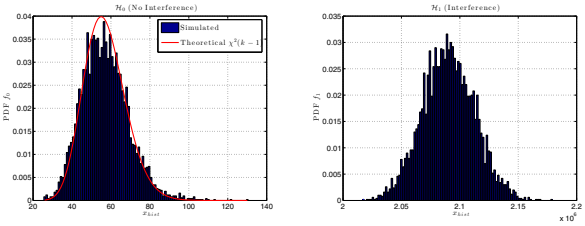

Fig. 4: Statistical characterisation of the chi-squared GoF test metric. Comparison between simulated (i.e. histogram) and theoretical PDF under $\mathcal{H}_0$ (left), and simulated PDF under $\mathcal{H}_1$ (right).

with $\mu_0^{(\text{kurt})}$, $\sigma_0^{2(\text{kurt})}$ and $\mu_1^{(\text{kurt})}$ defined as in (21)–(22), and $\hat{R}(m)$ the N-sample kurtosis estimate at snapshot m.

It is worth to mention that $\mu_1^{(\text{kurt})}$ is known to depend on the INR and duty cycle of the interference. Hence, a way to proceed is to fix a certain value for $\mu_1^{(\text{kurt})}$ according to the minimum INR that one expects to detect. Moreover, the duty cycle might be fixed to the value that produces the minimum change possible. In this way, a minimum change detection is set allowing the detection of any larger change caused by higher power interferences or with duty cycle that give rise to large changes in the kurtosis. Thereby, we can make use of the CUSUM algorithm decision rule, which can be written as follows:

$$g^{(\text{kurt})}(m) = \left(g^{(\text{kurt})}(m-1) + \text{LLR}_{\text{kurt}}(m)\right)^+ \geq h_{\text{kurt}} \quad (25)$$

which leads to the following performance in terms of false alarms and detection delay

$$\begin{array}{rcl} \overline{T}_{\text{kurt}} & \geq & e^{h_{\text{kurt}}} \\ \overline{\tau}_{\text{kurt}} & \leq & \frac{h_{\text{kurt}}}{\text{K}_{\text{kurt}}(f_1,f_0)} \end{array} \quad (26)$$

with $\text{K}_{\text{kurt}}(f_1,f_0) = \frac{\left(\mu_1^{(\text{kurt})} - \mu_0^{(\text{kurt})}\right)^2}{2\sigma_0^{2(\text{kurt})}}$ the Kullback-Leibler divergence for the kurtosis value.

These bounds are presented in Fig. 3, which are compared with simulated results. To do so, we use $N = 1e4$ and 1e5 Mote-Carlo runs. In addition, for the detection delay case we fix $\text{INR} = -10\text{dB}$. This low value is selected in order to show representative results, otherwise the detection delay would be one sample. The left plot shows that the simulated number of samples between false alarms is larger than the lower bound, which allows us to set a threshold h_{kurt} assuring certain desired false alarm rate. Moreover, the right plot shows similar values for the simulated and theoretical results.

B. CASE 2: Unknown LLR

Here, we present the case when the distribution after change is completely unknown. To do so, we apply the CUSUM algorithm to interference detection again, but using a different metric for measuring the Gaussianity of the received signal samples.

We know that in the interference free-case, the received signal is dominated by noise, so that the histogram of the received samples should follow a Gaussian shape. Meanwhile, when interference is present, the received signal is dominated by the interference, and then the histogram of the received samples should not follow a Gaussian shape. This detection problem is equivalent to a goodness-of-fit test (GoF), specifically, we are interested on determining whether our received signal follows a Gaussian distribution or not. To do so, we make use of the chi-square test, which uses a metric based on the histogram and have a known distribution under $\mathcal{H}_0$.

The chi-square test evaluates the following test statistic [13]:

$$x_{\text{hist}}(m) = \sum_{i=1}^{N_b} \frac{\left(O_i^{(m)} - E_i\right)^2}{E_i} \quad (27)$$

with E_i the i-th bin of the reference *theoretical* histogram evaluated under $\mathcal{H}_0$ with N_b bins. These values are obtained by calculating the theoretical PDF for each bin and transforming it to a histogram value (i.e. taking into account the snapshot samples N and the separation between bins). $O_i^{(m)}$ is the i-th bin of the measured histogram with N_b bins at snapshot m, where each snapshot contains N samples.

Pearson [13] claimed that for large N, the variable $x_{\text{hist}}(m)$ under $\mathcal{H}_0$ is approximately chi-squared distributed with $N_b - 1$ degrees of freedom, whereas under $\mathcal{H}_1$ it departs from a central chi-square distribution. Therefore, we can write the following hypotheses:

$$\begin{array}{ll} \mathcal{H}_0 : x_{\text{hist}}(m) \sim \chi^2\left(N_b - 1\right), & m < v \\ \mathcal{H}_1 : x_{\text{hist}}(m) \nsim \chi^2\left(N_b - 1\right), & m \geq v \end{array} \quad (28)$$

Fig. 4 shows the distribution of both hypotheses. We can see how under $\mathcal{H}_0$ (see left plot) the histogram of the simulated data almost fits the theoretical χ^2 with $N_b - 1$ degrees of freedom. On the other hand, in the right plot, it is shown the histogram for the case when the simulated interference is present. It can be seen how it departs from the $\chi^2(N_b - 1)$ distribution, which is expected under $\mathcal{H}_0$. The results have been obtained using a number of $N = 1e4$ samples, 1e5 Monte-Carlo runs, a number of bins $N_b = 50$, and a continuous wave (CW) with $\text{INR} = 20\text{dB}$ for the $\mathcal{H}_1$ case. Again, this INR value is selected in order to show a visible change on distribution.

Up to now, we have seen that the distribution after change is different to that before the change, but we have no knowledge about this distribution. Hence, since the distribution under $\mathcal{H}_1$ is unknown, the LLR cannot be completely defined, and then we are unable to apply the CUSUM algorithm directly to $x_{\text{hist}}(m)$. However, since $\mathcal{H}_0$ is known, we can use the Offset-CUSUM alternative (Sect. III-C). To do so we need to propose a metric ρ that has a negative mean before the change and a positive mean after the change. It is known that the mean of a χ^2 random variable is equal to the number of degrees of freedom of the χ^2. Therefore, $\text{E}_0\left[x_{\text{hist}}(m)\right] = N_b - 1 > 0$. Hence, we cannot use directly $x_{\text{hist}}(m)$ as the function ρ to be used in the Offset-CUSUM, but we define the following modified function

$$\rho_{\text{hist}}(m) \doteq x_{\text{hist}}(m) - b \quad (29)$$

with b a proper offset for which the mean of $\rho_{\text{hist}}(m)$ before change is negative, but it is positive after change.

Moreover, the choice of the offset b should be large enough to ensure a negative mean before change and to provide a certain false alarm rate, always maintaining a positive mean after change. From (20) and using $\rho_{\text{hist}}(m)$ we are able to adjust the false alarm rate through the nonzero root ω_0 with (18), which turns out to be the nonzero root of the next equation:

$$e^{\omega b} = (1 - 2\omega)^{-\frac{N_b - 1}{2}} \quad (30)$$

which can be solved numerically. Thus, the choice of the offset b will fix a value for ω_0 given by the equation above. Thereby, making

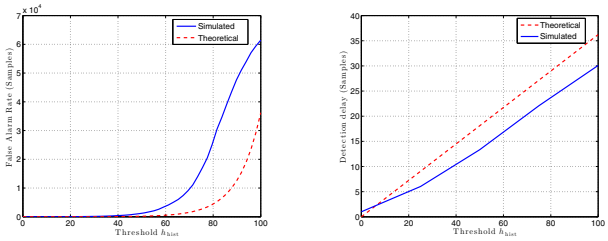

Fig. 5: Offset-CUSUM performance for the GoF metric as a function of detection threshold, with a fixed $N_b = 50$ and $b = 55$. False alarm rate (left) and detection delay for a CW with INR $= -20$dB (right).

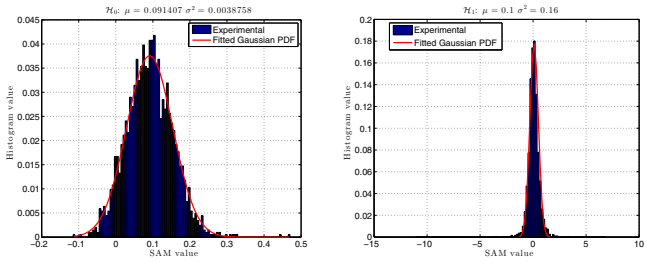

Fig. 6: Statistical characterisation of SAM with real data captured in London downtown under benign conditions (left) and under harsh conditions (right).

use of the following decision rule (i.e. Offset-CUSUM)

$$g^{(\rm hist)}_{\rm Offset}(m) = \left(g^{(\rm hist)}_{\rm Offset}(m-1) + \rho_{\rm hist}(m)\right)^+ \geq h_{\rm hist} \tag{31}$$

we obtain the following performance

$$\begin{array}{rcl} \overline{T}_{\rm hist} & \geq & e^{\omega_0 h_{\rm hist}} \\ \overline{\tau}_{\rm hist} & \leq & \frac{h_{\rm hist}}{\mathrm{E}_1[\rho_{\rm hist}(m)]} \end{array} \tag{32}$$

where ω_0 is the nonzero solution of (30).

These bounds are proved in Fig. 5, which shows the Offset-CUSUM performance using $\rho_{\rm hist}(m)$, and compares it with the theoretical bounds above. To do so, we set the offset to $b = 55$, the number of bins to $N_b = 50$ and for the case of the detection delay we simulate a CW interference with INR $= -20$dB. As for the previous case, this low value is chosen in order to show representative results. The left plot of Fig. 5 shows how the simulated number of samples between false alarms is greater than the lower bound given in (32). This observation allows us to set a threshold $h_{\rm hist}$ assuring certain desired false alarm rate. In addition, the right plot shows the detection delay measured in samples with respect to the set threshold, and shows similar values for the simulated results and theoretical ones, being the simulated delay below the upper bound given in (32).

C. CASE 3: Incompletely known LLR

Finally, we present the case when both the distribution before and after the change are known, but some of the parameters of them are unknown. To do so, we apply the CUSUM algorithm to multipath detection in GNSS. The problem of multipath detection must be carried out at the acquisition and/or tracking stage (i.e. after de-spreading), since it is there where multipath effects are visible. Multipath affects tracking measurements such as the estimated C/N_0, the code discriminator output and the shape of the correlation curve [2]. Therefore, we will able to detect NLOS and multipath based on the fluctuations of these measurements.

Next, we show the application of the CUSUM algorithm for the slope asymmetry metric (SAM), which is based on the correlation curves calculated in the tracking loop of any GNSS receiver. We know that under benign conditions (i.e. H_0) the correlation curve (i.e. correlation between local replica and received signal) is symmetrical, but it loses the symmetry under harsh conditions (i.e. H_1). This can be measured by the SAM, which under ideal conditions is close to cero (indicating symmetry), whereas when multipath is present it departs from cero (indicating asymmetry). Specifically, when multipath is present we experience two different effects [15]:

- Under LOS condition, the mean of the SAM departs from 0.
- Under NLOS condition, the variance of the SAM fluctuates.

Indeed, in both LOS and NLOS cases the mean and variance vary, but the mean change is predominant in LOS situations (i.e. deterministic component of LOS prevails), whereas the variance change is predominant in NLOS (i.e. random components due to MP prevails). Therefore, this is equivalent to have a change on both the variance and mean of the SAM. Since the SAM is calculated from the Lest Squares (LS) estimation of the slope at both sides of the correlation curve, we can assume that the distribution of the SAM is Gaussian, and then we can formulate the problem of multipath detection in a quickest detection framework as follows:

$$\begin{array}{l} \mathcal{H}_0 : x_{\rm S}(k) \sim \mathcal{N}\left(\mu_0^{(\rm S)}, \sigma_0^{2^{(\rm S)}}\right), \quad k < v \\ \mathcal{H}_1 : x_{\rm S}(k) \sim \mathcal{N}\left(\mu_1^{(\rm S)}, \sigma_1^{2^{(\rm S)}}\right), \quad k \geq v \end{array} \tag{33}$$

This is shown in Fig. 6, which shows the histogram of the SAM values for data gathered with a real GNSS receiver in urban environment under the framework of the Integrity GNSS Receiver (iGNSSrx) project, which was funded by the European Commission. This data was captured during 80 sec. in a scenario (i.e. London downtown) that was under benign conditions the first 40 sec., and then it changes to harsh conditions until the end of the captured data. We discriminate between benign and harsh conditions by analysing the positioning error (i.e. using a truth reference). For the data under benign conditions we obtain a mean positioning error of 2 meters, whereas for the data under harsh conditions we obtain a mean positioning error of 50 meters. In the left plot of Fig. 6 we present the histogram under benign conditions, where we see the Gaussianity of the SAM with a mean about 0.1. This value is due to the asymmetry introduced by the front-end filter and can be calibrated. On the other hand, in the right plot, we present the histogram under harsh conditions. We see how the histogram fits a Gaussian distribution quite well. This distribution has a mean that is also close to 0, but a much greater variance than under $\mathcal{H}_0$. This may be because in this case we are under NLOS conditions, and then the change on the variance is predominant.

We see that the SAM follows a Gaussian distribution with known mean before change (i.e. it must be calibrated) but unknown a priori variance before and after the change. In order to use the CUSUM algorithm we propose the following configuration of the Gaussian distribution parameters:

- $\mu_0^{(\rm S)}$: It should be equal to 0, but in practice it is slightly larger due to the shape of the front-end filter. Hence
$$\mu_0^{(\rm S)} = \kappa \sim 0 \tag{34}$$
with $\kappa = 0.1$ herein.
- $\sigma_0^{2^{(\rm S)}}$: This value is unknown a priori because it's difficult to have a perfect knowledge of the actual variance, even knowing the expression for the variance of the SAM. This is so because it ultimately will depend on the multipath parameters, which will be random and unknown. Hence, we propose to fix the variance under benign conditions according to the maximum allowable variations on the SAM values under $\mathcal{H}_0$, as follows:
$$\sigma_0^{2^{(\rm S)}} \doteq \left(\frac{(\Delta_0)_{\max}}{3}\right)^2 \tag{35}$$

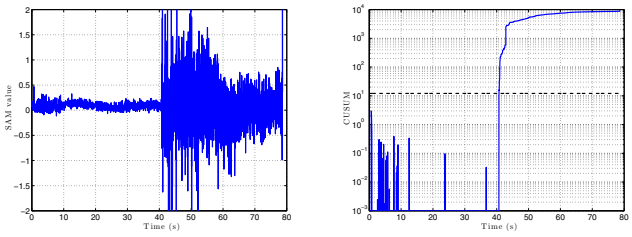

Fig. 7: SAM-based multipath detection for the real data characterised in Fig. 6. Metric (left) and CUSUM (right) time-evolution.

with $(\Delta_0)_{\max}$ the maximum allowable variations under $\mathcal{H}_0$. This is so because we know that for a Gaussian distribution the 99.86% of the values are comprised in the interval $\mu \pm 3\sigma$. For example, in our case we see that the SAM under $\mathcal{H}_0$ takes variations between -0.1–0.45. Therefore, a proper value for $(\Delta_0)_{\max}$ may be ± 0.4, which give rise to $\sigma_0^{2^{(\mathrm{S})}} = 1.78\text{e-}2$.

- $\mu_1^{\pm^{(\mathrm{S})}}$: This value is unknown, but it might be fixed as follows

$$\mu_1^{\pm^{(\mathrm{S})}} = \mu_0^{(\mathrm{S})} \pm \delta \tag{36}$$

with δ a proper value selected through experimental analysis.

- $\sigma_1^{2^{(\mathrm{S})}}$: Similarly as for the variance before change, we fix the variance under harsh conditions as the minimum detectable variability on the SAM due to multipath as follows:

$$\sigma_1^{2^{(\mathrm{S})}} \doteq \left(\frac{(\Delta_1)_{\min}}{3}\right)^2 \tag{37}$$

with $(\Delta_1)_{\min}$ the minimum detectable variation on the SAM under $\mathcal{H}_1$. For instance, in our case, a suitable value might be a variation equivalent to ± 0.6, which results in $\sigma_1^{2^{(\mathrm{S})}} = 4\text{e-}2$.

Therefore, since the SAM may present a change in both mean and variance, we suggest the use of two different CUSUM algorithms, one for detecting the change in variance (i.e. NLOS) and another for detecting the mean change (i.e. LOS). The expression for the mean change CUSUM would be like (24), but with the SAM parameters. For the variance change CUSUM the LLR expression is as follows:

$$\mathrm{LLR}_{\mathrm{S}}(k) = \ln\left(\frac{\sigma_0^{(\mathrm{S})}}{\sigma_1^{(\mathrm{S})}}\right) + \frac{\left(x_{\mathrm{S}}(k) - \mu_0^{(\mathrm{S})}\right)^2}{2} \cdot \left(\frac{1}{\sigma_0^{2^{(\mathrm{S})}}} - \frac{1}{\sigma_1^{2^{(\mathrm{S})}}}\right) \tag{38}$$

with $\mu_0^{(\mathrm{S})}$, $\sigma_0^{(\mathrm{S})}$, $\sigma_1^{(\mathrm{S})}$ defined as in (34)–(37), and $x_{\mathrm{S}}(k)$ the calculated SAM value at the k-th post-correlation snapshot.

Thereby, we can use the following decision rule:

$$g^{(\mathrm{S})}(k) = \left(g^{(\mathrm{S})}(k-1) + \mathrm{LLR}_{\mathrm{S}}(k)\right)^+ \geq h_{\mathrm{S}} \tag{39}$$

leading to the next performance

$$\begin{array}{rcl} \overline{T}_{\mathrm{S}} & \geq & e^{h_{\mathrm{S}}} \\ \overline{\tau}_{\mathrm{S}} & \leq & \frac{h_{\mathrm{S}}}{\mathrm{K}_{\mathrm{S}}(f_1, f_0)} \end{array} \tag{40}$$

with $\mathrm{K}_{\mathrm{S}}(f_1, f_0) = \ln\left(\frac{\sigma_0^{(\mathrm{S})}}{\sigma_1^{(\mathrm{S})}}\right) + \frac{\sigma_1^{2^{(\mathrm{S})}}}{\sigma_0^{2^{(\mathrm{S})}}} - \frac{1}{2}$.

Here, we show the obtained SAM and the CUSUM evolution for the real data characterised in Fig. 6. This is show in Fig. 7, where we see in the left plot how the SAM presents a change in variance just when multipath appears (i.e. second 40). The change of the variance is quite large and then it is promptly detected when it appears, as it is shown in the right plot, which shows how the CUSUM remains close to 0 until it start drifting upward at second 40 and it crosses the threshold. The threshold is set to fix a false alarm rate of 1h, which with a sampling rate of 10MHz and snapshot time of 20ms becomes equal to $h_{\mathrm{S}} = 12$ (see dashed black line in Fig. 7).

V. Conclusion

This paper presents a quickest detection framework for threat detection with the aim of improving integrity in single-antenna GNSS receivers. Two different cases have been presented: when the LLR is completely known and when the LLR is unknown. For the first case, the optimal solution is the CUSUM algorithm, whereas for the second case there is not an optimal solution, but a practical approach that works quite well may be the Offset-CUSUM. Applying the CUSUM to interference detection we cope with the cases of completely known and unknown LLR, which leads to the use of the CUSUM algorithm and the Offset-CUSUM, respectively. These algorithms are applied to two different metrics, the kurtosis and the chi-squared GoF test, respectively. The first is relevant for PW/CW interferences (i.e. it is restricted to the type of interference), but it uses the optimal algorithm. The GoF method, is not restricted to any kind of interference, but it does not use the optimal algorithm, and then it may be used as complement for the cases when the kurtosis is not useful. On the other hand, applying the CUSUM to multipath detection we cope with the case of incompletely known LLR. In this case, we see that we can discriminate between LOS and NLOS conditions. This is done by making use of two different CUSUM for detecting both the change in mean and variance of the SAM. Results of the proposed methods show the suitability of quickest detection for interference and multipath detection and the potential interest in practical applications involving GNSS signal integrity real-time monitoring. This is so because the suggested methods make use of simple techniques (i.e. easy for practical implementation) and allow tuning the CUSUM algorithm providing certain performance in terms of false alarms or detection delay fixed by the user through the detection threshold.

References

[1] B. W. Parkinson and J. J. Spilker, *Global Positioning System: Theory and Applications vol.2*. Aiaa, 1996.

[2] G. Seco-Granados, J. A. López-Salcedo *et al.*, "Challenges in Indoor Global Navigation Satellite Systems: Unveiling its core features in signal processing," *IEEE Sig. Process. Mag.*, vol. 29, no. 2, pp. 108–131, 2012.

[3] A. T. Balaei *et al.*, "A statistical inference technique for GPS interference detection," *IEEE Trans. on Aeros. and Electr. Syst.*, vol. 45, no. 4, 2009.

[4] H. K. Lee *et al.*, "GPS Multipath Detection Based on Sequence of Successive-Time Double-Differences," *IEEE Sig. Proc. Lett.*, vol. 11, no. 3, pp. 316–319, 2004.

[5] A. Broumandan *et al.*, "GNSS spoofing detection in handheld receivers based on signal spatial correlation," *IEEE PLANS*, pp. 479–487, 2012.

[6] S. P. Mertikas, "Automatic and Online Detection of Small but Persistent Shifts in GPS Station Coordinates by Statistical Process Control," *GPS Solutions*, vol. 5, no. 1, pp. 39–50, 2001.

[7] T. Osklper and H. V. Poor, "Quickest detection of a random signal in background noise using a sensor array," *Eurasip Journal on Applied Signal Processing*, no. 1, 2005.

[8] L. Lifeng, F. Yijia, and H. V. Poor, "Quickest Detection in Cognitive Radio: A Sequential Change Detection Framework," in *IEEE GLOBECOM*, 2008, pp. 1–5.

[9] D. Egea, J. A. López-Salcedo, and G. Seco-Granados, "Interference and multipath sequential tests for signal integrity in multi-antenna GNSS receivers," in *IEEE 8th Sensor Array and Multichanel Signal Processing Workshop*, 2014, pp. 117–120.

[10] D. Egea, G. Seco-Granados, and J. A. López-Salcedo, "Single- and Multi-Correlator Sequential Tests for Signal Integrity in Multi-Antenna GNSS Receivers," in *ICL-GNSS*, 2014, pp. 117–120.

[11] H. V. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge, 2009.

[12] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Englewood Cliffs: Prentice Hall, 1993.

[13] R. B. D'Agostino *et al.*, *Goodness-of-Fit Techniques*. CRC Press, 1986.

[14] R. D. De Roo *et al.*, "Sensitivity of the kurtosis statistic as a detector of pulsed sinusoidal radio frequency interference in a microwave radiometer receiver," in *IGARSS*, vol. 45, no. 7, 2007, pp. 2706–2709.

[15] J. A. López-Salcedo *et al.*, "Multipath detection metrics and attenuation analysis using a GPS snapshot receiver in harsh environments," in *IEEE EuCAP*, 2009, pp. 3692–3696.