Integrity Monitoring of Integrated Satellite/Inertial Navigation Systems Using the Likelihood Ratio

Jan Palmqvist, *Linköping University*

BIOGRAPHY

Jan Palmqvist received his M.Sc. in Applied Physics and Electrical Engineering from Linköping University, Sweden in 1986. He has been with Saab Military Aircraft since 1989 and in 1994 he also joined the Automatic Control Group at Linköping University as a part time Ph.D. student. His current research focuses on algorithms for integrity monitoring of integrated navigation systems.

ABSTRACT

Global Navigation Satellite Systems (GNSS) have the ability to fulfill the navigation accuracy requirements of most applications. The systems do however lack continuity and integrity to meet the requirements of high precision navigation applications. The use of a combination of Inertial Navigation Systems (INS) and GNSS information do however show promising results in fulfilling these requirements. Methods for monitoring the integrity of integrated INS-GNSS systems are investigated.

Integration of INS and GNSS is usually accomplished using a Kalman filter for recursive estimation of the parameters of interest. The residual used for integrity monitoring is the Kalman filter innovation.

The innovation signatures of different types of faults are analyzed. Since two of the most likely types of faults in an integrated solution are INS sensor bias shifts and satellite range bias drifts or jumps, these additive types of changes are studied in more detail. Taking the approach of hypothesis testing of the two hypotheses un-failed and failed system, fault detection methods based on the likelihood ratio are considered and the Generalized Likelihood Ratio (GLR) test is proposed to be used. This method uses the innovations of the Kalman filter to compute the maximum likelihood estimates of the time and magnitude of an assumed change. Using these estimates, it evaluates the log-likelihood ratio of a change versus no change. The GLR test uses a linearly in time increasing number of matched filters. Different ways of decreasing this computational burden are discussed, showing that fast detection can be achieved even with a small and constant number of matched filters.

A further advantage of the GLR test is that in addition to detecting the occurrence of a fault, it also estimates its magnitude, direction and time of occurrence, making it possible to identify the source of the fault, exclude faulty satellites and correct the Kalman filter estimate without re-processing the affected data.

INTRODUCTION

A combination of Global Navigation Satellite Systems (GNSS) and Inertial Navigation Systems (INS), has shown promising results in solving the continuity [5] and integrity [3, 11] demands of high precision navigation applications. The satellite system information and the INS have dissimilar but complementary characteristics that in many ways make them ideal complements.

The advantage of an integrated solution over stand alone systems regarding integrity is that the additional information available can be used to check for slowly drifting types of faults that can not be detected without this redundancy. It is well known that the proposed GPS RAIM algorithms, especially when not augmented with WADGPS and LADGPS, have problems detecting smaller range bias drift errors [2] that still would cause the navigation solution to drift off. Furthermore the need for a minimum number of satellites can be circumvented.

The integrity monitoring approach investigated in this paper uses test statistics based on the Kalman filter innovations and it can be used both in centralized and cascaded Kalman filter integration philosophies. However, there will be no satellite exclusion methodology for cascaded Kalman filter approaches where the output from the GNSS receiver, with its own RAIM, will be monitored for non detected faults.

While many proposed integrity monitoring methods are designed to detect all types of faults with one test, there are important advantages to be gained using specific tests for each type of considered fault. Even if the most likely types of faults are known there will always be other types, not known or considered. Hence these two approaches should be used in combination.

After introducing the notation of the Kalman filter, the innovation signatures of different types of additive faults in the state space model of the system will be analyzed.

Since two of the most likely types of faults in an integrated solution are INS sensor bias shifts and satellite range bias drifts and jumps these additive types of changes are studied in more detail.

The next two sections describe how different level of knowledge can be used to detect changes in the mean of a Gaussian sequence, the use of the Likelihood Ratio (LR) for hypothesis testing is introduced, and the Generalized Likelihood Ratio (GLR) test is studied in more detail.

Since a good integrity monitoring approach consists of *detection*, *identification* and *adaptation* the two further sections discuss how the test statistics can be changed to look for more specific faults and how the Kalman filter state estimates can easily be adapted to changes in the states or the measurements without re-processing the affected data.

Furthermore, on-line implementations are considered showing how the computational complexity can be decreased.

Finally the performance of the GLR test method is compared with some other types of integrity monitoring algorithms.

KALMAN FILTERING

An integrated navigation system uses a Kalman filter for recursive estimation of the parameters of interest. This filter uses a discrete time model of the underlying system of the form:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{F}_t \boldsymbol{x}_t + \boldsymbol{G}_t \boldsymbol{u}_t + \boldsymbol{w}_t \\ \boldsymbol{y}_t &= \boldsymbol{H}_t \boldsymbol{x}_t + \boldsymbol{e}_t, \end{aligned} \tag{1}$$

where $\boldsymbol{x}_t$ is the state vector, $\boldsymbol{y}_t$ is the measurement, $\boldsymbol{u}_t$ known inputs, $\boldsymbol{F}_t$, $\boldsymbol{G}_t$ and $\boldsymbol{H}_t$ are matrices that are known at time t. The noises $\boldsymbol{w}_t$ and $\boldsymbol{e}_t$ are assumed to be independent and Gaussian with covariances $\boldsymbol{Q}_t$ and $\boldsymbol{R}_t$, respectively. The Kalman filter for recursive estimation of the state vector $\boldsymbol{x}_t$ reads as follows:

$$\begin{aligned} \hat{\boldsymbol{x}}_{t+1|t} &= \boldsymbol{F}_t \hat{\boldsymbol{x}}_{t|t} + \boldsymbol{G}_t \boldsymbol{u}_t \\ \hat{\boldsymbol{x}}_{t|t} &= \hat{\boldsymbol{x}}_{t|t-1} + \boldsymbol{K}_t \boldsymbol{\epsilon}_t, \end{aligned} \tag{2}$$

where the indices $t|t$ and $t+1|t$ denotes the parameter at time t and $t+1$ respectively, based on measurement up to time t. The innovations $\boldsymbol{\epsilon}_t$ are given by

$$\boldsymbol{\epsilon}_t = \boldsymbol{y}_t - \boldsymbol{H}_t \hat{\boldsymbol{x}}_{t|t-1}. \tag{3}$$

The state estimate error covariance matrices $\boldsymbol{P}_{t+1|t}$ and $\boldsymbol{P}_{t|t}$, and Kalman gain matrix $\boldsymbol{K}_t$ are given by:

$$\begin{aligned} \boldsymbol{P}_{t+1|t} &= \boldsymbol{F}_t \boldsymbol{P}_{t|t} \boldsymbol{F}_t^T + \boldsymbol{Q}_t \\ \boldsymbol{P}_{t|t} &= (\boldsymbol{I} - \boldsymbol{K}_t \boldsymbol{H}_t) \boldsymbol{P}_{t|t-1} \\ \boldsymbol{K}_t &= \boldsymbol{P}_{t|t-1} \boldsymbol{H}_t^T \boldsymbol{S}_t^{-1} \\ \boldsymbol{S}_t &= \boldsymbol{H}_t \boldsymbol{P}_{t|t-1} \boldsymbol{H}_t^T + \boldsymbol{R}_t. \end{aligned} \tag{4}$$

If the Gauss-Markov model holds and the noise sequences are white and Gaussian the innovations, $\boldsymbol{\epsilon}_t$, will be independent, Gaussian distributed as $\mathcal{N}(\boldsymbol{0}, \boldsymbol{S}_t)$.

Even though a reduced order model of the real navigation system and an extended Kalman filter are used, causing the innovations not to be truly white and Gaussian, it is still appropriate to develop the theory as if they were. The innovations $\{\boldsymbol{\epsilon}_t\}$ are the appropriate parameter to study for detection of faults, in the state space system, or in the measurements. The parameters forming the innovations are based on all past and present measurements together with a model of the system. Hence the innovations will contain all the information needed for integrity monitoring, as will be seen in the next section.

INNOVATION SIGNATURE

In this section it will be shown that the innovations will be biased and distributed as:

$$\boldsymbol{\epsilon}_t(k) \in \mathcal{N}(\boldsymbol{\rho}_t(k), \boldsymbol{S}_t)$$

after an additive change.

Consider the discrete time description of the system (1) under general additive changes yielding:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{F}_t \boldsymbol{x}_t + \boldsymbol{G}_t \boldsymbol{u}_t + \boldsymbol{w}_t + \boldsymbol{C}_x \Upsilon_x(t,k) \\ \boldsymbol{y}_t &= \boldsymbol{H}_t \boldsymbol{x}_t + \boldsymbol{e}_t + \boldsymbol{C}_y \Upsilon_y(t,k), \end{aligned} \tag{5}$$

where $\boldsymbol{C}_x$ and $\boldsymbol{C}_y$ are vectors of dimension n and r representing change magnitudes and directions, and $\Upsilon_x(t,k)$ and $\Upsilon_y(t,k)$ are scalars representing the dynamic profiles of the assumed changes. The time instant k denotes the change time, making $\Upsilon_x(t,k)$ and $\Upsilon_y(t,k)$ identical to zero for $t < k$.

The dynamic signatures caused by these additive changes can recursively be written as in [1] using the following decomposition of the state, its estimate and the innovation:

$$\begin{aligned} \boldsymbol{x}_t(k) &= \boldsymbol{x}_t + \boldsymbol{\alpha}_t(k) \\ \hat{\boldsymbol{x}}_{t|t}(k) &= \hat{\boldsymbol{x}}_{t|t} + \boldsymbol{\beta}_t(k) \\ \boldsymbol{\epsilon}_t(k) &= \boldsymbol{\epsilon}_t + \boldsymbol{\rho}_t(k), \end{aligned} \tag{6}$$

where (k) denotes the parameter after a change. It can be shown that this yields the following recursions:

$$\begin{aligned}\boldsymbol{\alpha}_t(k) &= \boldsymbol{F}_t\boldsymbol{\alpha}_{t-1}(k) + \boldsymbol{C}_x\Upsilon_x(t-1,k)\\ \boldsymbol{\beta}_t(k) &= \boldsymbol{F}_{t-1}\boldsymbol{\beta}_{t-1}(k) + \boldsymbol{K}_t\boldsymbol{\rho}_t(k)\\ \boldsymbol{\rho}_t(k) &= \boldsymbol{H}_t[\boldsymbol{\alpha}_t(k) - \boldsymbol{F}_{t-1}\boldsymbol{\beta}_{t-1}(k)] +\\ &\qquad \boldsymbol{C}_y\Upsilon_y(t,k). \end{aligned} \qquad (7)$$

From these general equations all kinds of specific cases can be considered. We will here consider two specific types of additive changes:

1. A Kalman filter state shift with change magnitude $\boldsymbol{C}_x = \boldsymbol{\nu}_x$ and $\Upsilon_y(t,k) = 0$, corresponding to, e.g., a jump in one of the INS sensor biases.

2. A measurement variable bias drift or jump, with scalar drift rate or magnitude $\boldsymbol{C}_y = \boldsymbol{\nu}_y$ and $\Upsilon_x(t,k) = 0$, corresponding to, e.g., a GNSS satellite range bias drift or jump.

The signature on the innovation caused by a state bias shift at time k, corresponding to:

$$\Upsilon_x(t,k) = \delta(t-k)$$

i.e., equal to one at $t = k$ and zero elsewhere, can be expressed recursively as a linear regression in the change magnitude $\boldsymbol{\nu}_x$ as:

$$\begin{aligned}\hat{\boldsymbol{x}}_{t|t}(k) &= \hat{\boldsymbol{x}}_{t|t} + \boldsymbol{\beta}_t(k)\\ &= \hat{\boldsymbol{x}}_{t|t} + \boldsymbol{\mu}_{x,t}(k)\boldsymbol{\nu}_x\\ \boldsymbol{\epsilon}_t(k) &= \boldsymbol{\epsilon}_t + \boldsymbol{\rho}_t(k)\\ &= \boldsymbol{\epsilon}_t + \boldsymbol{\varphi}^T_{x,t-1}(k)\boldsymbol{\nu}_x \\ \boldsymbol{\varphi}^T_{x,t+1}(k) &= \boldsymbol{H}_t\left(\prod_{i=k}^{t}\boldsymbol{F}_i - \boldsymbol{F}_t\boldsymbol{\mu}_{x,t}(k)\right)\\ \boldsymbol{\mu}_{x,t+1}(k) &= \boldsymbol{F}_t\boldsymbol{\mu}_t(k) + \boldsymbol{K}_{t+1}\boldsymbol{\varphi}^T_{x,t+1}(k). \end{aligned} \qquad (8)$$

A typical example of the innovation signature due to a state change is shown in Figure 1.

A GNSS satellite range bias drift corresponds to

$$\Upsilon_y(t,k) = (t-k)\sigma(t-k),$$

yielding a drift with slope $\boldsymbol{\nu}_y$, starting at $t = k$. A satellite range bias jump instead corresponds to:

$$\Upsilon_y(t,k) = \sigma(t-k).$$

The signature on the innovation caused by a measurement variable drift or jump can now be expressed recursively as a linear regression in $\boldsymbol{\nu}_y$ as:

$$\begin{aligned}\hat{\boldsymbol{x}}_{t|t}(k) &= \hat{\boldsymbol{x}}_{t|t} + \boldsymbol{\beta}_t(k)\\ &= \hat{\boldsymbol{x}}_{t|t} + \boldsymbol{\mu}_{y,t}(k)\boldsymbol{\nu}_y\\ \boldsymbol{\epsilon}_t(k) &= \boldsymbol{\epsilon}_t + \boldsymbol{\rho}_t(k)\\ &= \boldsymbol{\epsilon}_t + \boldsymbol{\varphi}^T_{y,t-1}(k)\boldsymbol{\nu}_y \\ \boldsymbol{\varphi}^T_{y,t+1}(k) &= \Upsilon_y(t,k) - \boldsymbol{H}_t\boldsymbol{F}_t\boldsymbol{\mu}_{y,t}(k)\\ \boldsymbol{\mu}_{y,t+1}(k) &= \boldsymbol{F}_t\boldsymbol{\mu}_{y,t}(k) + \boldsymbol{K}_{t+1}\boldsymbol{\varphi}^T_{y,t+1}(k). \end{aligned} \qquad (9)$$

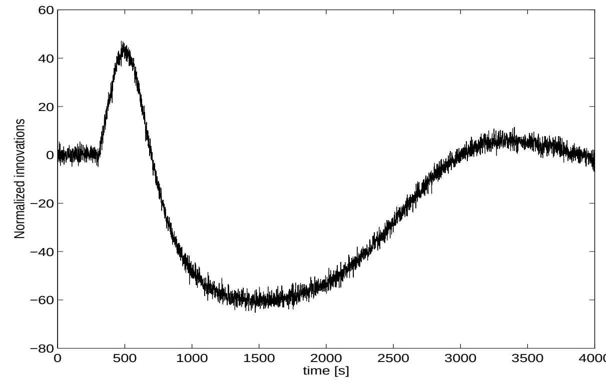

Fig. 1. Dynamic profile of normalized innovations after a state change at t=300.

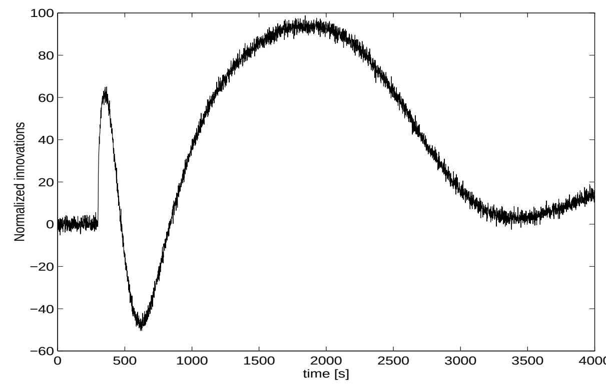

Fig. 2. Dynamic profile of normalized innovation after a measurement variable bias drift starting at t=300.

A typical innovation signature due to a ramp in a measurement variable is shown in Figure 2.

The expressions for $\boldsymbol{\varphi}_{x,t+1}(k)$ and $\boldsymbol{\mu}_{x,t+1}(k)$ in (8) and $\boldsymbol{\varphi}_{y,t+1}(k)$ and $\boldsymbol{\mu}_{y,t+1}(k)$ in (9) differs, but the same Greek letters are used in both cases since the test statistic will be formed in the same way, when deriving the GLR test.

It is possible to use $\boldsymbol{C}_x$ and $\boldsymbol{C}_y$ to constrain the possible faults to a subset of the state space or measurement variables. We will return to this in connection with diagnosis.

Once again, note that the expressions for the innovation signatures are linear regressions in the change. A fact that will be used when detecting the changes later on.

RECURSIVE DETECTION

It was shown in the last section that a change, not described by (1) will cause the innovations (3) to be biased. There are a number of approaches [1] for change detection in a Gaussian sequence and in this section some of them will be described.

The simplest approach for detecting changes in the innovation sequence is to filter under the hypothesis that there is no change and to check the whiteness of the innovations. This can be achieved by checking:

- Normalized innovations $s_t = \epsilon_t/\sqrt{S_t}$
- Squared normalized innovations $s_t = \epsilon_t^T S_t^{-1} \epsilon_t$.

The first one should sum up to approximately zero and the second one is the well-known χ^2-test, which also checks for changes in the variance. These statistics have well-known statistical distributions and the choice of thresholds becomes standard.

A more systematic approach is to consider different alternative hypotheses for the un-failed an the failed systems. In this paper only additive changes are considered and as was shown in the previous section this will cause a time-varying change in the mean of the innovation sequence. The two alternative hypotheses to consider are:

$$H_0 : \boldsymbol{\epsilon}_t \in \mathcal{N}(\mathbf{0}, \boldsymbol{S}_t), \qquad H_1 : \boldsymbol{\epsilon}_t \in \mathcal{N}(\boldsymbol{\rho}_t(k), \boldsymbol{S}_t)$$

Where H_0 is the hypothesis that no change has occured and H_1 is the hypothesis that a change has occured at time k resulting in a time-varying mean $\boldsymbol{\rho}_t(k)$ of the innovations. Following a statistical approach [1] the appropriate test is to look for a change in mean of a Gaussian sequence.

Motivated by the Neyman-Pearson lemma (see, e.g., [9]), saying that the likelihood ratio is the most powerful test statistics for testing for a change in a distribution, at a given time instant, the following test statistics is considered:

$$l_t(\boldsymbol{\rho}_t(k)) = \log \frac{p(\boldsymbol{\epsilon}_t | H_1(\boldsymbol{\rho}_t(k)))}{p(\boldsymbol{\epsilon}_t | H_0)}. \qquad (10)$$

Here $p(\boldsymbol{\epsilon}_t|H_0)$ is the likelihood for a given $\boldsymbol{\epsilon}_t$ assuming that there is no change and $p(\boldsymbol{\epsilon}_t|H_1(\boldsymbol{\rho}_t(k)))$ is the likelihood for $\boldsymbol{\epsilon}_t$ assuming that its mean is $\boldsymbol{\rho}_t(k)$.

The logarithm of the likelihood ratio is used to simplify the test statistics for Gaussian distributed variables.

The maximum likelihood estimate of the change time for a change of magnitude $\boldsymbol{\rho}_t(k)$ then is

$$\hat{k}^{\mathrm{ML}} = \arg\max_k l_t(\boldsymbol{\rho}_t(k)) \qquad (11)$$

Example 1 *Let us consider the particular case of testing for a known change in the mean of a one dimensional Gaussian distributed sequence.*

$$y_t = \mu_t + \omega_t$$

The problem is to estimate the unknown change time k, when μ_t changes from μ_0 to μ_1, and the hypotheses to consider are:

$$\begin{aligned} H_0 : \quad & y_t \in \mathcal{N}(\mu_0, \sigma^2) \\ H_1 : \quad & y_t \in \mathcal{N}(\mu_1, \sigma^2). \end{aligned}$$

The probability density function for $y_t \in \mathcal{N}(\mu, \sigma^2)$ is:

$$p_\mu(y_t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y_t-\mu)^2}{2\sigma^2}}.$$

This yields the logarithmic likelihood ratio for testing H_1 against H_0 as:

$$\begin{aligned} l_t(k) &= \log \prod_{i=k}^{t} \frac{p_{\mu_1}(y_t)}{p_{\mu_0}(y_t)} \\ &= \frac{\mu_1 - \mu_0}{\sigma^2} \sum_{i=k}^{t} \left(y_i - \frac{\mu_0 + \mu_1}{2} \right) \\ &= \frac{b}{\sigma} \sum_{i=k}^{t} \left(y_i - \mu_0 - \frac{\nu}{2} \right), \end{aligned}$$

where $\nu = \mu_1 - \mu_0$ is the change magnitude and $b = \frac{\nu}{\sigma}$ is the signal-to-noise ratio. The estimated change time is

$$\hat{k}^{ML} = \arg\max_k l_t(k).$$

If we instead want to check if a change has occured, the stopping rule would be defined as:

$$t_a = \arg\min_k (l_t(k) > h).$$

□

The case with a known change magnitude is a very special case, but the example shows the relatively simple test statistics resulting from the logarithmic likelihood ratio. This method can also be interpreted as if ν is the smallest change to detect.

There are a number of change detection algorithms based on this approach and two of the most important ones will be mentioned here showing their appealing simplicity. The algorithms will also be used for comparison with the GLR test later on.

The first algorithm is the Geometric Moving Average (GMA) test [8] based on the following decision function:

$$\begin{aligned} g_t &= \sum_{i=0}^{\infty} \alpha_i \log \frac{p(\epsilon_{t-i}|H_1(k))}{p(\epsilon_{t-i}|H_0))} \\ &= \sum_{i=0}^{\infty} \alpha_i s_{t-i} \end{aligned} \qquad (12)$$

which is a weighted sum of logarithmic likelihood ratios. The weights are chosen to be exponentially decreasing, namely

$$\alpha_i = \gamma^i, 0 < \gamma \leq 1.$$

The decision function can then be written recursively as:

$$g_t = \gamma g_{t-1} + s_t, \qquad (13)$$

and the stopping rule is defined by:

$$t_a = \arg\min_t (g_t > h).$$

The second algorithm that will be mentioned is the cumulative sum (CUSUM) algorithm [7] for which the decision function and the stopping time is defined as:

$$\begin{aligned} g_t &= \max(0, g_{t-1} + s_t - \nu) \\ t_a &= \arg\min_t (g_t > h), \end{aligned} \qquad (14)$$

This can be interpreted as a cumulative sum of logarithmic likelihood ratios with an adaptive threshold or as a Repeated Sequential Probability Ratio test (SPRT) [12]. The decision function g_t will remain zero until there is a $s_t > \nu$ and it will then grow until $s_t < \nu$ again or until $g_t > h$.

Note that the test statistics s_t in both these methods can be either the normalized or the squared normalized innovations.

In the next section we will derive the appropriate test for a time varying bias change, with unknown magnitude, in a Gaussian sequence.

THE GLR TEST

The Generalized Likelihood Ratio (GLR) test was first proposed by Willsky and Jones in [13] and it has been widely used in many different applications. In this section the test is derived and explained. The GLR test is global in time and tests for changes at any past time instant, but it will be shown that it can be restricted to the last M time instants in the next section.

In the previous section it was concluded that the appropriate test statistics was the likelihood ratio as in (10) and the estimated change time was given by (11) if the change magnitude was known. But since the change magnitude is unknown it must be eliminated. This can be achieved by taking the maximum likelihood estimate yielding the GLR test or by marginalization as in [4] yielding the MLR test. The GLR test is thus given by double maximization over $\boldsymbol{\nu}$ and k

$$\hat{\boldsymbol{\nu}}(k) = \arg\max_{\nu} \log \frac{p(\epsilon_t | k, \boldsymbol{\nu})}{p(\epsilon_t | H_0)} \qquad (15)$$

$$\hat{k} = \arg\max_{k} \log \frac{p(\epsilon_t | k, \hat{\boldsymbol{\nu}}(k))}{p(\epsilon_t | H_0)}. \qquad (16)$$

Now the jump candidate $\hat{k}$ is accepted if

$$l_t(\hat{k}, \hat{\boldsymbol{\nu}}(\hat{k})) > h.$$

From the results of (8) and (9) we noted that the original state space problem can be transformed into a linear regression framework. Suppose that we are using information up to time t, then the compact quantities of the least-squares (LS) estimator for the linear regression are given by:

$$\boldsymbol{f}_t(k) = \sum_{i=1}^{t} \boldsymbol{\varphi}_i(k) \boldsymbol{S}_i^{-1} \boldsymbol{\epsilon}_i \qquad (17)$$

$$\boldsymbol{R}_t(k) = \sum_{i=1}^{t} \boldsymbol{\varphi}_i(k) \boldsymbol{S}_i^{-1} \boldsymbol{\varphi}_i^T(k). \qquad (18)$$

From this the maximum likelihood estimate of $\boldsymbol{\nu}$, given the jump instant at k, can be written as:

$$\hat{\boldsymbol{\nu}}(k) = \boldsymbol{R}_t^{-1}(k) \boldsymbol{f}_t(k), \qquad (19)$$

and the test statistics for each k is now given by:

$$\begin{aligned} l_t(k, \hat{\nu}(k)) &= 2 \log \frac{p(\epsilon_t | k, \hat{\nu}(k))}{p(\epsilon_t | H_0)} \\ &= \sum_{i=k+1}^{t} \boldsymbol{\epsilon}_i^T \boldsymbol{S}_i^{-1} \boldsymbol{\epsilon}_i \\ &\quad - (\boldsymbol{\epsilon}_i - \boldsymbol{\varphi}_i^T(k) \hat{\boldsymbol{\nu}}_i(k))^T \boldsymbol{S}_i^{-1} \\ &\quad \cdot (\boldsymbol{\epsilon}_i - \boldsymbol{\varphi}_i^T(k) \hat{\boldsymbol{\nu}}_i(k)) \\ &= \boldsymbol{f}_t^T(k) \boldsymbol{R}_t^{-1}(k) \boldsymbol{f}_t(k), \end{aligned} \qquad (20)$$

where the second equality follows from the fact that $\boldsymbol{\epsilon}_t(k) \in \mathcal{N}(\boldsymbol{\varphi}_{t-1}^T(k)\boldsymbol{\nu}, \boldsymbol{S}_t)$ and the Gaussian probability density function. The third equality follows from (17), (18) and (19). The factor 2 is included for notational convenience.

The test is computed for each k giving the maximum likelihood estimation of the change time as:

$$\hat{k} = \arg\max_{k} l_t(k, \hat{\boldsymbol{\nu}}(k)). \qquad (21)$$

A fault is declared if $l_t(\hat{k}, \hat{\boldsymbol{\nu}}(\hat{k}))$ is larger that a predetermined threshold.

Note that the test is computed for each k making it linearly increasing with time.

The above formulation is compact but it is an off-line expression. In an on-line situation it is more efficient to calculate a recursive least-squares (RLS) estimate of $\hat{\boldsymbol{\nu}}(k)$ as:

$$\hat{\boldsymbol{\nu}}_{t+1}(k) = \hat{\boldsymbol{\nu}}_t(k) + \boldsymbol{L}_{t+1}(\boldsymbol{\epsilon}_{t+1} - \boldsymbol{\varphi}_{t+1}^T \hat{\boldsymbol{\nu}}_t(k)). \qquad (22)$$

with the gain $\boldsymbol{L}_t$ and the estimate error covariance $\boldsymbol{P}_t^{\nu}$ given by:

$$\begin{aligned} \boldsymbol{L}_t &= \boldsymbol{P}_{t-1}^{\nu} \boldsymbol{\varphi}_t [\boldsymbol{\varphi}_t^T \boldsymbol{P}_{t-1}^{\nu} \boldsymbol{\varphi}_t + \boldsymbol{S}_t]^{-1} \\ \boldsymbol{P}_t^{\nu} &= \boldsymbol{P}_{t-1}^{\nu} - \boldsymbol{L}_t \boldsymbol{\varphi}_t^T \boldsymbol{P}_{t-1}^{\nu}, \end{aligned} \qquad (23)$$

with the test statistics now given by:

$$l_t(k, \hat{\boldsymbol{\nu}}(k)) = \boldsymbol{f}_t^T(k) \hat{\boldsymbol{\nu}}_t(k). \qquad (24)$$

The GLR test is now derived and it can be implemented as follows:

Algorithm 1 *Assume that the signal model (5) is given and that a Kalman filter (2) is used.*

At each time step do the following:

1. *Calculate the innovations* $\boldsymbol{\epsilon}_t$ *from the Kalman filter assuming no change.*
2. *Update the regressors* $\boldsymbol{\mu}_t$ *and* $\boldsymbol{\varphi}_t$ *and the quantity* $\boldsymbol{f}_t(k)$ *for each* t *and* $1 \le k \le t$*.*
3. *Compute estimates of the jump magnitudes* $\hat{\boldsymbol{\nu}}_t(k)$ *recursively.*
4. *Compute estimates of the log likelihood ratios*

$$l_t(k) = \boldsymbol{f}_t^T(k)\hat{\boldsymbol{\nu}}_t(k).$$

5. *The jump candidate is now given by*

$$\hat{k} = \arg\max_k l_t(k, \hat{\boldsymbol{\nu}}_t(k)),$$

which should be compared with a predetermined threshold to determine if a jump should be declared.

□

Algorithm 1 is basically the same as proposed by Teunissen [10].

Some remarks should be made about the algorithm.

- The optimal test requires t parallel RLS schemes in addition to the original Kalman filter.
- The test statistic $l_t(k, \hat{\boldsymbol{\nu}}(k))$ at each time instant is distributed as [6] $\chi^2(b, 0)$ under H_0, where b is the size of $\boldsymbol{\nu}$ and as $\chi^2(b, \lambda)$ under H_1, where $\lambda = \boldsymbol{\nu}^T \boldsymbol{R}_t(k)\boldsymbol{\nu}$, making it possible to pre-compute the threshold given the false alarm rate and the probability of missed detection. Note that this is not possible for the total GLR test since it includes multiple hypotheses.
- The choice of the threshold for the total GLR test is tricky since it is connected to the knowledge of the noise variance [4]. This is however a problem that is present with most of the suggested integrity monitoring methods where the test threshold depends on the noise variance.

The latter remark and the fact that there does not exist any explicit formula for computing the threshold for the GLR test makes the design somewhat involved and one has to use simulations or real data to set the threshold.

DECREASING THE COMPLEXITY

Since the implementation of the full GLR test involves a growing bank of filters, it must in some way be restricted. The most convenient way to do this is to restrict the attention to the last $M-1$ time instants, i.e., only considering change times in the interval $t - M < k \le t$. This window can be further reduced by also skipping the last N units of time only considering change times in the interval $t - M < k \le k - N$. The later is also motivated by the fact that the system may not be completely observable for less than N measurements [13] or that the test statistics may be too insensitive for detecting changes if $k > t - N$. The extreme of this reduction is when only $k = t - M + 1$ is considered. The latter cases will however always give an extra delay in the detection.

If only the last time instant is considered the test statistics would be reduced to:

$$\boldsymbol{\epsilon}_t^T \boldsymbol{S}_t^{-1} \boldsymbol{\epsilon}_t, \tag{25}$$

which is the well known χ^2-test, proposed for integrity monitoring in the AIME test [3].

Do however note that the power of the test increases with increasing M.

DIAGNOSIS

When a fault has been detected the next step is to diagnose the cause. The best way to do this, and at the same time decrease the computational complexity and increase the probability for detection, is to constrain the possible change directions to lie among a fixed subset of the states [6] or the measurement variables. So far we have considered possible changes in all the states and measurement variables.

As we saw in (8) and (9) the innovations will be biased as

$$\boldsymbol{\epsilon}_t(k) = \boldsymbol{\epsilon}_t + \boldsymbol{\varphi}_{t-1}^T(k)\boldsymbol{\nu},$$

after a state or measurement variable change, if all change directions are considered.

If we instead only consider a constrained set of possible change directions one could factorize $\boldsymbol{C}_x$ and $\boldsymbol{C}_y$ as:

$$\boldsymbol{C}_x = \boldsymbol{T}_x \boldsymbol{\nu}_x, \qquad \boldsymbol{C}_y = \boldsymbol{T}_y \boldsymbol{\nu}_y, \tag{26}$$

where $\boldsymbol{T}_x$ and $\boldsymbol{T}_y$ are matrices of dimension $n \times b_x$ and $r \times b_y$ respectively, in which the columns are the basis vectors that span the space of all possible change directions and $\boldsymbol{\nu}_x$ and $\boldsymbol{\nu}_y$ will be vectors of dimension $b_x \le n$ and $b_y \le r$ respectively, representing the change magnitudes.

Two typical examples are when only changes in one of the states are considered, making $\boldsymbol{T}_x$ a vector and $\boldsymbol{\nu}_x$ a scalar, or when only changes in one of the measurement variables is considered, making $\boldsymbol{T}_y$ a vector and $\boldsymbol{\nu}_y$ a scalar. Since the fault occur in continuous time we must represent the considered states and measurement variables in the continuous time state space model:

$$\begin{aligned}\dot{\boldsymbol{x}}(t) &= \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)\boldsymbol{u}(t) + \boldsymbol{w}(t) + \\ &\quad \delta(t-t_0)[0\cdots1\cdots0]^T\nu_x \\ \boldsymbol{y}(t) &= \boldsymbol{C}(t)\boldsymbol{x}(t) + \boldsymbol{e}(t) + \\ &\qquad \sigma(t-t_0)[0\cdots1\cdots0]^T\nu_y. \end{aligned} \tag{27}$$

But as the computations are done in discrete time the $\boldsymbol{T}$-vectors must be the sampled equivalent making:

$$\begin{aligned}\boldsymbol{T}_x &= \int_0^{\Delta t} e^{\boldsymbol{A}(t)h}\,dh\cdot[0\cdots1\cdots0]^T \\ \boldsymbol{T}_y &= [0\cdots1\cdots0]^T,\end{aligned}$$

Now the innovation signature of the possible changes could instead be expressed as:

$$\boldsymbol{\epsilon}_t(k) = \boldsymbol{\epsilon}_t + \boldsymbol{\varphi}_{t-1}^T(k)\boldsymbol{T}\boldsymbol{\nu}. \tag{28}$$

The test statistics will change accordingly to be:

$$\begin{aligned}\boldsymbol{f}_t(k)' &= \boldsymbol{T}^T\boldsymbol{f}_N(k) \\ \boldsymbol{R}_t(k)' &= \boldsymbol{T}^T\boldsymbol{R}_N(k)\boldsymbol{T},\end{aligned} \tag{29}$$

or when computed "on-line" the estimate of $\hat{\boldsymbol{\nu}}$ should be computed as:

$$\hat{\boldsymbol{\nu}}_{t+1}(k) = \hat{\boldsymbol{\nu}}_t(k) + \boldsymbol{L}_t'(\boldsymbol{\epsilon}_t - \boldsymbol{\varphi}_t^T\boldsymbol{T}\hat{\boldsymbol{\nu}}_t(k). \tag{30}$$

The constraint will increase the "signal-to-noise" ratio of the GLR test and for a given probability of false alarm the probability for detection will increase. If multiple cases are considered the one that yields the highest test statistics, $l_t(\hat{k},\hat{\boldsymbol{\nu}}(\hat{k}))$ should be chosen as the most likely fault.

ADAPTATION OF THE ESTIMATES

When a satellite range bias drift is detected, an estimate of the change magnitude and direction will be given by (22), with corresponding error covariance given by (23), making it possible to identify and exclude the faulty satellite or a set of satellites including the faulty one. This exclusion should also be accompanied by a statistical measure of the probability of false exclusion.

In addition to the exclusion we would also like to adapt the Kalman filter estimate as quickly and correctly as possible. This is possible in a straightforward manner without the need for smoothing the estimates between the fault onset time and the present time.

If we consider a change in the measurements it can be seen from (6) and (7) that if we exactly know the time instant k of a measurement variable change $\boldsymbol{\nu}_y$ the estimate should be corrected as:

$$\begin{aligned}\hat{\boldsymbol{x}}_{t|t}^{corr} &= \hat{\boldsymbol{x}}_{t|t} + \boldsymbol{\alpha}_t(k) - \boldsymbol{\beta}_t(k) \\ &= \hat{\boldsymbol{x}}_{t|t} - \boldsymbol{\mu}_t(k)\boldsymbol{\nu}_y,\end{aligned} \tag{31}$$

since $\boldsymbol{\alpha}_t(k)$ equals zero in this case. If we have good estimates of k and $\boldsymbol{\nu}_x$ they can be used to correct the state estimates. The uncertainty of the estimate shall be reflected by increasing the state estimate error covariance by:

$$\boldsymbol{P}_{t|t}^{corr} = \boldsymbol{P}_{t|t} + \boldsymbol{\mu}_t(k)\boldsymbol{P}_t^{\nu}\boldsymbol{\mu}_t(k)^T \tag{32}$$

For a state bias change there is no exclusion, but rather an adaption to the new state level. From (6) and (7) it can be seen that if we exactly know the time instant k of a state change with magnitude $\boldsymbol{\nu}_x$ the state estimate should be corrected as:

$$\begin{aligned}\hat{\boldsymbol{x}}_{t|t}^{corr} &= \hat{\boldsymbol{x}}_{t|t} + \boldsymbol{\alpha}_t(k) - \boldsymbol{\beta}_t(k) \\ &= \hat{\boldsymbol{x}}_{t|t} + \left(\prod_{i=k}^{t-1}\boldsymbol{F}_i - \boldsymbol{\mu}_t(k)\right)\boldsymbol{\nu}_x.\end{aligned} \tag{33}$$

The state estimate covariance should also be increased by:

$$\begin{aligned}&\boldsymbol{P}_{t|t}^{corr} = \boldsymbol{P}_{t|t} + \\ &\left(\prod_{i=k}^{t-1}\boldsymbol{F}_i - \boldsymbol{\mu}_t(k)\right)\boldsymbol{P}_t^{\nu}\left(\prod_{i=k}^{t-1}\boldsymbol{F}_i - \boldsymbol{\mu}_t(k)\right)^T.\end{aligned} \tag{34}$$

When using methods that do not calculate an estimate of the fault magnitude one have to proceed differently. If there are redundant measurement sources one should exclude the faulty one and re-process the affected data. For a state change the best thing to do is to increase the covariance matrix of the Kalman filter which helps the filter to track the new state level faster.

COMPARISON WITH OTHER METHODS

To demonstrate the performance of the discussed algorithms they are compared with the commonly used χ^2-test (25).

A fault detection procedure would be considered as optimal if it has the shortest mean delay for detection for a given false alarm rate. In order to compare the different methods the Average Runtime Length (ARL) function has been computed by a Monte Carlo simulation of an INS-GNSS system. The ARL(ν) function is the mean delay for detection versus the change magnitude ν. Note that ARL(0) is the false alarm rate.

The probability of missed detection is also a good performance measure, why that also have been computed for the different cases.

The false alarm rate has for all methods been set to 2 per hour, which has been used to set the test thresholds using simulations or explicit expressions, when available. The unrealistic value of 2 false alarms per hour has been used to decrease the computational time since for some of the methods there is no explicit expression for setting the threshold.

To decrease the computation time in the Monte Carlo simulation only the north horizontal channel of an integrated INS-GNSS system is studied instead of all three dimensions. The acceleration and gyro errors are assumed to consist of white noise plus a first order Gauss-Markov process for the acceleration bias and the gyro drift. The state-space description is chosen to include states for accelerometer bias, gyro drift and time correlated noise from the GNSS, giving the continuous state space equation as:

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{w}(t) \\ \boldsymbol{y}(t) &= \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{e}(t) \end{aligned} \tag{35}$$

where

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{g}{R} & \frac{1}{R} & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{\tau_a} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{\tau_g} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_{gps}} \end{bmatrix}$$

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Corresponding to the following set of states:

$$\begin{bmatrix} \text{position error} \\ \text{position rate error} \\ \text{east angular error} \\ \text{acceleration bias} \\ \text{gyro bias} \\ \text{correlated measurement noise} \end{bmatrix}$$

Since high performance applications are considered small errors has been simulated for the satellite navigation system corresponding to differential GNSS or military systems. The performance data that was used in order to simulate the INS and the GNSS are given in Table 1. The Kalman filter is executed at 1 Hz. The influence of the initial state was allowed to die out before the acceleration bias or the GNSS position error changed.

The following methods have been compared:

1. The χ^2-test with using only the last time instant and summed over the last 10 instants.

2. The GMA test (13), both with normalized and squared normalized innovations as test statistics and with $\gamma = 0.1$ and 0.05.

		INS	
Gyro	bias	0.004	deg/hr
	corr noise	0.001	deg/$\sqrt{\text{hr}}$
	time const	3600	s
	white noise	0.001	deg/$\sqrt{\text{hr}}$
Accel.	bias	50	μg
	corr noise	5	μg/$\sqrt{\text{Hz}}$
	time const	7200	s
	white noise	5	μg/$\sqrt{\text{Hz}}$
		GNSS	
	corr noise	1.5	m
	time const	10	s
	white noise	1	m

Table 1 INS and GNSS performance used in simulations

3. The CUSUM test (14), both with normalized (ν = 12,13, & 13.5) and squared normalized innovations (ν = 9, 11 & 12) as test statistics.

4. The GLR test as described in the last sections, windowed with M = 10 and 20. The considered faults have not been constrained but changes in the full state vector nave been considered.

They have all been tested on the following changes:

1. Accelerometer bias jump with eight different jump magnitudes between 450 μg and 1 g.

2. Measurement drift with eight different drift rates between 0.5 m/s and 34 m/s.

100 different noise realizations were simulated for each change magnitude.

The ARL function from the simulations with the accelerometer bias jump is plotted in Figure 3 and

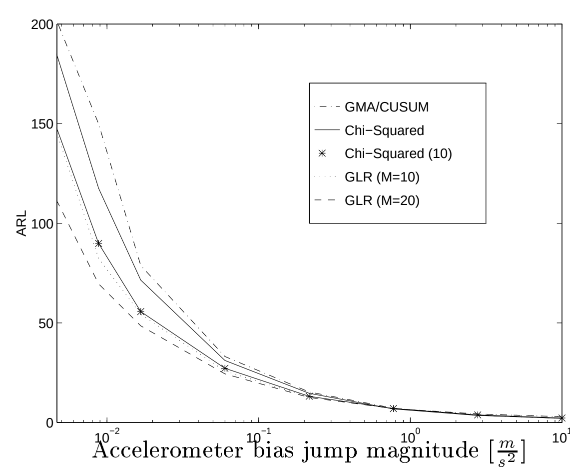

Fig. 3. Average delay in detection of state bias change.

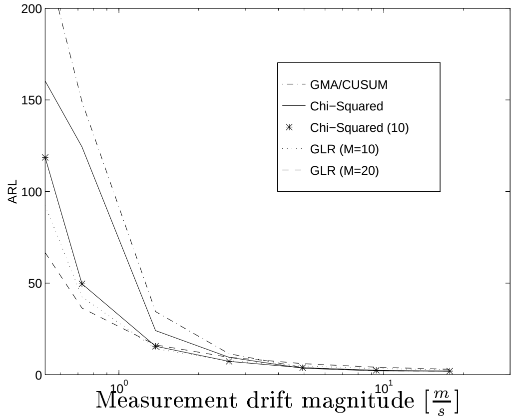

Fig. 4. Average delay in detection of measurement drift.

Acceleration bias jump		
test method	jump magnitude 450 μg	
GMA/CUSUM	44	
χ^2	17	
χ^2 (10)	3	
GLR (M=10)	0	
GLR (M=20)	0	
Measurement drift		
test method	drift magnitude	
	0.5 m/s	0.7 m/s
GMA/CUSUM	72	51
χ^2	43	13
χ^2 (10)	22	0
GLR (M=10)	24	0
GLR (M=20)	12	0

Table 2 Number of missed detections in simulations

for the measurement drift in Figure 4. Note that the horizontal axes are scaled logarithmically. The different GMA and CUSUM tests performed almost identically and worse than the other tests why only the one that performed best is given for comparison. As can be seen the GLR-test performed best of the test methods on small change magnitudes and it performed slightly better than the χ^2-test when the same number (10) of time instants were used. The GLR test with $M = 20$ did however get a slightly larger ARL for detection of measurement drifts with larger magnitudes. That is a prize one may have to pay for the advantages of diagnosis and adaption. The GLR test did not miss any detections of accelerometer bias jumps, but the other methods did with the smallest jump magnitude, as can be seen in Table 2. In the second case with the measurement change all methods did miss some detections with the smallest drift magnitudes.

The conclusion that can be drawn from these simulations is that if one is looking for small or slowly drifting faults the test would be more powerful if multiple time instants are considered. If multiple time instants are considered the use of the GLR test will help in decreasing the delay for detection.

These simulation have only compared the detection phase but with the GLR test one would also yield a method for diagnosis and adaption.

CONCLUSIONS

The use of integrated INS - GNSS navigation systems have been proposed for solving the integrity requirements of high precision navigation applications.

The position estimate in an integrated navigation system is usually calculated using an error state Kalman filter. By using the Kalman filter innovations as test statistics all past and present measurement are used to monitor the integrity. This means that there always will be enough redundancy for integrity monitoring and the need of a minimum number of satellites and a sufficient geometric constellation are circumvented.

The use of integrity monitoring methods based on hypothesis testing using the likelihood ratio and the use of specific tests for each type of considered fault are recommended.

The most powerful of the studied integrity monitoring methods is the Generalized Likelihood Ratio (GLR) test which uses the innovations of the Kalman filter to compute the maximum likelihood estimate of the change time and change magnitude. The GLR test then uses these estimates to compute the logarithmic likelihood of a fault versus no fault.

A comparison of the described integrity monitoring methods with the χ^2-test is performed by a Monte Carlo simulation showing that the GLR test is best in detecting small or slowly growing faults even when only a small and constant number of matched filters are used.

A further advantage of the GLR test is that, in addition to detecting the occurrence of a fault, it also estimates the change magnitude and change time, making it possible to identify the source of the fault, exclude faulty satellites and correct the Kalman filter estimates, without re-processing the affected data.

ACKNOWLEDGMENTS

This work was supported by NUTEK's Center of Excellence ISIS and Saab Military Aircraft. NUTEK is the Swedish National Board for Industrial and Technical Development. The author also which to thank Fredrik Gustafsson for valuable suggestions and comments.

REFERENCES

[1] M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. Information and System Sciences Series. Prentice-Hall, Englewood Cliffs, N.J. USA, 1993.

[2] R. G. Brown. A Baseline GPS RAIM Scheme and a Note on the Equivalence of Three RAIM Methods. *NAVIGATION: Journal of the Institute of Navigation*, 39(3):301–316, Fall 1992.

[3] J. Diesel and S. Luu. GPS/IRS AIME: Calculation of Thresholds and Protection Radius Using Chi-Squared Methods. In *Proc. of ION GPS-95*, volume 2, pages 1959–1964. ION, 1995.

[4] F. Gustafsson. The Marginalized Likelihood Ratio Test for Detection Abrupt Changes. *IEEE Transactions on Automatic Control*, 41(1):66–78, January 1996.

[5] P-Y. Ko, P. Enge, and D. J. Powell. Continuity Improvements via Inertial Augumentation of GPS-Based Landing System. In *Proc. of IEEE 1996 Position, Location and Navigation Symposium*, pages 153–159. IEEE, 1996.

[6] J. S. Liu and H. L. Jones. Linear Manifold Constrained GLR. *IEEE Transactions on Automatic Control*, AC-22(6):988–989, December 1977.

[7] E. S. Page. Continuous Inspection Schemes. *Biometrika*, 41:100–115, 1954.

[8] S. W. Roberts. Control Charts Based on Geometric Moving Averages. *Technometrics*, 1:239–250, 1959.

[9] L. L. Scharf. *Statistical Signal Processing: Detection, Estimation and Time Series Analysis.* Addison-Wesley, 1991.

[10] P. J. G. Teunissen. An Integrity and Quality Control Procedure for Use in Multi Sensor Integration. In *Proc. of ION GPS-90*, pages 513–522. ION, 1990.

[11] S. Vieweg and C. Butzmuehlen. Solving Integrity Insufficiencies of Sat Nav Systems with Complementary Navigation Information. In *Proc. of ION GPS-95*, volume 2, pages 2005–2013. ION, 1995.

[12] A. Wald. *Statistical Decision Functions.* Wiley, New York, 1950.

[13] A. S. Willsky and H. L. Jones. A Generalized Likelihood Ratio Approach to State Estimation in Linear Systems Subjected to Abrupt Changes. In *Proc. of the 1974 IEEE Conference on Decision and Control*, pages 846–853, November 1974.