Risking to underestimate the integrity risk

Zaminpardaz, Safoora; Teunissen, Peter J.G.; Tiberius, Christiaan C.J.M.

DOI
10.1007/s10291-018-0812-0

Publication date
2019

Document Version
Final published version

Published in
GPS Solutions

GPS Solutions (2019) 23:29
https://doi.org/10.1007/s10291-018-0812-0

ORIGINAL ARTICLE

Risking to underestimate the integrity risk

Safoora Zaminpardaz[1] · Peter J. G. Teunissen[2,3] · Christiaan C. J. M. Tiberius[3]

Received: 16 January 2018 / Accepted: 7 December 2018

Abstract
As parameter estimation and statistical testing are often intimately linked in the processing of observational data, the uncertainties involved in both estimation and testing need to be properly propagated into the final results produced. This necessitates the use of conditional distributions when evaluating the quality of the resulting estimator. As the conditioning should be on the identified hypothesis as well as on the corresponding testing outcome, omission of the latter will result in an incorrect description of the estimator's distribution. In this contribution, we analyse the impact this omission or approximation has on the considered distribution of the estimator and its integrity risk. For a relatively simple observational model it is mathematically proven that the rigorous integrity risk exceeds the approximation for the contributions under the null hypothesis, which typically has a much larger probability of occurrence than an alternative. Actual GNSS-based positioning examples confirm this finding. Overall we observe a tendency of the approximate integrity risk being smaller than the rigorous one. The approximate approach may, therefore, provide a too optimistic description of the integrity risk and thereby not sufficiently safeguard against possibly hazardous situations. We, therefore, strongly recommend the use of the rigorous approach to evaluate the integrity risk, as underestimating the integrity risk in practice, and also the risk to do so, cannot be acceptable particularly in critical and safety-of-life applications.

Keywords Detection, identification and adaptation (DIA) · DIA estimator · Integrity risk · Statistical testing · Conditional distribution

Introduction

The DIA method for the detection, identification and adaptation of model misspecifications combines estimation with testing. Parameter estimation is conducted to determine estimates for the parameters of interest, and statistical testing is conducted to validate the results with the aim of removing any unwanted biases that may be present. To rigorously capture the estimation–testing combination, the DIA estimator has recently been introduced by Teunissen (2017) together with a unifying probabilistic framework. This allows one to take into account the intricacies of the combination when evaluating the contributions of the decisions and estimators involved. Procedures followed in practice are usually conditional ones implying that the quality and the performance of the resulting estimator must be described based on its subsequent conditional distribution. Hence, employing the distribution of the estimator under an identified hypothesis without regard to the conditioning process that led to the decision of accepting this hypothesis may impact the quality description of the resulting estimator in terms of precision, unbiasedness, confidence region and integrity risk.

In this contribution, we turn our attention specifically to the integrity risk, which—in short—is the probability that, whatever hypothesis is true in reality, the estimator of unknown parameters, directed by the testing outcome, is outside an acceptable area or volume around its targeted value. The estimator's integrity risk is thus one minus its confidence level. As integrity plays a crucial role in critical and safety-of-life applications, for instance in aviation, when GNSS positioning is used to fly an approach to an airport, stringent requirements on integrity obviously apply.

✉ Peter J. G. Teunissen
p.j.g.teunissen@tudelft.nl

[1] School of Science, RMIT University, GPO Box 2476V, Melbourne, VIC 3001, Australia

[2] GNSS Research Centre, Curtin University, GPO Box U1987, Perth, WA 6845, Australia

[3] Department of Geoscience and Remote Sensing (DIA), Delft University of Technology, PO Box 5048, 2600 GA Delft, The Netherlands

Published online: 09 January 2019

We compare, using a number of GNSS positioning examples, the integrity risk using unconditional distributions with the one obtained by rigorous evaluation of the correct conditional ones. We demonstrate that with the approximate approach of using unconditional distributions to evaluate the integrity risk, for instance when accounting for the event of removing from the solution an observation identified as faulty, one may obtain too optimistic figures, and thereby compromise the whole concept of integrity. The actual integrity risk, evaluated using the correct conditional distributions, may, in fact, be significantly larger.

This contribution is organized as follows. We start with a brief review of the detection, identification, and adaptation procedure, including the DIA estimator and its statistical distribution. Next, the integrity risk is defined, rigorous as well as approximate, with the latter following from neglecting the conditioning on the testing outcome. We then demonstrate in graphical form, using a simple observational model with just a single unknown parameter, both the unconditional and conditional distributions, so that the different contributions to the integrity risk, as well as the differences between the two approaches, are understood. We also prove in this section that, under the null hypothesis, the rigorous integrity risk always exceeds the approximate one. The integrity risk comparison is then continued, but now for a number of actual satellite-based single-point positioning examples. These findings show that one indeed runs a serious risk of underestimating the actual integrity risk (or overestimating the confidence level) when using the unconditional distributions instead of the conditional ones. We hereby note that our findings, although demonstrated by means of an application of the DIA procedure, are equally valid for any other method of fault detection and exclusion, and, therefore, hold true for a wide variety of different applications, such as geodetic quality control (Kösters and van der Marel 1990; Amiri Simkooei 2001; Perfetti 2006), navigational integrity (Teunissen 1990b; Gillissen and Elema 1996; Yang et al. 2014), structural health integrity (Verhoef and de Heus 1995; Yavaşoğlu et al. 2017; Durdag et al. 2018), and integrity monitoring of GNSS (Jonkman and de Jong 2000; Kuusniemi et al. 2004; Hewitson and Wang 2006). Finally, a summary with conclusions are presented.

Detection, identification and adaptation (DIA)

A brief recap of the DIA-datasnooping procedure is provided, and the DIA estimator introduced in Teunissen (2017) is presented. Then an inventory of all possible testing decisions is compiled, and the distribution of the DIA estimator is decomposed into contributions, conditioned on the testing outcome and the hypothesis.

Statistical hypotheses

We first formulate the null- and alternative hypotheses, denoted by $\mathcal{H}_0$ and $\mathcal{H}_\mathrm{i}$, respectively. Throughout the paper, as alternative hypotheses, we consider those describing outliers in individual observations. Here we restrict ourselves to the case of one outlier at a time. In that case there are as many alternative hypotheses as there are observations. Therefore, the observational model under $\mathcal{H}_0$ and $\mathcal{H}_\mathrm{i}$ is given as

$$\mathcal{H}_0 : E(y) = Ax; \quad D(y) = Q_{yy} \tag{1}$$

$$\mathcal{H}_\mathrm{i} : E(y) = Ax + c_i b_i; \quad D(y) = Q_{yy} \tag{2}$$

with $E(\cdot)$ the expectation operator, $D(\cdot)$ the dispersion operator, $y \in \mathbb{R}^m$ the normally distributed random vector of observables linked to the estimable unknown parameters through the design matrix $A \in \mathbb{R}^{m\times n}$ of $\mathrm{rank}(A) = n$, and $Q_{yy} \in \mathbb{R}^{m\times m}$ the positive-definite variance matrix of y. The redundancy of $\mathcal{H}_0$ is $r = m - \mathrm{rank}(A) = m - n$. c_i is the canonical unit vector having one as its ith entry and zeros elsewhere, and b_i is the scalar bias. Note that $[A\ c_i]$ is a known matrix of full rank. As the number of observations is equal to m, there are also m alternative hypotheses $\mathcal{H}_i$ defined in (2); $i = 1, \ldots, m$.

The best linear unbiased estimator (BLUE) of the unknown parameters x is given by

$$\mathcal{H}_0 : \hat{x}_0 = A^+ y \quad \mathcal{H}_{i\neq 0} : \hat{x}_i = \bar{A}_i^+ y \tag{3}$$

with $A^+ = (A^T Q_{yy}^{-1} A)^{-1} A^T Q_{yy}^{-1}$ the BLUE-inverse of A, $\bar{A}_i^+ = (\bar{A}_i^T Q_{yy}^{-1} \bar{A}_i)^{-1} \bar{A}_i^T Q_{yy}^{-1}$ the BLUE-inverse of $\bar{A}_i = P_{c_i}^\perp A$ and $P_{c_i}^\perp = I_m - c_i (c_i^T Q_{yy}^{-1} c_i)^{-1} c_i^T Q_{yy}^{-1}$ being an orthogonal projector that projects onto the orthogonal complement of the range space of c_i.

DIA-datasnooping procedure

The DIA method has been widely employed in a variety of applications, such as the quality control of geodetic networks and the integrity monitoring of GNSS models, see, e.g., Teunissen (1990a) and Amiri Simkooei (2001). The DIA steps are realized using the *misclosure* vector $t \in \mathbb{R}^r$ given as

$$t = B^T y; \quad Q_{tt} = B^T Q_{yy} B \tag{4}$$

where the $m \times r$ matrix B is a full-rank matrix, with $\mathrm{rank}(B) = r$, of which the range space is an orthogonal complement of that of A, i.e., $[A\ B] \in \mathbb{R}^{m\times m}$ is invertible and $A^T B = 0$. Assuming that the measurement errors are normally distributed, i.e., $y \overset{\mathcal{H}_i}{\sim} \mathcal{N}\left(Ax + c_i b_i, Q_{yy}\right)$ for $i = 0, 1, \ldots, m$ and $c_0 b_0 = 0$, the misclosure vector is then distributed as

$$t \overset{\mathcal{H}_i}{\sim} \mathcal{N}\left(\mu_{t_i} = B^T c_i b_i, Q_{tt}\right) \quad \text{for} \quad i = 0, 1, \ldots, m \tag{5}$$

As t is zero-mean under $\mathcal{H}_0$ and also independent of $\hat{x}_0$, it provides all the available information useful for validation of $\mathcal{H}_0$ (Teunissen 2017). Thus, an unambiguous testing procedure can be established through assigning the outcomes of t to the statistical hypotheses $\mathcal{H}_i$ for $i = 0, 1, \ldots, m$.

The DIA-datasnooping procedure is specified as follows.

Detection: Accept $\mathcal{H}_0$if $t \in \mathcal{P}_0$ with

$$\mathcal{P}_0 = \left\{ t \in \mathbb{R}^r \middle| ||t||^2_{Q_{tt}} \leqslant k_{\alpha,r} \right\} \tag{6}$$

in which $||\cdot||^2_{Q_{tt}} = (\cdot)^T Q_{tt}^{-1}(\cdot)$ and $k_{\alpha,r}$ is the α-percentage of the central Chi-square distribution with r degrees of freedom. If $\mathcal{H}_0$ is accepted, then $\hat{x}_0$is provided as the estimate of x. Otherwise, go to next step.

Identification: Compute Baarda's test statistic for all alternative hypotheses as (Baarda 1967; Teunissen 2000)

$$w_i = \frac{c_{t_i}^T Q_{tt}^{-1} t}{\sqrt{c_{t_i}^T Q_{tt}^{-1} c_{t_i}}} \tag{7}$$

in which $c_{t_i} = B^T c_i$ is the ith column of matrix B^T since c_i is a canonical unit vector. Select $\mathcal{H}_{i\neq 0}$ if $t \in \mathcal{P}_{i\neq 0}$ with

$$\mathcal{P}_{i\neq 0} = \left\{ t \in \mathbb{R}^r / \mathcal{P}_0 \middle| |w_i| = \max_{j\in\{1,\ldots,m\}} |w_j| \right\} \tag{8}$$

Adaptation: When $\mathcal{H}_i$ is selected, then $\hat{x}_i$ is provided as the estimate of x.

The partitioning $\mathcal{P}_i$ in terms of the (original) misclosure vector is introduced in Teunissen (2017), and an example is shown in Fig. 3 in [Ibid]. Note that the above datasnooping partitioning would need modification in case of 'iterated datasnooping' (Kok et al. 1984) for multiple outlier testing which involves consecutive rounds of detection and identification until no further outlier is detected.

DIA estimator

Given the above three steps, estimation and testing are *combined* in DIA-datasnooping. Teunissen (2017) presents a unifying framework to rigorously capture the probabilistic properties of this combination, see also Teunissen et al. (2017). As such, the *DIA estimator* $\bar{x}$ was introduced, which captures the whole estimation–testing scheme and it is given as

$$\bar{x} = \sum_{j=0}^{m} \hat{x}_j p_j(t) \tag{9}$$

with $p_j(t)$ being the indicator function of region $\mathcal{P}_j$, i.e., $p_j(t) = 1$ for $t \in \mathcal{P}_j$ and $p_j(t) = 0$ elsewhere. Therefore, the DIA estimator $\bar{x}$ is a combination of $\hat{x}_j$ for $j = 0, 1, \ldots, m$ and the misclosure vector t. The probability density function (PDF) of $\bar{x}$ under $\mathcal{H}_i$ reads (Teunissen 2017)

$$\begin{aligned} f_{\bar{x}}(\theta|\mathcal{H}_i) &= \sum_{j=0}^{m} \int_{\mathcal{P}_j} f_{\hat{x}_j,t}(\theta, \tau|\mathcal{H}_i)\, \mathrm{d}\tau \\ &= \sum_{j=0}^{m} \int_{\mathcal{P}_j} f_{\hat{x}_0}(\theta + L_j\tau|\mathcal{H}_i) f_t(\tau|\mathcal{H}_i)\, \mathrm{d}\tau \end{aligned} \tag{10}$$

where the second equality is a consequence of

$$\hat{x}_j = \hat{x}_0 - L_j t \tag{11}$$

with

$$L_j = \frac{1}{||c_{t_j}||^2_{Q_{tt}}} A^+ c_j c_{t_j}^T Q_{tt}^{-1}. \tag{12}$$

Testing decisions

As was shown above, the decisions of the testing procedure are driven by the outcome of the misclosure vector t. If $\mathcal{H}_i$ is true, then the decision is correct if $t \in \mathcal{P}_i$, and wrong if $t \in \mathcal{P}_{j\neq i}$. We, therefore, discriminate between the following events

$$\begin{aligned} \mathrm{CA} &= (t \in \mathcal{P}_0|\mathcal{H}_0) = \text{correct acceptance} \\ \mathrm{FA} &= (t \notin \mathcal{P}_0|\mathcal{H}_0) = \text{false alarm} \\ \mathrm{MD}_i &= (t \in \mathcal{P}_0|\mathcal{H}_i) = \text{missed detection} \\ \mathrm{CD}_i &= (t \notin \mathcal{P}_0|\mathcal{H}_i) = \text{correct detection} \\ \mathrm{WI}_i &= (t \in \cup_{j\neq 0,i}\mathcal{P}_j|\mathcal{H}_i) = \text{wrong identification} \\ \mathrm{CI}_i &= (t \in \mathcal{P}_i|\mathcal{H}_i) = \text{correct identification} \end{aligned} \tag{13}$$

With $* = \{\mathrm{CA}, \mathrm{FA}, \mathrm{MD}_i, \mathrm{CD}_i, \mathrm{WI}_i, \mathrm{CI}_i\}$, we denote the probability of $*$ by P_*. satisfying

$$\begin{aligned} \mathrm{P}_{\mathrm{CA}} + \mathrm{P}_{\mathrm{FA}} &= 1, \\ \mathrm{P}_{\mathrm{MD}_i} + \mathrm{P}_{\mathrm{CD}_i} &= 1, \\ \mathrm{P}_{\mathrm{WI}_i} + \mathrm{P}_{\mathrm{CI}_i} &= \mathrm{P}_{\mathrm{CD}_i} \end{aligned} \tag{14}$$

Computation of P_* requires information about the misclosure PDF which is given in (5). Here, it is important to note the difference between the CD- and CI-probability, i.e.$\mathrm{P}_{\mathrm{CD}_i} \geqslant \mathrm{P}_{\mathrm{CI}_i}$. They would be the same if there were only one single alternative hypothesis, say $\mathcal{H}_1$, since then $\mathcal{P}_1 = \mathbb{R}^r / \mathcal{P}_0$.

Decomposition of $f_{\bar{x}}(\theta|\mathcal{H}_i)$

Given the events in (13) and using the total probability rule, $f_{\bar{x}}(\theta|\mathcal{H}_i)$ can be decomposed as follows:

$$f_{\bar{x}}(\theta|\mathcal{H}_i) = \sum_{j=0}^{m} f_{\bar{x}|t\in\mathcal{P}_j}\big(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\big)\,\mathrm{P}\big(t\in\mathcal{P}_j|\mathcal{H}_i\big) \tag{15}$$

where

$$f_{\bar{x}|t\in\mathcal{P}_j}\big(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\big) = f_{\hat{x}_j|t\in\mathcal{P}_j}\big(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\big) \tag{16}$$

Therefore, in case $j=0$, due to $\hat{x}_0$ being independent of t, we have

$$f_{\bar{x}|t\in\mathcal{P}_0}\big(\theta|t\in\mathcal{P}_0,\mathcal{H}_i\big) = f_{\hat{x}_0}\big(\theta|\mathcal{H}_i\big) \tag{17}$$

The proof of (16) and (17) is given as follows. The conditional PDF $f_{\bar{x}|t\in\mathcal{P}_j}\big(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\big)$ is obtained through the following general expression

$$f_{\bar{x}|t\in\mathcal{P}_j}\big(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\big) = \int_{\mathcal{P}_j}\frac{f_{\bar{x},t}(\theta,\tau|\mathcal{H}_i)}{\mathrm{P}(t\in\mathcal{P}_j|\mathcal{H}_i)}\mathrm{d}\tau \tag{18}$$

in which (Teunissen 2017)

$$f_{\bar{x},t}\big(\theta,\tau|\mathcal{H}_i\big) = \sum_{j=0}^{m} f_{\hat{x}_j,t}\big(\theta,\tau|\mathcal{H}_i\big)p_j(\tau) \tag{19}$$

Substituting (19) into (18) gives

$$f_{\bar{x}|t\in\mathcal{P}_j}\big(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\big) = \int_{\mathcal{P}_j}\frac{f_{\hat{x}_j,t}\big(\theta,\tau|\mathcal{H}_i\big)}{\mathrm{P}\big(t\in\mathcal{P}_j|\mathcal{H}_i\big)}\mathrm{d}\tau \tag{20}$$

Comparing the structure of (20) with that of (18), we achieve (16). For $j=0$, we have $f_{\hat{x}_0,t}\big(\theta,\tau|\mathcal{H}_i\big) = f_{\hat{x}_0}(\theta|\mathcal{H}_i)f_t\big(\tau|\mathcal{H}_i\big)$ of which the substitution into (20) gives (17). Note that the conditional PDFs $f_{\hat{x}_j|t\in\mathcal{P}_j}\big(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\big)$ for $j\neq 0$ are *non-normal*, which is further discussed in the following sections.

Integrity risk

Being an estimator of x, the DIA estimator $\bar{x}$ is likely to be considered a good estimator if it is close to x with a sufficiently large probability. Defining 'closeness' as 'lying in an x-centered region $\mathcal{B}_x$', and 'sufficiently large' as '$1-\varepsilon$' for a very small ε, then $\bar{x}$ is an acceptable estimator of x if $\mathrm{P}(\bar{x}\in\mathcal{B}_x)\geqslant 1-\varepsilon$ or equivalently

$$\mathrm{P}(\bar{x}\in\mathcal{B}_x^c) < \varepsilon \tag{21}$$

with $\mathcal{B}_x^c = \mathbb{R}^n/\mathcal{B}_x$. We refer to the above probability as integrity risk, see also, e.g., Schuster et al. (2007) and Salós et al. (2010). The integrity risk (21) is thus one minus the probability of $\bar{x}$ lying inside the confidence region $\mathcal{B}_x$. In the sequel, we denote this probability by $\overline{\mathrm{IR}}$.

Assuming that $\sum_{j=0}^{m}\mathrm{P}\big(\mathcal{H}_j\big)=1$, with $\mathrm{P}\big(\mathcal{H}_j\big)$ being the probability of occurrence of $\mathcal{H}_j$, describing a bias in the single jth measurement, cf. (2), the integrity risk can be decomposed, using the total probability rule, as

$$\overline{\mathrm{IR}} = \sum_{j=0}^{m}\mathrm{P}\big(\bar{x}\in\mathcal{B}_x^c|\mathcal{H}_j\big)\mathrm{P}\big(\mathcal{H}_j\big) \tag{22}$$

Here it is important to realize that the above expression depends on the bias value b_i under $\mathcal{H}_i$ for $i=1,\ldots,m$. One may take a conservative route by computing each term in the summation for a bias value b_i which maximizes the product $\mathrm{P}\big(\bar{x}\in\mathcal{B}_x^c|\mathcal{H}_i\big)\mathrm{P}\big(\mathcal{H}_i\big)$. The probabilities $\mathrm{P}\big(\bar{x}\in\mathcal{B}_x^c|\mathcal{H}_i\big)$ in the summation (22), using (15) and (16), can be expressed as

$$\overline{\mathrm{IR}}|\mathcal{H}_i = \mathrm{P}\big(\bar{x}\in\mathcal{B}_x^c|\mathcal{H}_i\big) = \sum_{j=0}^{m}\mathrm{P}\big(\hat{x}_j\in\mathcal{B}_x^c|t\in\mathcal{P}_j,\mathcal{H}_i\big)\mathrm{P}\big(t\in\mathcal{P}_j|\mathcal{H}_i\big) \tag{23}$$

The conditional probabilities in the above equation are computed based on the PDFs $f_{\hat{x}_j|t\in\mathcal{P}_j}\big(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\big)$.

Would one neglect the correlation between $\hat{x}_j$ and t, and use the unconditional PDFs $f_{\hat{x}_j}\big(\theta|\mathcal{H}_i\big)$ instead, an approximation of the rigorous integrity risk $\overline{\mathrm{IR}}|\mathcal{H}_i$ is obtained as

$$\mathrm{IR}^o|\mathcal{H}_i = \sum_{j=0}^{m}\mathrm{P}\big(\hat{x}_j\in\mathcal{B}_x^c|\mathcal{H}_i\big)\mathrm{P}\big(t\in\mathcal{P}_j|\mathcal{H}_i\big) \tag{24}$$

In $\big(\overline{\mathrm{IR}}|\mathcal{H}_i\big)$, one conditions on both the hypothesis and the testing outcome, while in $\big(\mathrm{IR}^o|\mathcal{H}_i\big)$, one conditions only on the hypothesis and *not* on the testing outcome.

The difference between the rigorous and approximate integrity risks under $\mathcal{H}_i$ reads

$$\begin{aligned}\overline{\mathrm{IR}}|\mathcal{H}_i - \mathrm{IR}^o|\mathcal{H}_i = \sum_{j=0}^{m}\big[&\mathrm{P}\big(\hat{x}_j\in\mathcal{B}_x^c|t\in\mathcal{P}_j,\mathcal{H}_i\big)\\ &-\mathrm{P}\big(\hat{x}_j\in\mathcal{B}_x^c|\mathcal{H}_i\big)\big]\mathrm{P}\big(t\in\mathcal{P}_j|\mathcal{H}_i\big)\end{aligned} \tag{25}$$

Note, in the above summation, that j runs from 1 to m as $\mathrm{P}\big(\hat{x}_0\in\mathcal{B}_x^c|t\in\mathcal{P}_0,\mathcal{H}_i\big) = \mathrm{P}\big(\hat{x}_0\in\mathcal{B}_x^c|\mathcal{H}_i\big)$, cf. (17). For a given region $\mathcal{B}_x$, the difference within square brackets depends on the difference between $f_{\hat{x}_j|t\in\mathcal{P}_j}\big(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\big)$ and $f_{\hat{x}_j}\big(\theta|\mathcal{H}_i\big)$. The conditional PDF $f_{\hat{x}_j|t\in\mathcal{P}_j}\big(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\big)$ would become equal to the PDF $f_{\hat{x}_j}\big(\theta|\mathcal{H}_i\big)$, if the correlation between $\hat{x}_j$ and t would be zero. With (11), however, there is a non-zero correlation between $\hat{x}_j$ and t, driven by $Q_{\hat{x}_0\hat{x}_0}, Q_{tt}$

and L_j. In addition, the conditional PDF $f_{\hat{x}_j|t\in\mathcal{P}_j}\left(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\right)$, using the total probability rule, can be written as

$$f_{\hat{x}_j|t\in\mathcal{P}_j}(\theta|t\in\mathcal{P}_j,\mathcal{H}_i)=\frac{1}{\mathrm{P}(t\in\mathcal{P}_j|\mathcal{H}_i)}\times\left(f_{\hat{x}_j}(\theta|\mathcal{H}_i)-\sum_{k\neq j}^{m}f_{\hat{x}_j|t\in\mathcal{P}_k}\left(\theta|t\in\mathcal{P}_k,\mathcal{H}_i\right)\mathrm{P}\left(t\in\mathcal{P}_k|\mathcal{H}_i\right)\right) \tag{26}$$

which reveals that

$$\mathrm{P}\left(t\in\mathcal{P}_j|\mathcal{H}_i\right)\to 1\Rightarrow f_{\hat{x}_j|t\in\mathcal{P}_j}\left(\theta|t\in\mathcal{P}_j,\mathcal{H}_i\right)\to f_{\hat{x}_j}\left(\theta|\mathcal{H}_i\right) \tag{27}$$

If $j=i$ and $|b_i|\to\infty$ then we have $\mathrm{P}\left(t\in\mathcal{P}_i|\mathcal{H}_i\right)\to 1$. As a result, the difference $\left(\overline{\mathrm{IR}}|\mathcal{H}_i-\mathrm{IR}^o|\mathcal{H}_i\right)$ is mainly driven by the term in the summation corresponding with $j=i$. This in tandem with (27) gives

$$|b_i|\to\infty\Rightarrow\overline{\mathrm{IR}}|\mathcal{H}_i-\mathrm{IR}^o|\mathcal{H}_i\to 0 \tag{28}$$

saying that for a very large bias magnitude, the difference between rigorous and approximate integrity risk vanishes.

Finally, we note that the issue of correlation-neglect between $\hat{x}_j$ and t also comes up if one would use the outcomes of testing in an a posteriori evaluation. In that case one would have to work with the PDF of $\left(\bar{x}|t\in\mathcal{P}_j\right)$ which is different from that of $\hat{x}_j$, despite the fact that both random vectors, $\left(\bar{x}|t\in\mathcal{P}_j\right)$ and $\hat{x}_j$, have the same sample outcome (Teunissen 2017). For instance, if $\mathcal{H}_j$ is the identified hypothesis, confidence levels are typically evaluated in practice as $\mathrm{P}\left(\hat{x}_j\in\mathcal{B}_x|\mathcal{H}_i\right)$, see e.g. (Wieser 2004; Devoti et al. 2011; Dheenathayalan et al. 2016), while they should be evaluated as $\mathrm{P}\left(\hat{x}_j\in\mathcal{B}_x|t\in\mathcal{P}_j,\mathcal{H}_i\right)$. The difference between their hypothesis averaged versions will then provide differences as those between (23) and (24).

Numerical analysis: single alternative hypothesis

In this section, we evaluate both the "rigorous" and "approximate" integrity risks defined by (23) and (24). To get a better understanding of their characteristics, we consider a simple observational model with only a single alternative hypothesis. Suppose that in (1), there is only one unknown parameter ($n=1$) and also the redundancy of the model is one ($r=1$), i.e., $x\in\mathbb{R}$ and $t\in\mathbb{R}$. The canonical form of such a model, applying the Tienstra-transformation $\mathcal{T}$ to the (assumed) normally distributed vector of observables y (Teunissen 2017), reads

$$\mathcal{T}y=\begin{bmatrix}A^+\\B^T\end{bmatrix}y=\begin{bmatrix}\hat{x}_0\\t\end{bmatrix}\overset{\mathcal{H}_i}{\sim}\mathcal{N}\left(\begin{bmatrix}x+b_{\hat{x}_{0,i}}\\b_{t_i}\end{bmatrix},\begin{bmatrix}\sigma^2_{\hat{x}_0}&0\\0&\sigma^2_t\end{bmatrix}\right) \tag{29}$$

which is specified for $i\in\{0,a\}$ as

$$\begin{aligned}&\mathcal{H}_0:\ b_{\hat{x}_{0,0}}=0,\ b_{t_0}=0\\&\mathcal{H}_a:\ b_{\hat{x}_{0,a}}=L_ab_a,\ b_{t_a}=b_a\end{aligned} \tag{30}$$

for some $b_a\in\mathbb{R}/\{0\}$, and also $L_a\in\mathbb{R}$ which establishes the following link:

$$\hat{x}_a=\hat{x}_0-L_at \tag{31}$$

With the mean of $\hat{x}_0$ and t given by (29) and (30), we have $E(\hat{x}_a|\mathcal{H}_0)=E(\hat{x}_a|\mathcal{H}_a)=x$.

The corresponding DIA-datasnooping procedure is defined as

Detection: Accept $\mathcal{H}_0$ if $t\in\mathcal{P}_0$ with

$$\mathcal{P}_0=\left[-\sqrt{k_{\alpha,1}},\sqrt{k_{\alpha,1}}\right] \tag{32}$$

Provide $\hat{x}_0$ as the estimate of x.

Identification: Select $\mathcal{H}_a$ if $t\in\mathcal{P}_0^c$ with $\mathcal{P}_0^c=\mathbb{R}/\mathcal{P}_0$.

Adaptation: When $\mathcal{H}_a$ is selected, $\hat{x}_a$ is provided as the estimate of x.

With the above three steps, the DIA estimator and its PDF under $\mathcal{H}_i, i\in\{0,a\}$, are given by

$$\bar{x}=\hat{x}_0p_0(t)+\hat{x}_a(1-p_0(t)) \tag{33}$$

and

$$\begin{aligned}f_{\bar{x}}\left(\theta|\mathcal{H}_i\right)&=f_{\hat{x}_0}\left(\theta|\mathcal{H}_i\right)\int_{\mathcal{P}_0}f_t\left(\tau|\mathcal{H}_i\right)\mathrm{d}\tau+\int_{\mathcal{P}_0^c}f_{\hat{x}_0}\left(\theta+L_a\tau|\mathcal{H}_i\right)f_t\left(\tau|\mathcal{H}_i\right)\mathrm{d}\tau\\&=f_{\hat{x}_0}\left(\theta|\mathcal{H}_i\right)+\int_{\mathcal{P}_0^c}\left[f_{\hat{x}_0}\left(\theta+L_a\tau|\mathcal{H}_i\right)-f_{\hat{x}_0}\left(\theta|\mathcal{H}_i\right)\right]f_t\left(\tau|\mathcal{H}_i\right)\mathrm{d}\tau\end{aligned} \tag{34}$$

In (34), the second equality follows from $\mathcal{P}_0^c = \mathbb{R}/\mathcal{P}_0$.

Decomposition of $f_{\bar{x}}(\theta|\mathcal{H}_0)$ and $f_{\bar{x}}(\theta|\mathcal{H}_a)$

As there is only one alternative hypothesis $\mathcal{H}_a$, the events in (13) are reduced to four events CA, FA, MD and CD, and the decomposition of $f_{\bar{x}}(\theta|\mathcal{H}_i)$ in (15) is then simplified to

$$\begin{aligned} f_{\bar{x}}(\theta|\mathcal{H}_0) &= f_{\hat{x}_0}(\theta|\mathcal{H}_0)\mathrm{P_{CA}} + f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})\mathrm{P_{FA}} \\ f_{\bar{x}}(\theta|\mathcal{H}_a) &= f_{\hat{x}_0}(\theta|\mathcal{H}_a)\mathrm{P_{MD}} + f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})\mathrm{P_{CD}} \end{aligned} \tag{35}$$

Note that the subscripts of MD and CD as in (13) are dropped, as $\mathcal{H}_a$ is the only alternative. In Fig. 1, assuming (for example) that $\sigma^2_{\hat{x}_0} = 0.5\mathrm{m}^2, \sigma_t^2 = 2\,\mathrm{m}^2$ and $L_a = 0.5$, we show how the PDFs $f_{\bar{x}}(\theta|\mathcal{H}_0)$ (top) and $f_{\bar{x}}(\theta|\mathcal{H}_a)$ (middle and bottom) are formed according to (35). The solid and dashed blue curves, respectively, depict $f_{\hat{x}_0}(\theta|\mathcal{H}_0) = f_{\bar{x}|\mathrm{CA}}(\theta|\mathrm{CA})$ and $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$ in the top panel, and $f_{\hat{x}_0}(\theta|\mathcal{H}_a) = f_{\bar{x}|\mathrm{MD}}(\theta|\mathrm{MD})$ and $f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})$ in the middle and bottom panels. The *black* curve shows $f_{\hat{x}_a}(\theta|\mathcal{H}_0)$, which is also equal to $f_{\hat{x}_a}(\theta|\mathcal{H}_a)$. The probability of false alarm $\mathrm{P_{FA}}$ is usually user defined by setting the appropriate size of $\mathcal{P}_0$, hence an input to the DIA procedure both under null and alternative hypotheses. To assess the PDF of $\hat{x}_0$ and $\bar{x}$ under the alternative H_a, one additionally needs to set the size of the bias b_a, or alternatively, one may choose to set the correct detection probability $\mathrm{P_{CD}}$ as we did here.

As was mentioned earlier, the conditional PDFs $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$ and $f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})$ are *non-normal*, which for the case of one single alternative can be expressed as

$$f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA}) = \frac{1}{\mathrm{P_{FA}}}\left[f_{\hat{x}_a}(\theta|\mathcal{H}_0) - f_{\hat{x}_a|\mathrm{CA}}(\theta|\mathrm{CA})\mathrm{P_{CA}}\right] \tag{36}$$

$$f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD}) = \frac{1}{\mathrm{P_{CD}}}\left[f_{\hat{x}_a}(\theta|\mathcal{H}_a) - f_{\hat{x}_a|\mathrm{MD}}(\theta|\mathrm{MD})\mathrm{P_{MD}}\right] \tag{37}$$

Substituting (36) and (37) into (35), given $\mathrm{P_{FA}} = 1 - \mathrm{P_{CA}}$ and $\mathrm{P_{CD}} = 1 - \mathrm{P_{MD}}$, we have

$$\begin{aligned} &\mathrm{P_{FA}} \to 1 \Rightarrow f_{\bar{x}}(\theta|\mathcal{H}_0) \to f_{\hat{x}_a}(\theta|\mathcal{H}_0) \\ &\mathrm{P_{FA}} \to 0 \Rightarrow f_{\bar{x}}(\theta|\mathcal{H}_0) \to f_{\hat{x}_0}(\theta|\mathcal{H}_0) \\ &\mathrm{P_{CD}} \to 1 \Rightarrow f_{\bar{x}}(\theta|\mathcal{H}_a) \to f_{\hat{x}_a}(\theta|\mathcal{H}_a) \\ &\mathrm{P_{CD}} \to 0 \Rightarrow f_{\bar{x}}(\theta|\mathcal{H}_a) \to f_{\hat{x}_0}(\theta|\mathcal{H}_a) \end{aligned} \tag{38}$$

In Fig. 1, in agreement with (38), as P_{FA} decreases, in the graphs on top, the curve of $f_{\bar{x}}(\theta|\mathcal{H}_0)$ in red gets close to that of $f_{\hat{x}_0}(\theta|\mathcal{H}_0)$ in blue. Also, when P_{CD} increases from 0.40 to 0.99, in the graphs in the middle and at the bottom, the curve of $f_{\bar{x}}(\theta|\mathcal{H}_a)$ in red gets close to that of $f_{\hat{x}_a}(\theta|\mathcal{H}_a)$ in black and to that of $f_{\bar{x}|\mathrm{CD}}(\theta|\mathrm{CD})$ as well.

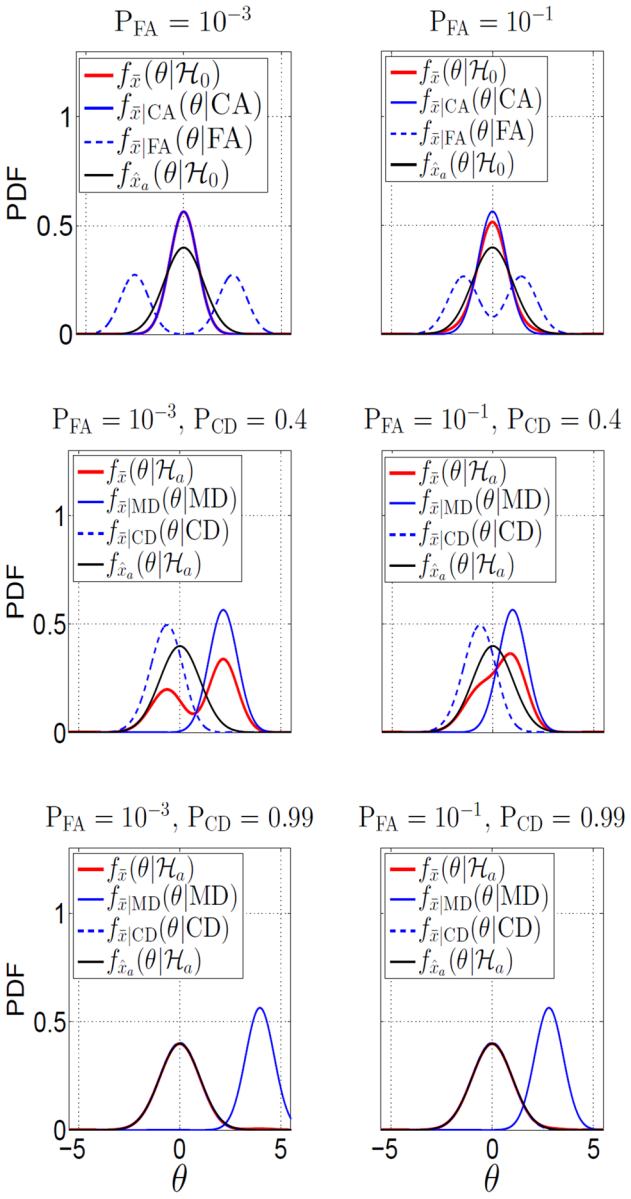

Fig. 1 Illustration of the DIA estimator PDF [top] under null hypothesis $\mathcal{H}_0$ and [middle and bottom] under alternative hypothesis $\mathcal{H}_a$. The contributing parameters are set to be $\sigma^2_{\hat{x}_0} = 0.5\mathrm{m}^2$, $\sigma_t^2 = 2\,\mathrm{m}^2$, and $L_a = 0.5$. Panels, from left to right, correspond to $\mathrm{P_{FA}} = 10^{-3}$ and $\mathrm{P_{FA}} = 10^{-1}$. Panels in the middle and at the bottom correspond to, resp., $\mathrm{P_{CD}} = 0.4$ and $\mathrm{P_{CD}} = 0.99$

Non-normality of $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$ and $f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})$

To appreciate the non-normality of the two PDFs $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$ and $f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})$, we show them for different values of the contributing factors, namely $\mathrm{P_{FA}}$,

$\mathrm{P_{MD}}\big(=1-\mathrm{P_{CD}}\big)$, $\sigma_{\hat{x}_0}$, σ_t and L_a, in Fig. 2. To highlight the non-normality of $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$ and $f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})$, we have also plotted (for reference) their normal counterparts having the same mean and variance in black. These normal PDFs, respectively, correspond with the random variables

$$\begin{aligned}\hat{x}_{\mathrm{fa}} &\sim \mathcal{N}\Big(E\big(\hat{x}_a|\mathrm{FA}\big),\sigma^2_{\hat{x}_a|\mathrm{FA}}\Big)\\ \hat{x}_{\mathrm{cd}} &\sim \mathcal{N}\Big(E\big(\hat{x}_a|\mathrm{CD}\big),\sigma^2_{\hat{x}_a|\mathrm{CD}}\Big)\end{aligned} \tag{39}$$

Note that the above random variables are only introduced here to illustrate the departure from normality of $\big(\hat{x}_a|\mathrm{FA}\big)$ and $\big(\hat{x}_a|\mathrm{CD}\big)$.

The panels to the left side of Fig. 2 demonstrate the behavior of $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$ in blue, in comparison to $f_{\hat{x}_{\mathrm{fa}}}(\theta)$ in black. We note that the following situations $\sigma_{\hat{x}_0}\uparrow, \sigma_t\downarrow, L_a\downarrow$ and $\mathrm{P_{FA}}\uparrow$ make the PDF $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$ get closer to a normal one. This can be explained as follows. We first consider the impact of $\sigma_{\hat{x}_0}, \sigma_t$ and L_a which drive the correlation between $\hat{x}_a$ and t as

$$\Big|\rho_{\hat{x}_a,t}\Big| = \left(1+\frac{\sigma^2_{\hat{x}_0}}{L_a^2\sigma_t^2}\right)^{-\frac{1}{2}} \tag{40}$$

If this correlation becomes zero, then the PDF $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$ becomes identical to the unconditional normal PDF $f_{\hat{x}_a}\big(\theta|\mathcal{H}_0\big)$. As (40) suggests, this would be realized if $L_a\to 0$ and $\sigma_t\to 0$ as well as $\sigma_{\hat{x}_0}\to\infty$. Now, we consider the impact of $\mathrm{P_{FA}}$, which can be explained through (36). Since $\mathrm{P_{CA}}=1-\mathrm{P_{FA}}$, increasing $\mathrm{P_{FA}}$ (thus decreasing $\mathrm{P_{CA}}$) leads to smaller differences between $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$ and $f_{\hat{x}_a}\big(\theta|\mathcal{H}_0\big)$. This can also be seen by comparing the dashed blue curves with the solid black ones in the first row of Fig. 1.

Shown to the right of Fig. 2 are the graphs of $f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})$ together with those of $f_{\hat{x}_{\mathrm{cd}}}(\theta)$. The response of $f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})$ to the changes in the parameters $\sigma_{\hat{x}_0}$, σ_t and L_a is similar to that of $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$. $f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})$ in addition depends on $\mathrm{P_{CD}}$ according to (37). Since $\mathrm{P_{CD}}=1-\mathrm{P_{MD}}$, increasing $\mathrm{P_{CD}}$ (thus decreasing $\mathrm{P_{MD}}$) leads to smaller differences between $f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})$ and $f_{\hat{x}_a}\big(\theta|\mathcal{H}_a\big)$. This can also be seen by comparing the dashed blue curves with the solid black ones in the second and third rows of Fig. 1.

Rigorous vs approximate integrity risk

The rigorous and approximate integrity risks for the case of working with a single alternative hypothesis $\mathcal{H}_a$ are formulated in, respectively, the first and the second rows of Table 1. The difference between them reads

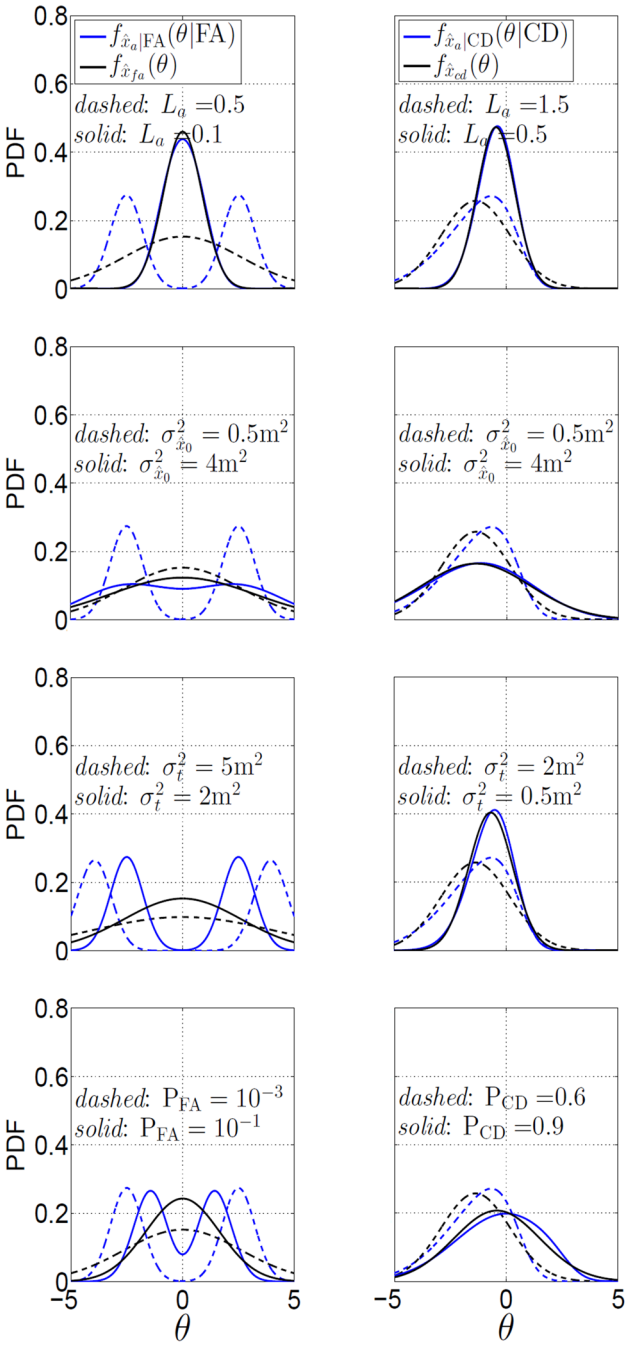

Fig. 2 Illustration of the non-normal distribution of the components of the DIA estimator (left) $f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA})$ in blue and the normal distribution $f_{\hat{x}_{\mathrm{fa}}}(\theta)$, cf. (39), for reference, in black; (right) $f_{\hat{x}_a|\mathrm{CD}}(\theta|\mathrm{CD})$ in blue and the normal distribution $f_{\hat{x}_{\mathrm{cd}}}(\theta)$, cf. (39), for reference, in black. Unless otherwise mentioned in each panel, the default settings for the left panels are $\sigma^2_{\hat{x}_0}=0.5\mathrm{m}^2$, $\sigma^2_t=2\,\mathrm{m}^2$, $L_a=0.5$ and $\mathrm{P_{FA}}=10^{-3}$, and for the right panels are $\sigma^2_{\hat{x}_0}=0.5\mathrm{m}^2$, $\sigma^2_t=2\mathrm{m}^2$, $L_a=1.5$, $\mathrm{P_{FA}}=10^{-3}$ and $\mathrm{P_{CD}}=0.6$

Table 1 Integrity risk based on the DIA estimator ($\overline{\text{IR}}$) and its approximation by ignoring the correlation between $\hat{x}_a$ and t (IR^o) for the case of a null- and a single alternative hypotheses. $\mathcal{B}_x^c = \mathbb{R}^n/\mathcal{B}_x$

$$\begin{aligned}\overline{\text{IR}} &= \text{P}(\mathcal{H}_0)\times\int_{\mathcal{B}_x^c}\left[f_{\hat{x}_0}(\theta|\mathcal{H}_0)\text{P}_{\text{CA}}+f_{\hat{x}_a|\text{FA}}(\theta|\text{FA})\text{P}_{\text{FA}}\right]\text{d}\theta\\ &\quad+\text{P}(\mathcal{H}_a)\times\int_{\mathcal{B}_x^c}\left[f_{\hat{x}_0}(\theta|\mathcal{H}_a)\text{P}_{\text{MD}}+f_{\hat{x}_a|\text{CD}}(\theta|\text{CD})\text{P}_{\text{CD}}\right]\text{d}\theta\\ \text{IR}^o &= \text{P}(\mathcal{H}_0)\times\int_{\mathcal{B}_x^c}\left[f_{\hat{x}_0}(\theta|\mathcal{H}_0)\text{P}_{\text{CA}}+f_{\hat{x}_a}(\theta|\mathcal{H}_0)\text{P}_{\text{FA}}\right]\text{d}\theta\\ &\quad+\text{P}(\mathcal{H}_a)\times\int_{\mathcal{B}_x^c}\left[f_{\hat{x}_0}(\theta|\mathcal{H}_a)\text{P}_{\text{MD}}+f_{\hat{x}_a}(\theta|\mathcal{H}_a)\text{P}_{\text{CD}}\right]\text{d}\theta\end{aligned}$$

$$\begin{aligned}\overline{\text{IR}}-\text{IR}^o &= \text{P}(\mathcal{H}_0)\times\left(\overline{\text{IR}}|\mathcal{H}_0-\text{IR}^o|\mathcal{H}_0\right)\\ &\quad+\text{P}(\mathcal{H}_a)\times\left(\overline{\text{IR}}|\mathcal{H}_a-\text{IR}^o|\mathcal{H}_a\right)\end{aligned}\tag{41}$$

where

$$\overline{\text{IR}}|\mathcal{H}_0-\text{IR}^o|\mathcal{H}_0=P_{\text{FA}}\int_{\mathcal{B}_x^c}\left[f_{\hat{x}_a|\text{FA}}(\theta|\text{FA})-f_{\hat{x}_a}(\theta|\mathcal{H}_0)\right]d\theta\tag{42}$$

and

$$\overline{\text{IR}}|\mathcal{H}_a-\text{IR}^o|\mathcal{H}_a=\text{P}_{\text{CD}}\int_{\mathcal{B}_x^c}\left[f_{\hat{x}_a|\text{CD}}(\theta|\text{CD})-f_{\hat{x}_a}(\theta|\mathcal{H}_a)\right]\text{d}\theta\tag{43}$$

The difference between $\overline{\text{IR}}$ and IR^o lies in the difference between $f_{\hat{x}_a|\text{FA}}(\theta|\text{FA})$ and $f_{\hat{x}_a}\left(\theta|\mathcal{H}_0\right)$ and the difference between $f_{\hat{x}_a|\text{CD}}(\theta|\text{CD})$ and $f_{\hat{x}_a}\left(\theta|\mathcal{H}_a\right)$. These differences, as discussed previously, depend on $\rho_{\hat{x}_a,t}$, P_{FA} and, in case of the latter, on P_{CD}. In case $\rho_{\hat{x}_a,t}=0$, then $f_{\hat{x}_a|\text{FA}}(\theta|\text{FA})=f_{\hat{x}_a}\left(\theta|\mathcal{H}_0\right)$ and $f_{\hat{x}_a|\text{CD}}(\theta|\text{CD})=f_{\hat{x}_a}\left(\theta|\mathcal{H}_a\right)$, thereby $\overline{\text{IR}}=\text{IR}^o$. As nominal conditions are more likely than anomalies incurring measurement biases, the probability of the occurrence of $\mathcal{H}_0$ is far larger than that of $\mathcal{H}_a$, see e.g. Wu et al. (2013). Thus, the impact of $(\overline{\text{IR}}|\mathcal{H}_a-\text{IR}^o|\mathcal{H}_a)$ is downweighted by the small value of $\text{P}\left(\mathcal{H}_a\right)$. Therefore, in the following, we first discuss the behavior of $(\overline{\text{IR}}|\mathcal{H}_0-\text{IR}^o|\mathcal{H}_0)$ and then the behavior of $(\overline{\text{IR}}-\text{IR}^o)$.

Evaluation of $\left(\overline{\text{IR}}|\mathcal{H}_0-\text{IR}^o|\mathcal{H}_0\right)$

It can be shown that the difference $(\overline{\text{IR}}|\mathcal{H}_0-\text{IR}^o|\mathcal{H}_0)$ in (42) is always positive. For a proof, see the Appendix. Assuming $\sigma^2_{\hat{x}_0}=0.5\text{m}^2$ and $\sigma_t^2=2\,\text{m}^2$, Fig. 3 illustrates graphs of $(\overline{\text{IR}}|\mathcal{H}_0-\text{IR}^o|\mathcal{H}_0)$ (solid lines) and those of $(\overline{\text{IR}}|\mathcal{H}_0)$ (dashed lines) as a function of AL, for $\text{P}_{\text{FA}}=10^{-1}$ (top) and

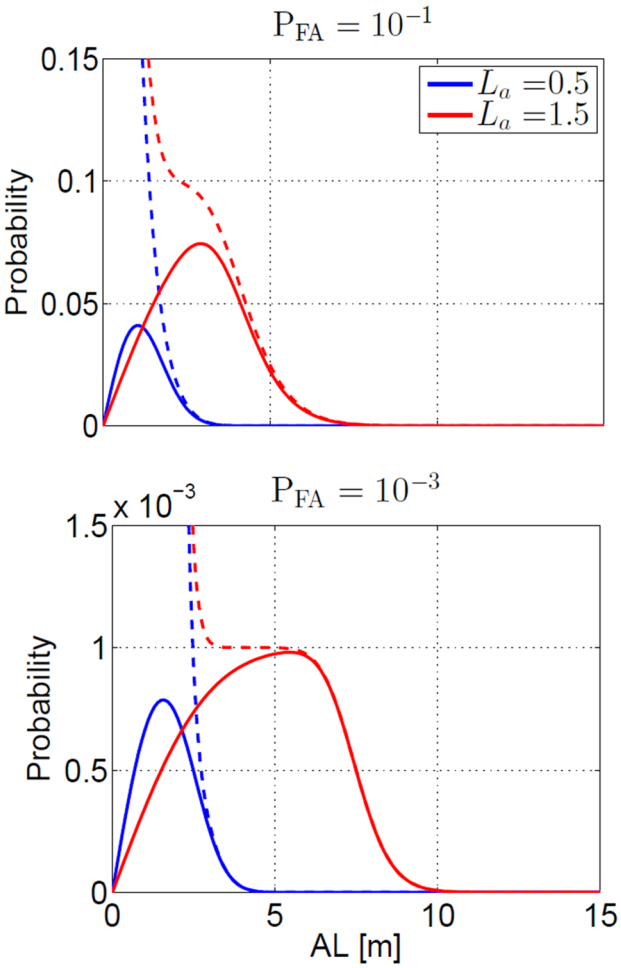

Fig. 3 Illustration of the behavior of integrity risks corresponding with $\mathcal{B}_x=[x-\text{AL},x+\text{AL}]$ as a function of AL, assuming $\sigma^2_{\hat{x}_0}=0.5\,\text{m}^2$ and $\sigma_t^2=2\,\text{m}^2$, for the case of a null- and a single alternative hypotheses. The solid lines show $(\overline{\text{IR}}|\mathcal{H}_0-\text{IR}^o|\mathcal{H}_0)$, cf. (42), while the dashed lines show the corresponding $(\overline{\text{IR}}|\mathcal{H}_0)$

$\text{P}_{\text{FA}}=10^{-3}$ (bottom), and $L_a=0.5,1.5$ (in blue and red, resp.). Comparing the solid lines with their corresponding dashed lines, we note, depending on the values of L_a and P_{FA}, that after a certain alert limit, the values of $(\overline{\text{IR}}|\mathcal{H}_0)$ and $(\overline{\text{IR}}|\mathcal{H}_0-\text{IR}^o|\mathcal{H}_0)$ approach each other, implying that the approximate integrity risk $\left(\text{IR}^o|\mathcal{H}_0\right)$ gets very small indeed. We explain this behavior for the *blue* curves at the bottom when AL = 4 m. The probability mass of the PDFs $f_{\hat{x}_a|\text{FA}}(\theta|\text{FA})$ and $f_{\hat{x}_a}\left(\theta|\mathcal{H}_0\right)$, the dashed blue curve and the black curve in the upper left panel in Fig. 1, are at the level of 2×10^{-2} and 6×10^{-5}, respectively, outside $\mathcal{B}_x=[x-4,x+4]$. In addition, the probability mass of the PDF $f_{\hat{x}_0}\left(\theta|\mathcal{H}_0\right)$ outside $\mathcal{B}_x=[x-4,x+4]$ is at the level of 10^{-8}. These values, given $P_{\text{FA}}=10^{-3}$, will then result in a difference at the level of 8×10^{-8} between $(\overline{\text{IR}}|\mathcal{H}_0)$ and $(\overline{\text{IR}}|\mathcal{H}_0-\text{IR}^o|\mathcal{H}_0)$. As a consequence of this case, the

approximate integrity risk under the null hypothesis is too optimistic by a factor of 300.

It can be seen that the graphs of $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ take only positive values. We note that all curves as a function of AL show almost the same signature. They first increase and then decrease to zero. Since this behavior depends on AL, it can be explained by looking at the integral part in (42), which is the difference between $\text{P}\big(|\hat{x}_a - x| > \text{AL}|\text{FA}\big)$ and $\text{P}\big(|\hat{x}_a - x| > \text{AL}|\mathcal{H}_0\big)$. We have

$$\begin{aligned} \text{AL} \to 0 &\Rightarrow \begin{cases} \text{P}(|\hat{x}_a - x| > \text{AL}|\text{FA}) \to 1 \\ \text{P}(|\hat{x}_a - x| > \text{AL}|\mathcal{H}_0) \to 1 \end{cases} \\ \text{AL} \to \infty &\Rightarrow \begin{cases} \text{P}(|\hat{x}_a - x| > \text{AL}|\text{FA}) \to 0 \\ \text{P}(|\hat{x}_a - x| > \text{AL}|\mathcal{H}_0) \to 0 \end{cases} \end{aligned} \tag{44}$$

As shown in the Appendix, $\text{P}(|\hat{x}_a - x| \le \text{AL}|\mathcal{H}_0)$ is always greater than $\text{P}(|\hat{x}_a - x| \le \text{AL}|\text{FA})$. Therefore, it can be concluded that the PDF $f_{\hat{x}_a}\big(\theta|\mathcal{H}_0\big)$ is more peaked around x than $f_{\hat{x}_a|\text{FA}}(\theta|\text{FA})$. Therefore, when AL increases, $\text{P}\big(|\hat{x}_a - x| > \text{AL}|\mathcal{H}_0\big)$ decreases more rapidly than $\text{P}\big(|\hat{x}_a - x| > \text{AL}|\text{FA}\big)$. This, together with (44) and the fact that the probabilities $\text{P}\big(|\hat{x}_a - x| > \text{AL}|\mathcal{H}_0\big)$ and $\text{P}\big(|\hat{x}_a - x| > \text{AL}|\text{FA}\big)$ are continuous functions of AL, results in an increasing and then decreasing behavior for $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$.

In Fig. 3, when L_a increases, the curve of $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ stretches over a larger range of values of AL. This is due to the fact that increasing L_a reduces the peakedness of the PDFs $f_{\hat{x}_a}\big(\theta|\mathcal{H}_0\big)$ and $f_{\hat{x}_a|\text{FA}}(\theta|\text{FA})$ around x. The impact of changing L_a on $f_{\hat{x}_a}\big(\theta|\mathcal{H}_0\big)$ and $f_{\hat{x}_a|\text{FA}}(\theta|\text{FA})$ is demonstrated in (31) and Fig. 2 (top-left), respectively. Therefore, both the probabilities $\text{P}\big(|\hat{x}_a - x| > \text{AL}|\mathcal{H}_0\big)$ and $\text{P}\big(|\hat{x}_a - x| > \text{AL}|\text{FA}\big)$ behave more smoothly as function of AL, and so does their difference.

Decreasing α by a factor of 10^2, we note that the values of $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ in Fig. 3 also decrease by almost a factor of 10^2. Reducing α by a factor of 10^2 will reduce $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ by the same amount if $\text{P}\big(|\hat{x}_a - x| > \text{AL}|\text{FA}\big)$ remains invariant. However, as Fig. 2 (bottom-left) shows, reducing α by a factor of 10^2 increases $\text{P}\big(|\hat{x}_a - x| > \text{AL}|\text{FA}\big)$, which results in $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ decreasing by a factor slightly smaller than 10^2.

Evaluation of $\left(\overline{\text{IR}} - \text{IR}^o\right)$

Here, we investigate the behavior of the integrity risks when both the null- and alternative hypotheses are taken into account. The difference $(\overline{\text{IR}} - \text{IR}^o)$ in (41), due to the contribution of $\mathcal{H}_a$, depends on the bias value b_a. Considering $\sigma^2_{\hat{x}_0} = 0.5\,\text{m}^2, \sigma^2_t = 2\,\text{m}^2$ and $\text{P}\big(\mathcal{H}_0\big) = 0.9$ (thus $\text{P}\big(\mathcal{H}_a\big) = 0.1$; the latter probability is usually much smaller in practice), Fig. 4 shows the curves of $(\overline{\text{IR}} - \text{IR}^o)$ as a function of b_a for two different values of AL in $\mathcal{B}_x = [x - \text{AL}, x + \text{AL}]$, L_a and P_{FA}. Decreasing L_a, although the correlation $\rho_{\hat{x}_a,t}$ decreases, the difference $(\overline{\text{IR}} - \text{IR}^o)$ may increase or decrease depending on AL and P_{FA}. It can be seen that $(\overline{\text{IR}} - \text{IR}^o)$ does not change too much as a function of b_a. This is due to the fact that any change in b_a will change $(\overline{\text{IR}}|\mathcal{H}_a - \text{IR}^o|\mathcal{H}_a)$, of which the impact is downweighted by $\text{P}\big(\mathcal{H}_a\big) = 0.1$. Note that $\overline{\text{IR}}$ is larger than IR^o in most cases, revealing that using the approximate integrity risk instead of the rigorous one generally will be hazardous. The IR^o does not provide a safe bound to the actual $\overline{\text{IR}}$.

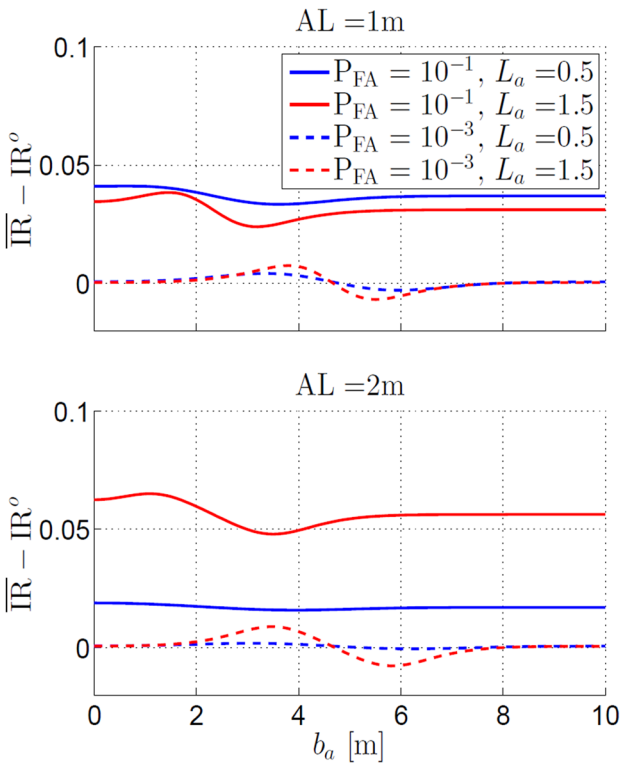

Fig. 4 Illustration of the difference between rigorous and approximate integrity risks $(\overline{\text{IR}} - \text{IR}^o)$ in (41) as a function of bias value b_a corresponding with $\mathcal{B}_x = [x - \text{AL}, x + \text{AL}]$. The contributing parameters are set to be $\sigma^2_{\hat{x}_0} = 0.5\,\text{m}^2$, $\sigma^2_t = 2\,\text{m}^2$ and $\text{P}\big(\mathcal{H}_0\big) = 0.9$. The variation of P_{FA} and L_a is highlighted through changing the style and color of the curve, respectively: $\text{P}_{\text{FA}} = 10^{-1}$ in solid line and $\text{P}_{\text{FA}} = 10^{-3}$ in dashed line, $L_a = 0.5$ in blue and $L_a = 1.5$ in red

Numerical analysis: multiple alternative hypotheses

So far, for simplicity, we have been working with an observational model with one unknown parameter and one redundancy. In this section, we work with a satellite-based

single-point positioning (SPP) model based on the observations of m satellites with four unknown parameters ($n = 4$) and $r = m - 4$ redundancy. As alternative hypotheses, we consider those given in (2). In that case there are as many alternative hypotheses as there are observations.

We first present the observational model, and then we analyze, by means of three practical examples, the difference between the rigorous and approximate integrity risk for the contributions under the null hypothesis ($\mathcal{H}_0$).

SPP observational model

Assuming there are m pseudorange observations, the observational model under $\mathcal{H}_i$ for $i = 0, 1, \ldots, m$ is given as

$$\mathcal{H}_i : E(y) = [G \quad e_m] \begin{bmatrix} x \\ dt \end{bmatrix} + c_i b_i,\ Q_{yy} = \sigma_p^2 I_m \tag{45}$$

where the $m \times 3$ matrix $G = \left[-u_1^T, \ldots, -u_m^T\right]^T$ contains the receiver-satellite unit direction vectors u_i as its rows, e_m is the m-vector of ones, and again $c_0 b_0 = 0$. The unknown receiver coordinate components and clock error are, respectively, denoted by the 3-vector x and scalar dt. The dispersion of the observables is characterized through the standard deviation σ_p and the identity matrix I_m. At this stage, to simplify our analysis, we do not consider a satellite-elevation-dependent variance matrix. In the following, we only concentrate on the $\mathcal{H}_0$-driven difference between $\overline{\text{IR}}$ and IR^o, as the probability of the occurrence of $\mathcal{H}_0$ is by far larger than that of the alternative hypotheses.

Evaluation of $\left(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0\right)$

Setting $i = 0$ in (25), we have

$$\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0 = \sum_{j=1}^{m} \left[P(\hat{x}_j \in \mathcal{B}_x^c | t \in \mathcal{P}_j, \mathcal{H}_0) - P(\hat{x}_j \in \mathcal{B}_x^c | \mathcal{H}_0)\right] P\left(t \in \mathcal{P}_j | \mathcal{H}_0\right) \tag{46}$$

which shows, for a given satellite geometry (design matrix), that the difference $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ depends on $\text{P}_{\text{FA}} = \alpha$, $\mathcal{B}_x$ and Q_{yy}. In the following, for a few satellite geometries, we illustrate $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ as function of the contributing factors. Note, we will compute the integrity risk for the horizontal components and the vertical component separately. As such, the integrity risk for the vertical component is computed for

$$\mathcal{B}_{x_V} = \left\{\theta_V \in \mathbb{R} \middle| |\theta_V - x_V| \leq \text{AL}\right\} \tag{47}$$

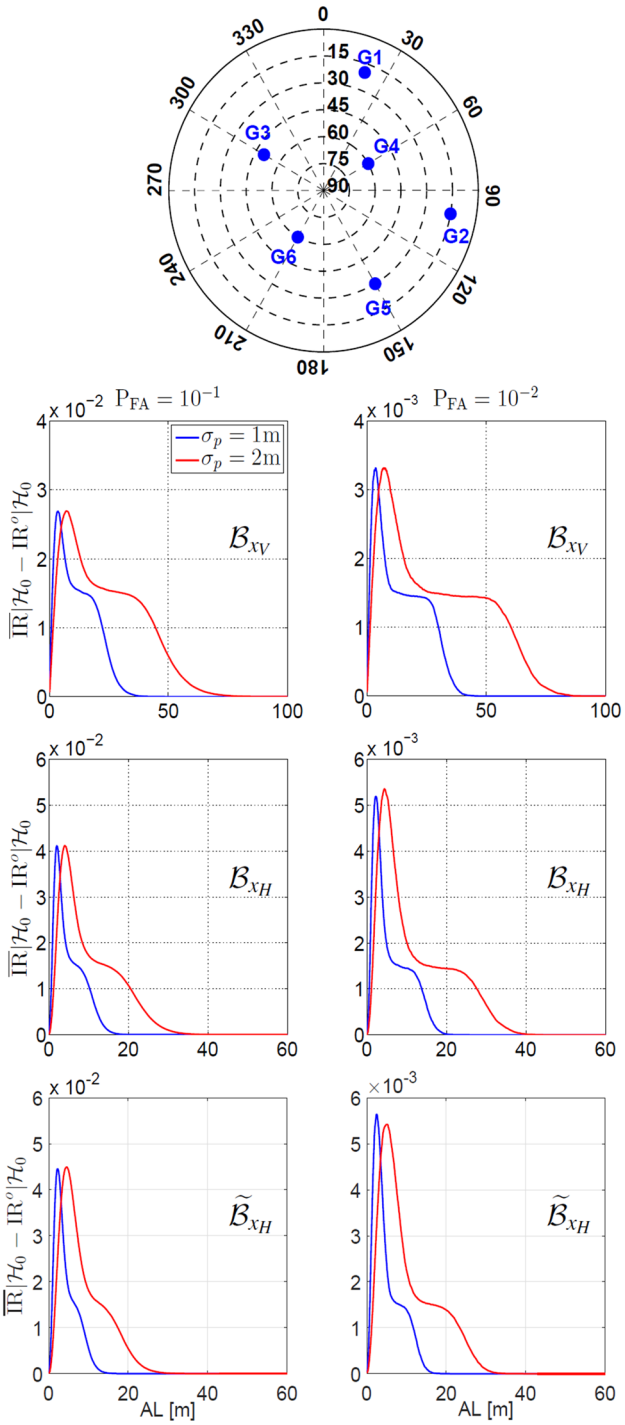

Fig. 5 Skyplot of the satellites (top), the difference $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ for vertical component (second row) and horizontal components (third and fourth rows) in the model given in (45), as a function of alert limit AL, for different values of σ_p (cf. 45) and $\text{P}_{\text{FA}} = \alpha$

and for the horizontal components for

$$\mathcal{B}_{x_H} = \left\{ \theta_H \in \mathbb{R}^2 \middle| ||\theta_H - x_H||_{l_2} \leq \text{AL} \right\} \tag{48}$$

$$\tilde{\mathcal{B}}_{x_H} = \left\{ \theta_H \in \mathbb{R}^2 \middle| ||\theta_H - x_H||_{C_{\hat{x}_{H_0}\hat{x}_{H_0}}} \leq \text{AL} \right\} \tag{49}$$

where $x_H \in \mathbb{R}^2$ and $x_V \in \mathbb{R}$ are, respectively, the horizontal components and the vertical component of x. Region $\mathcal{B}_{x_H}$ is a circle with radius AL, and $\tilde{\mathcal{B}}_{x_H}$ is an ellipse driven by cofactor matrix $C_{\hat{x}_{H_0}\hat{x}_{H_0}}$. Without loss of generality, our illustrations will be depicted for $x = 0$, thus $x_V = 0$ and $x_H = 0$.

Example 1: Fig. 5

The skyplot in Fig. 5 (top) shows a geometry of six satellites. The graphs of the difference $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ as a function of AL are shown for the vertical component [second row] and the horizontal components (third and fourth rows) for different values of σ_p (cf. 45) and $\text{P}_{\text{FA}} = \alpha$. The third row corresponds with $\mathcal{B}_{x_H}$, cf. (48), while the fourth row corresponds with $\tilde{\mathcal{B}}_{x_H}$, cf. (49). It is important to note that $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ has always 'positive' values for all components, meaning that employing the approximate integrity risk instead of the rigorous one could be dangerous, depending on the application at hand. To get a better appreciation of such danger, the values of $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ together with $(\overline{\text{IR}}|\mathcal{H}_0)$ are tabulated in Table 2 for some values of AL. For example, for the vertical component with the settings of $\text{P}_{\text{FA}} = 10^{-1}$, $\sigma_p = 1\text{m}$ and $\text{AL} = 15$ m, we have $(\overline{\text{IR}}|\mathcal{H}_0) = 0.0168$ and $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0) = 0.0148$, implying that the approximate integrity risk $(\overline{\text{IR}}|\mathcal{H}_0)$ is too optimistic by a factor of 8.

We also note that all graphs as a function of AL show almost the same signature. They first increase and then decrease to zero at different slopes. Since this behavior depends on the AL values, it can be explained by looking at $\text{P}(\hat{x}_{*_j} \in \mathcal{B}^c_{x_*}|t \in \mathcal{P}_j, \mathcal{H}_0) - \text{P}(\hat{x}_{*_j} \in \mathcal{B}^c_{x_*}|\mathcal{H}_0)$ as a function of AL with $*= \{H, V\}$. We have

$$\begin{aligned} \text{AL} \to 0 &\Rightarrow \begin{cases} \text{P}(\hat{x}_{*_j} \in \mathcal{B}^c_{x_*}|t \in \mathcal{P}_j, \mathcal{H}_0) \to 1 \\ \text{P}(\hat{x}_{*_j} \in \mathcal{B}^c_{x_*}|\mathcal{H}_0) \to 1 \end{cases} \\ \text{AL} \to \infty &\Rightarrow \begin{cases} \text{P}(\hat{x}_{*_j} \in \mathcal{B}^c_{x_*}|t \in \mathcal{P}_j, \mathcal{H}_0) \to 0 \\ \text{P}(\hat{x}_{*_j} \in \mathcal{B}^c_{x_*}|\mathcal{H}_0) \to 0 \end{cases} \end{aligned} \tag{50}$$

We now, as an example, take the vertical component integrity risk. Assuming that $x_V = 0$ and $\sigma_p = 1\text{m}$, Fig. 6 illustrates the PDFs $f_{\hat{x}_{V_j}}\left(\theta_V|\mathcal{H}_0\right)$ (dashed blue curves) and $f_{\hat{x}_{V_j}|t\in\mathcal{P}_j}\left(\theta_V|t \in \mathcal{P}_j, \mathcal{H}_0\right)$ for $\alpha = 10^{-1}$ (solid *blue* curves) and $\alpha = 10^{-2}$ (solid red curves), for all six alternative hypotheses $j = 1, \ldots, 6$. It can be seen that the PDF $f_{\hat{x}_{V_j}}\left(\theta_V|\mathcal{H}_0\right)$ (dashed) is more peaked around zero than $f_{\hat{x}_{V_j}|t\in\mathcal{P}_j}\left(\theta_V|t \in \mathcal{P}_j, \mathcal{H}_0\right)$. Therefore, when the AL increases, then $\text{P}(\hat{x}_{V_j} \in \mathcal{B}^c_{x_V}|\mathcal{H}_0)$ decreases more rapidly than $\text{P}(\hat{x}_{V_j} \in \mathcal{B}^c_{x_V}|t \in \mathcal{P}_j, \mathcal{H}_0)$. This, together with (50) and the fact that the probabilities $\text{P}(\hat{x}_{V_j} \in \mathcal{B}^c_{x_V}|\mathcal{H}_0)$ and $\text{P}(\hat{x}_{V_j} \in \mathcal{B}^c_{x_V}|t \in \mathcal{P}_j, \mathcal{H}_0)$ are continuous functions of AL, results in increasing and then decreasing

Table 2 Values of the rigorous integrity risk $(\overline{\text{IR}}|\mathcal{H}_0)$ and its difference with its approximation $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$, taken from the graphs in Fig. 5

		AL (m)	$\overline{\text{IR}}\|\mathcal{H}_0$	$\overline{\text{IR}}\|\mathcal{H}_0 - \text{IR}^o\|\mathcal{H}_0$	AL (m)	$\overline{\text{IR}}\|\mathcal{H}_0$	$\overline{\text{IR}}\|\mathcal{H}_0 - \text{IR}^o\|\mathcal{H}_0$
$\text{P}_{\text{FA}} = 10^{-1}$, $\sigma_p = 1$ m							
	$\mathcal{B}_{x_V}$	15	0.0168	0.0148	30	0.0015	0.0015
	$\mathcal{B}_{x_H}$	10	0.0112	0.0107	18	0.0001	0.0001
	$\tilde{\mathcal{B}}_{x_H}$	5	0.0211	0.0180	10	0.0055	0.0054
$\text{P}_{\text{FA}} = 10^{-1}$, $\sigma_p = 2$ m							
	$\mathcal{B}_{x_V}$	15	0.0260	0.0179	30	0.0168	0.0148
	$\mathcal{B}_{x_H}$	10	0.0213	0.0171	18	0.0137	0.0129
	$\tilde{\mathcal{B}}_{x_H}$	5	0.1011	0.0440	10	0.0215	0.0184
$\text{P}_{\text{FA}} = 10^{-2}$, $\sigma_p = 1$ m							
	$\mathcal{B}_{x_V}$	15	0.0017	0.0015	30	0.0011	0.0011
	$\mathcal{B}_{x_H}$	10	0.0015	0.0015	18	0.0001	0.0001
	$\tilde{\mathcal{B}}_{x_H}$	5	0.0027	0.0024	10	0.0014	0.0014
$\text{P}_{\text{FA}} = 10^{-2}$, $\sigma_p = 2$ m							
	$\mathcal{B}_{x_V}$	15	0.0041	0.0021	30	0.0017	0.0015
	$\mathcal{B}_{x_H}$	10	0.0025	0.0021	18	0.0015	0.0014
	$\tilde{\mathcal{B}}_{x_H}$	5	0.0506	0.0055	10	0.0026	0.0023

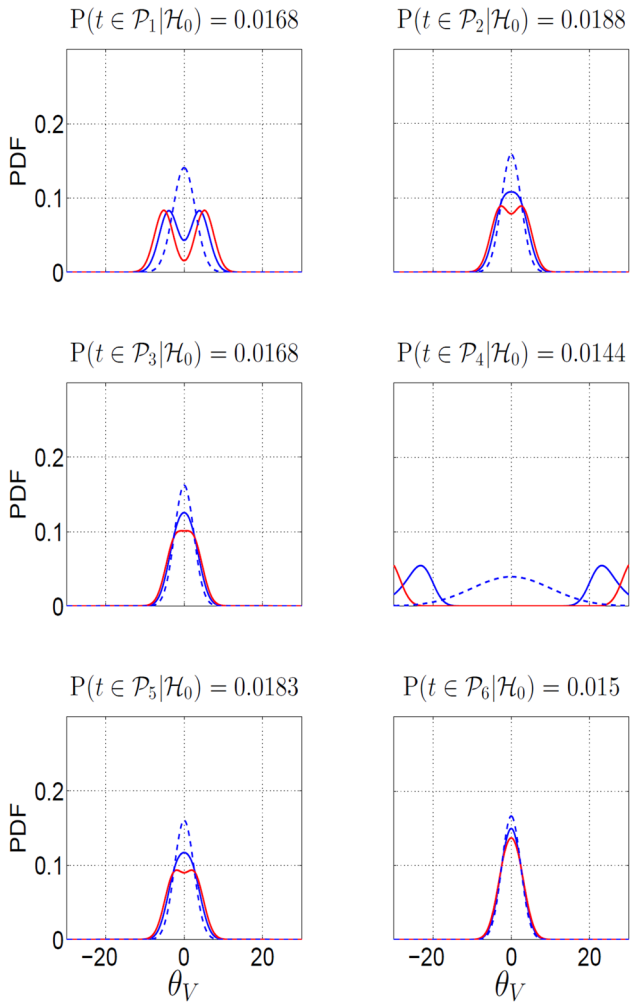

Fig. 6 Illustration of the PDFs $f_{\hat{x}_{V_j}}(\theta_V|\mathcal{H}_0)$ (dashed *blue* curves) and $f_{\hat{x}_{V_j}|t\in\mathcal{P}_j}(\theta_V|t \in \mathcal{P}_j, \mathcal{H}_0)$ for $\alpha = 10^{-1}$ (blue curves) and $\alpha = 10^{-2}$ (*red* curves), for the model in (45), corresponding to the satellite geometry in Fig. 5 and $\sigma_p = 1$m, for all six alternative hypotheses $j = 1, \ldots, 6$. The probability given on top of each panel is computed for $\alpha = 10^{-1}$

behavior for $\mathrm{P}(\hat{x}_{V_j} \in \mathcal{B}^c_{x_V}|t \in \mathcal{P}_j, \mathcal{H}_0) - \mathrm{P}(\hat{x}_{V_j} \in \mathcal{B}^c_{x_V}|\mathcal{H}_0)$ as function of AL.

In Fig. 5, when σ_p increases, the graph of $(\overline{\mathrm{IR}}|\mathcal{H}_0 - \mathrm{IR}^o|\mathcal{H}_0)$ stretches over a larger range of AL. This is due to the fact that increasing σ_p reduces the peakedness of the PDFs $f_{\hat{x}_{*_j}}(\theta_*|\mathcal{H}_0)$ and $f_{\hat{x}_{*_j}|t\in\mathcal{P}_j}(\theta_*|t \in \mathcal{P}_j, \mathcal{H}_0)$ around zero. Therefore, both the probabilities $\mathrm{P}(\hat{x}_{*_j} \in \mathcal{B}^c_{x_*}|t \in \mathcal{P}_j, \mathcal{H}_0)$ and $\mathrm{P}(\hat{x}_{*_j} \in \mathcal{B}^c_{x_*}|\mathcal{H}_0)$ behave more smoothly as function of AL, and so does their difference.

Decreasing $\alpha = \mathrm{P_{FA}}$ by a factor of 10, we note that the graphs of $(\overline{\mathrm{IR}}|\mathcal{H}_0 - \mathrm{IR}^o|\mathcal{H}_0)$ for both the vertical and horizontal components also decrease by almost a factor of 10. In (46), the dependence on the false alarm probability α is introduced by $\mathrm{P}(t \in \mathcal{P}_j|\mathcal{H}_0)$ and $\mathrm{P}(\hat{x}_j \in \mathcal{B}^c_x|t \in \mathcal{P}_j, \mathcal{H}_0)$. With the definition of $\mathcal{P}_0$ in (6) and $\mathcal{P}_{j\neq 0}$ in (8), for some scalar $0 < s < 1$, we have

$$\alpha \to s \times \alpha \Rightarrow \mathrm{P}(t \in \mathcal{P}_{j\neq 0}|\mathcal{H}_0) \to s \times \mathrm{P}\big(t \in \mathcal{P}_{j\neq 0}|\mathcal{H}_0\big). \quad (51)$$

Therefore, if α reduces, for instance, by a factor of 10, then the probability $\mathrm{P}\big(t \in \mathcal{P}_{j\neq 0}|\mathcal{H}_0\big)$ also decreases by a factor of 10. For the vertical component, as Fig. 6 shows, changing α from 10^{-1} to 10^{-2} does not significantly affect the shape of $f_{\hat{x}_{V_j}|t\in\mathcal{P}_j}\big(\theta_V|t \in \mathcal{P}_j, \mathcal{H}_0\big)$ for $j = 2, 3, 5, 6$. Therefore, the leading factors driving the difference between $(\overline{\mathrm{IR}}|\mathcal{H}_0 - \mathrm{IR}^o|\mathcal{H}_0)$ for $\alpha = 10^{-1}$ and $(\overline{\mathrm{IR}}|\mathcal{H}_0 - \mathrm{IR}^o|\mathcal{H}_0)$ for $\alpha = 10^{-2}$ are the probabilities $\mathrm{P}\big(t \in \mathcal{P}_{j\neq 0}|\mathcal{H}_0\big)$ which, according to (51), get 10 times smaller when reducing α from 10^{-1} to 10^{-2}. Therefore, the reduction of $(\overline{\mathrm{IR}}|\mathcal{H}_0 - \mathrm{IR}^o|\mathcal{H}_0)$ due to reducing α from 10^{-1} to 10^{-2} is expected to be by almost one order of magnitude.

From Fig. 6, we note that the difference between the two PDFs $f_{\hat{x}_{V_j}}\big(\theta_V|\mathcal{H}_0\big)$ and $f_{\hat{x}_{V_j}|t\in\mathcal{P}_j}\big(\theta_V|t \in \mathcal{P}_j, \mathcal{H}_0\big)$ can be arranged in an ascending order for $j = 6, 3, 5, 2, 1, 4$. To understand the impact of conditioning on $t \in \mathcal{P}_j$ on $\hat{x}_j$, we consider (11), which describes the link between $\hat{x}_j$ and t established through L_j, cf. (12). For the observational model at hand, in which the redundancy is $r = 2$ and also $x \in \mathbb{R}^3$, L_j is a 3×2 matrix which can be decomposed as $L_j = [l_{j,1}, l_{j,2}, l_{j,3}]^T$ with $l_{j,i} \in \mathbb{R}^2 (i = 1, 2, 3)$. The conditional PDF $f_{\hat{x}_{V_j}|t\in\mathcal{P}_j}\big(\theta_V|t \in \mathcal{P}_j, \mathcal{H}_0\big)$ is then given by

$$f_{\hat{x}_{V_j}|t\in\mathcal{P}_j}(\theta_V|t \in \mathcal{P}_j, \mathcal{H}_0) = \frac{1}{\mathrm{P}(t \in \mathcal{P}_j|\mathcal{H}_0)} \times \int_{\mathcal{P}_j} f_t(\tau|\mathcal{H}_0) f_{\hat{x}_{V_0}}(\theta_V + l^T_{j,3}\tau|\mathcal{H}_0)\,\mathrm{d}\tau \quad (52)$$

in which

$$l^T_{j,3}\tau = -\frac{w_j(\tau)}{||c_{t_j}||_{Q_{tt}}} q^T_{\hat{x}_0\hat{x}_{V_0}}\big(u_j - \bar{u}\big) \quad (53)$$

with $q_{\hat{x}_0\hat{x}_{V_0}} \in \mathbb{R}^3$ the covariance vector between $\hat{x}_0$ and $\hat{x}_{V_0}$, $w_j(\tau)$ given by (7), and the average receiver-satellite unit direction vector $\bar{u} = \frac{1}{m}\sum_{j=1}^m u_j$. The above equations reveal that the impact of L_j on the vertical component $\hat{x}_{V_j}$ is governed by various factors. Table 3 gives the values of $|q^T_{\hat{x}_0\hat{x}_{V_0}}(u_j - \bar{u})| \Big/ ||c_{t_j}||_{Q_{tt}}$ in (53) for $\sigma_p = 1$m for all hypotheses $j = 1, \ldots, 6$. As the value of $\mathrm{P}\big(t \in \mathcal{P}_j|\mathcal{H}_0\big)$ does not

Table 3 Evaluation of (53) for Example 1 (Fig. 5) assuming $\sigma_p = 1\text{m}$

j	6	3	5	2	1	4
$\lvert q^T_{\hat{x}_0\hat{x}_{V_0}}(u_j - \bar{u})\rvert \Big/ \lvert\lvert c_{t_j}\rvert\rvert_{Q_{tt}}$	0.47	0.73	0.80	0.89	1.58	9.79

change too much for different j (see Fig. 6), it can be stated that the shapes of the regions $\mathcal{P}_j$ are similar to each other. Therefore, the leading factor driving the difference in the shape of $f_{\hat{x}_{V_j}|t\in\mathcal{P}_j}(\theta_V|t \in \mathcal{P}_j, \mathcal{H}_0)$ is $\lvert q^T_{\hat{x}_0\hat{x}_{V_0}}(u_j - \bar{u})\rvert \Big/ \lvert\lvert c_{t_j}\rvert\rvert_{Q_{tt}}$ which is by far greatest for $j = 4$. This explains the discrepancy between $f_{\hat{x}_{V_4}|t\in\mathcal{P}_4}(\theta_V|t \in \mathcal{P}_4, \mathcal{H}_0)$ and $f_{\hat{x}_{V_{j\neq4}}|t\in\mathcal{P}_{j\neq4}}(\theta_V|t \in \mathcal{P}_{j\neq4}, \mathcal{H}_0)$. The red and blue curves for $j = 4$ in Fig. 6 (middle right) really differ from those in the other panels.

Example 2: Fig. 7

Figure 7 presents the same type of information as Fig. 5 but for a different geometry of six satellites. For this example, again we note the 'positive' values for $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$ in case of both the vertical and horizontal components. The difference in integrity risk (rigorous minus approximate) in general behaves similar to the earlier example.

Example 3: Fig. 8

In Fig. 8, we present the same type of information as in Figs. 5 and 7, but now for a multi-constellation example given in Blanch et al. (2012), Appendix J. This constellation is made of five GPS (G) and five Galileo (E) satellites. In this case, we have three position parameters and two receiver clock parameters (one for GPS and one for Galileo). The misclosure vector is of the dimension of five. Note that the results in Fig. 8 use an elevation-dependent variance (C_{int} in the mentioned paper) for the observations. Also here, like in the previous examples, we note the 'positive' values for $(\overline{\text{IR}}|\mathcal{H}_0 - \text{IR}^o|\mathcal{H}_0)$.

Summary and conclusion

The message of this contribution finds its origin in the combination of parameter estimation and statistical testing. These two activities are typically disconnected in practice when it comes to describing the quality of the eventual estimator. That is, the distribution of the estimator under an identified statistical hypothesis is used without regard to the conditioning process that led to the decision to accept this hypothesis as the working model. We analyzed what the contribution of this simplification is to the actual integrity risk.

Considering a null hypothesis and a single alternative, the different distributions were first shown graphically for a simple observational model with a one-dimensional unknown parameter and a one-dimensional misclosure for statistical testing. It was demonstrated that, with normally distributed observables and linear models, the distributions of the estimators conditioned on false alarm and correct detection turn out to be no longer normal. To compute the integrity risk rigorously, one needs to condition on both the hypothesis and the testing outcome. An approximate risk is obtained, however, when one omits the connection between testing and estimation and thus only conditions on the hypothesis and not on the testing outcome.

For the simple observational model, it was mathematically proven that the rigorous integrity risk exceeds the approximate one. This comparison of the rigorous and approximate integrity risk was then continued by means of a number of satellite-based single-point positioning examples, focusing again on the contributions under the null hypothesis. Although a mathematical proof for the multi-dimensional case does not yet exist, these examples support the previously obtained conclusion that the approximate integrity risk has a tendency of being smaller than its rigorous counterpart. Thus, by including the uncertainty of the decision process driven by statistical testing and using conditional distributions instead of unconditional ones, the actual integrity risk may end up being larger than the computed approximate one. In other words, the approximate approach may provide a too optimistic description of the integrity risk and thereby not sufficiently safeguard against possibly hazardous situations.

We, therefore, clearly advocate the use of the rigorous approach to evaluate the integrity risk, as underestimating the risk, or knowingly allowing this possibility to exist, cannot be acceptable particularly in critical and safety-of-life applications.

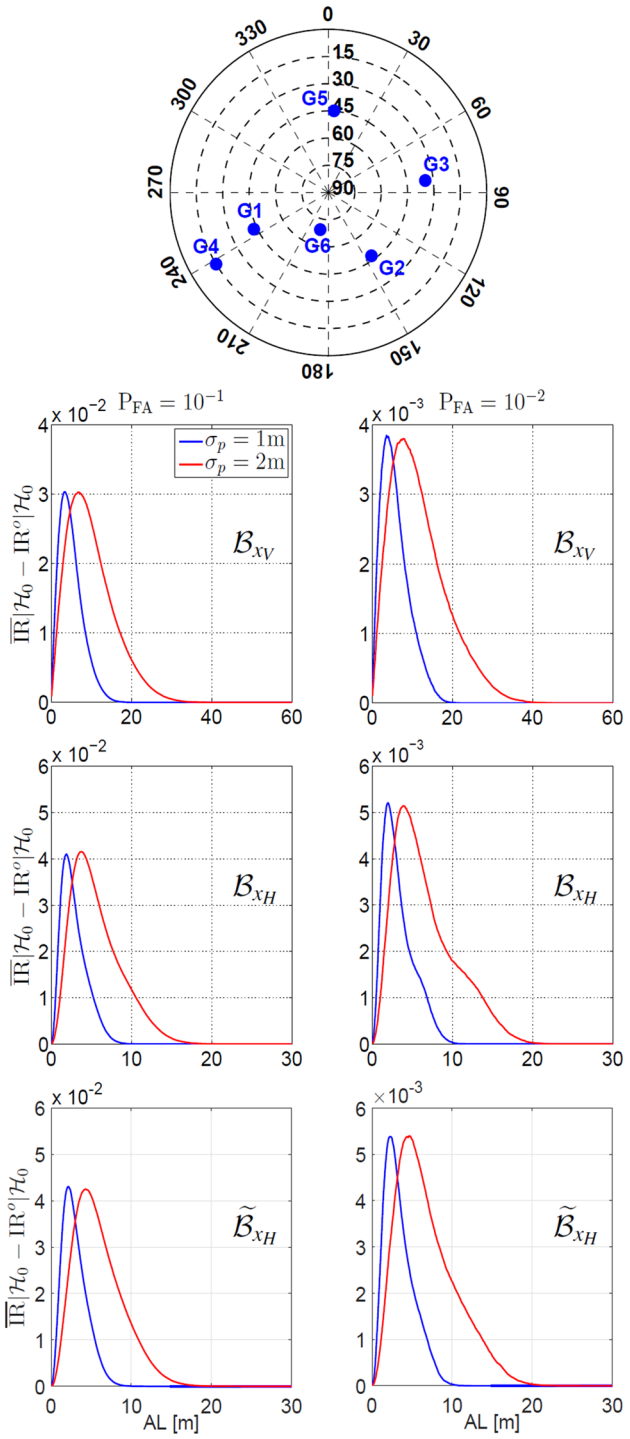

Fig. 7 Skyplot of the satellites (top), the difference $(\overline{\mathrm{IR}}|\mathcal{H}_0 - \mathrm{IR}^o|\mathcal{H}_0)$ for vertical component (second row) and horizontal components (third and fourth rows) in the model given in (45), as a function of alert limit AL, for different values of σ_p (cf. 45) and $\mathrm{P_{FA}} = \alpha$

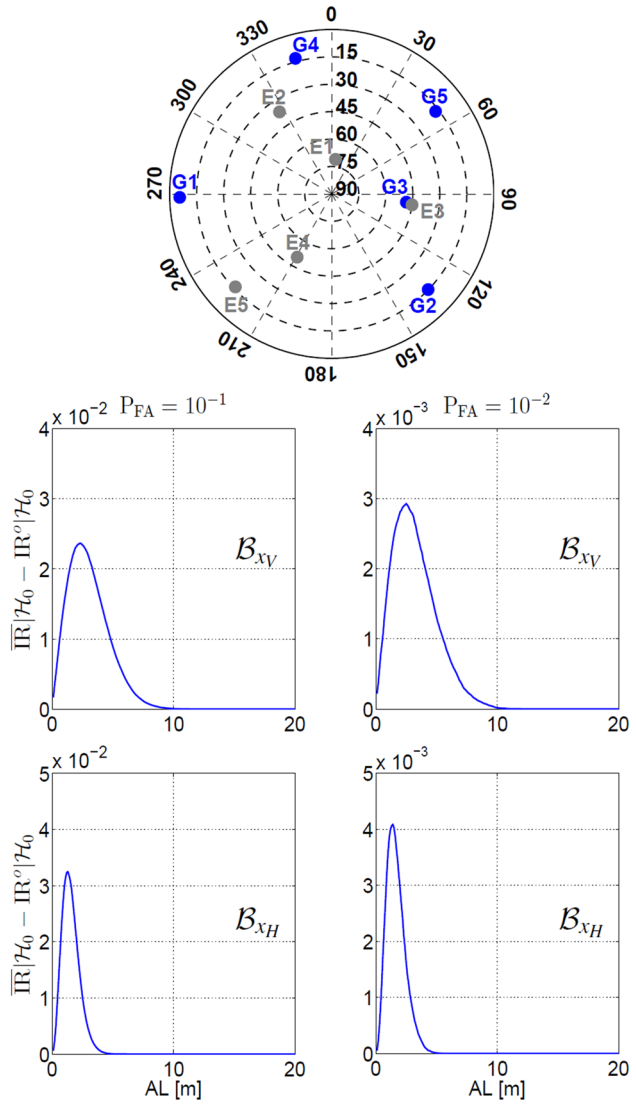

Fig. 8 Skyplot of the satellites (top), the difference $(\overline{\mathrm{IR}}|\mathcal{H}_0 - \mathrm{IR}^o|\mathcal{H}_0)$ for vertical component (second row) and horizontal components (third row) in the model given in Blanch et al. (2012), Appendix J, as a function of alert limit AL, for different values of $\mathrm{P_{FA}} = \alpha$

Acknowledgements The second author is the recipient of an Australian Research Council (ARC) Federation Fellowship (project number FF0883188). This support is greatly acknowledged.

Appendix

To prove that the difference $(\overline{\mathrm{IR}}|\mathcal{H}_0 - \mathrm{IR}^o|\mathcal{H}_0)$ in (42) is always positive, consider the following expression:

$$I = \int_{\mathcal{B}_x} \left[f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA}) - f_{\hat{x}_a}\big(\theta|\mathcal{H}_0\big) \right] \mathrm{d}\theta \tag{54}$$

which, given $\mathrm{P}(t \notin \mathcal{P}_0|\mathcal{H}_0) = \alpha$ and

$$\begin{aligned} f_{\hat{x}_a|\mathrm{FA}}(\theta|\mathrm{FA}) &= \frac{1}{\alpha} \int_{\mathcal{P}_0^c} f_{\hat{x}_a,t}\big(\theta,\tau|\mathcal{H}_0\big)\mathrm{d}\tau \\ f_{\hat{x}_a}\big(\theta|\mathcal{H}_0\big) &= \int_{\mathbb{R}} f_{\hat{x}_a,t}\big(\theta,\tau|\mathcal{H}_0\big)\mathrm{d}\tau \end{aligned} \tag{55}$$

can be worked out as

$$I = I_1 - I_2 \tag{56}$$

with

$$\begin{aligned} I_1 &= \int_{\mathcal{P}_0^c} \Big(\frac{1}{\alpha} - 1\Big) f_t\big(\tau|\mathcal{H}_0\big) \int_{\mathcal{B}_x} f_{\hat{x}_0}\big(\theta + L_a\tau|\mathcal{H}_0\big)\mathrm{d}\theta\mathrm{d}\tau \\ I_2 &= \int_{\mathcal{P}_0} f_t\big(\tau|\mathcal{H}_0\big) \int_{\mathcal{B}_x} f_{\hat{x}_0}\big(\theta + L_a\tau|\mathcal{H}_0\big)\mathrm{d}\theta\mathrm{d}\tau \end{aligned} \tag{57}$$

As $\hat{x}_0 \overset{\mathcal{H}_0}{\sim} \mathcal{N}(x, \sigma^2_{\hat{x}_0})$ and also $\mathcal{B}_x = [x - \mathrm{AL}, x + \mathrm{AL}]$ is a convex x-centered region, for some Alert Limit $\mathrm{AL} > 0$, we have

$$|\tau_1| \geq |\tau_2| \Rightarrow \int_{\mathcal{B}_x} f_{\hat{x}_0}\big(\theta + L_a\tau_1|\mathcal{H}_0\big)\mathrm{d}\theta \leq \int_{\mathcal{B}_x} f_{\hat{x}_0}\big(\theta + L_a\tau_2|\mathcal{H}_0\big)\mathrm{d}\theta \tag{58}$$

which, given the definition of $\mathcal{P}_0$ in (32), results in

$$\begin{aligned} I_1 &\leqslant \int_{\mathcal{P}_0^c} \Big(\frac{1}{\alpha} - 1\Big) f_t\big(\tau|\mathcal{H}_0\big)\mathrm{d}\tau \int_{\mathcal{B}_x} f_{\hat{x}_0}\Big(\theta + L_a\sqrt{k_{\alpha,1}}\Big|\mathcal{H}_0\Big)\mathrm{d}\theta \\ &= \int_{\mathcal{P}_0} f_t\big(\tau|\mathcal{H}_0\big)\mathrm{d}\tau \int_{\mathcal{B}_x} f_{\hat{x}_0}\Big(\theta + L_a\sqrt{k_{\alpha,1}}\Big|\mathcal{H}_0\Big)\mathrm{d}\theta \\ &< \int_{\mathcal{P}_0} f_t\big(\tau|\mathcal{H}_0\big) \int_{\mathcal{B}_x} f_{\hat{x}_0}\Big(\theta + L_a\tau\Big|\mathcal{H}_0\Big)\mathrm{d}\theta\mathrm{d}\tau \\ &= I_2 \end{aligned} \tag{59}$$

in which the second equality follows from $\mathrm{P}\big(t \in \mathcal{P}_0|\mathcal{H}_0\big) = \Big(\frac{1}{\alpha} - 1\Big) \times \mathrm{P}\big(t \in \mathcal{P}_0^c|\mathcal{H}_0\big)$, and the third inequality follows from (48) given that $|\tau| < \sqrt{k_{\alpha,1}}$ for any $\tau \in \mathcal{P}_0$. From (59), we have $I < 0$, and with $\mathcal{B}_x^c = \mathbb{R}^n/\mathcal{B}_x$ we get for (42)

$$\Big(\overline{\mathrm{IR}}|\mathcal{H}_0 - \mathrm{IR}^o|\mathcal{H}_0\Big) > 0 \tag{60}$$

References

Amiri Simkooei A (2001) Comparison of reliability and geometrical strength criteria in geodetic networks. J Geodesy 75(4):227–233

Baarda W (1967) Statistical concepts in geodesy. Netherlands Geodetic Commission, Publications on Geodesy, New Series, Vol. 2, No. 4, pp 74

Blanch J, Walter T, Enge P, Lee Y, Pervan B, RipplM, Spletter A (2012) Advanced RAIM user algorithm description: integrity support message processing, fault detection, exclusion, and protection level calculation. In: Proc. of ION ITM-2012, ION, pp 2828–2849

Devoti R, Esposito A, Pietrantonio G, Pisani AR, Riguzzi F (2011) Evidence of large scale deformation patterns from GPS data in the Italian subduction boundary. Earth Planet Sci Lett 311:230–241

Dheenathayalan P, Small D, Schubert A, Schubert A, Hanssen RF (2016) High-precision positioning of radar scatterers. J Geodesy 90(5):403–422

Durdag UM, Hekimoglu S, Erdoğan B (2018) Reliability of models in kinematic deformation analysis. ASCE J Surv Eng 144(3):040118004

Gillissen I, Elema IA (1996) Test results of DIA: a real-time adaptive integrity monitoring procedure, used in an integrated navigation system. Int Hydrogr Rev 73(1):75–100, March 1996

Hewitson S, Wang J (2006) GNSS receiver autonomous integrity monitoring (RAIM) performance analysis. GPS Solut 10:155–170

Jonkman N, de Jong K (2000) Integrity monitoring of IGEX-98 data, part II: cycle slip and outlier detection. GPS Solut 3(4):24–34

Kok JJ, NOAA, National Oceanic, Atmospheric Administration (1984) On data snooping and multiple outlier testing. NOAA technical report NOS NGS:30;1984, U.S. Department ofCommerce, National Oceanic and Atmospheric Administration (NOAA). https://books.google.com.au/books?id=tJqqMQEACAAJ

Kösters AJM, van der Marel H (1990) Statistical testing and quality analysis of 3-D networks (parts I and II). In: Bock Y, Leppard N (eds) Global positioning system: an overview, vol 102. International Association of Geodesy Symposia, Springer, New York

Kuusniemi H, Lachapelle G, Takala JH (2004) Position and velocity reliability testing in degraded GPS signal environments. GPS Solut 8(4):226–237

Perfetti N (2006) Detection of station coordinate discontinuities within the Italian GPS Fiducial Network. J Geodesy 80(7):381–396

Salós D, Macabiau C, Martineau A, Bonhoure B, Kubrak D (2010) Analysis of GNSS integrity requirements for road user charging applications. In: Satellite navigation technologies and European workshop on GNSS signals and signal processing (NAVITEC), 2010 5th ESA Workshop on. IEEE, 2010, IEEE, pp 1–8

Schuster W, Bai J, Feng S, Ochieng W (2007) Airport surface movement–performance requirements, architecture considerations and integrity algorithms. In: Proc. of ION GNSS 2007, ION, pp 2950–2961

Teunissen PJG (1990a) An integrity and quality control procedure for use in multi-sensor integration. In: Proc. of ION GPS 1990, ION, pp 513–522

Teunissen PJG (1990b) Quality control in integrated navigation systems. IEEE Aerosp Electron Syst Mag 5(7):35–41

Teunissen PJG (2000) Testing theory: an introduction. Series on Mathematical Geodesy. Delft University Press, Delft
Teunissen PJG (2017) Distributional theory for the DIA method. J Geodesy. https://doi.org/10.1007/s00190-017-1045-7
Teunissen PJG, Imparato D, Tiberius CCJM (2017) Does RAIM with correct exclusion produce unbiased positions? Sensors 17(7):1508
Verhoef HME, de Heus HM (1995) On the estimation of polynomial breakpoints in the subsidence of the Groningen gasfield. Surv Rev 33(255):17–30
Wieser A (2004) Reliability checking for GNSS baseline and network processing. GPS Solut 8(3):55–66
Wu Y, Wang J, Jiang Y (2013) Advanced receiver autonomous integrity monitoring (ARAIM) schemes with GNSS time offsets. Adv Space Res 52:52–61
Yang L, Li Y, Wu Y, Rizos C (2014) An enhanced MEMS-INS/GNSS integrated system with fault detection and exclusion capability for land vehicle navigation in urban areas. GPS Solut 18:593–603
Yavaşoğlu HH, Kalkan Y, Tiryakioğlu İ, Yigit CO, Özbey V, Alkan MN, Bilgi S, Alkan RM (2017) Monitoring the deformation and strain analysis on the Ataturk Dam, Turkey. Geomat Nat Hazards Risk 9(1):94–107

Safoora Zaminpardaz is a lecturer in Geospatial Sciences at RMIT University, Melbourne, Australia. Her research interests include array-based multi-GNSS positioning, integrity monitoring and ionosphere sensing.

Peter J. G. Teunissen is Professor of Geodesy and Navigation at Curtin University, Perth, Australia, and Delft University of Technology, Delft, The Netherlands. His current research focuses on multi-GNSS and the modeling of next-generation GNSS for high-precision positioning, navigation and timing applications.

Christiaan C. J. M. Tiberius is an associate professor at Delft University of Technology. He has been involved in GNSS positioning and navigation research since 1991, currently with emphasis on data quality control, satellite-based augmentation systems and precise point positioning.