Fault Detection and Isolation in Multiple MEMS-IMUs Configurations

STÉPHANE GUERRIER
University of Geneva

ADRIAN WAEGLI
JAN SKALOUD
Ecole Polytechnique fédérale de Lausanne

MARIA-PIA VICTORIA-FESER
University of Geneva

This research presents methods for detecting and isolating faults in multiple micro-electro-mechanical system inertial measurement unit (MEMS-IMU) configurations. First, geometric configurations with n sensor triads are investigated. It is proved that the relative orientation between sensor triads is irrelevant to system optimality in the absence of failures. Then, the impact of sensor failure or decreased performance is investigated. Three fault detection and isolation (FDI) approaches (i.e., the parity space method, Mahalanobis distance method and its direct robustification) are reviewed theoretically and in the context of experiments using reference signals. It is shown that in the presence of multiple outliers the best performing detection algorithm is the robust version of the Mahalanobis distance.

Manuscript received November 18, 2010; revised May 5, 2011; released for publication June 9, 2011.

IEEE Log No. T-AES/48/3/943999.

Refereeing of this contribution was handled by M. Blanke.

This research was partially financed by TracEdge, based at Grenoble, France.

Authors' addresses: S. Guerrier and M-P. Victoria-Feser, Research Center for Statistics, HEC, University of Geneva, 40, Bd du Pont d'Arve, 1211 Geneva 4, Switzerland, E-mail: (Stephane.Guerrier@unige.ch); A. Waegli and J. Skaloud, Geodetic Engineering Lab (TOPO), Ecole Polytechnique fédérale de Lausanne, Route Cantonale, 1015 Lausanne, Switzerland.

I. INTRODUCTION

Because of their small size and cost, micro-electro-mechanical system (MEMS)-type inertial and magnetic sensors have a great potential for navigation [1]. Therefore, these sensors are increasingly used and the research focuses on modeling and compensation of their large and variable measurement errors. Although sensor calibration by the manufacturers (temperature, misalignments) removes its important part, the residual effects do not provide sufficiently accurate measurements for standalone inertial navigation, and permanent sensor error compensation during navigation becomes mandatory. However, the quality of the navigation parameters can be considerably improved when several MEMS inertial measurement unit (MEMS-IMUs) triads works in parallel, provided faulty measurements are detected and isolated from the estimation process [2, 3]. The resulting cost and volume of such system still remains very attractive.

In general, the task of fault detection and isolation (FDI) includes the detection of the presence of failure (or outlier) and the isolation of the component responsible for the irregularity. FDI is applied in a wide range of domains such as in the process industries (e.g. chemical plants), the aerospace industry (e.g. high-performance airplanes) and more recently in some mass-produced equipments (e.g. automobiles) [4].

Generally, there is a conceptual difference between a faulty observation and an outlier. A faulty observation corresponds to a measurement that is generated from another unknown statistical process than the rest of the observations. An outlier is a measurement that is "far" from the rest of the data which may or may not be a faulty observation. However, in practice, faulty observations and outliers are impossible to discriminate. In addition, even if it were possible to discriminate them, their impacts on the estimation of the model parameters is equivalent. Thus, faulty observations or outliers receive statistically equivalent treatment. For these reasons, we do not make any distinction in this article between the two concepts.

In the inertial navigation framework, FDI algorithms were traditionally applied to safety-critical operations such as in the control of military or space aircrafts. These aircrafts are designed to be dynamically unstable, in order to increase manoeuvrability, while inertial accelerations and rotation rates are used to observe the vehicle's stabilizing parameters [5]. Such systems require sensor redundancy for obvious safety considerations. However, multiple MEMS-IMUs systems are not yet used for safety-critical applications. In this context, FDI aims at detecting gross errors, enhancing navigation performance, and improving the stability

of the filters used in the GPS (Global Positioning System)/INS (inertial navigation system) integration.

FDI algorithms applied to INS have been thoroughly investigated. The most commonly used approach is by far the parity space (vector) method, but other approaches such as artificial neural networks have also been examined [6]. Indeed, the first FDI algorithms were applied strapdown redundant inertial systems in a dodecahedron configuration [7–9]. For example, [9] developed the "minimax" FDI algorithm which enables to test and compare any set of four sensors. A review and a comparison of earlier FDI algorithms can be found in [10]. Unfortunately, these algorithms did not allow to link the integrity of the (navigation) system with the employed FDI method. This limitation was overcome with the generalized likelihood ratio test (GLRT) approach introduced in [11] and adapted for the detection of failures in redundant inertial systems in [12], [13]. This approach is based on ratio of the (log) likelihood given that no failure and given that a failure occurred. The main advantage of this method is that the performance of the FDI algorithms can be directly related to the integrity of the system. Since then, various improved versions of this method have been developed to detect measurement failures in GPS and redundant IMU systems [14–16].

In [17], [18] the parity space approach was introduced to detect failures in redundant inertial systems and GPS signals. This approach relies on the basic assumption that the residuals used in FDI are independent and identically normally distributed. When applied to IMUs of higher accuracy, this method gives satistifactory results, however it performs poorly with low-cost sensors [2, 3]. Indeed, the complexity of implementation of an efficient FDI system is increased when using MEMS-type IMUs. Their poor performances (i.e., noise density variations in time and among sensors, large systematic measurement errors compared with the random errors) often creates false alarms and increases the possibility of misdetection of faulty measurements. Additionally, it has been shown in [2], [19] that MEMS-IMUs are strongly influenced by variation of the environmental conditions (e.g. increased vibrations or temperature variations) which can introduce correlation among the residuals. Note that recent developments proposed improvements of the parity approach [20, 21]. For example, [20] proposed to combine the parity equation approach with a wavelet-based signal decomposition. A complete and up-to-date review of FDI algorithms applied to inertial sensors can be found in [22].

Nevertheless, from a statistical point of view, the already presented FDI approaches are relatively similar. Indeed, most of these methods rely on the same assumption (i.e., normally independently and identically distributed errors) even if they are based on different test statistics. In this article, we intend to relax the assumption of independent and homoscedastic errors. We propose employing a method based on Mahalanobis distances to detect erroneous measurements and we compare its performance with the classical parity space approach proposed in [18]. Moreover, we also intend to show the importance of robust estimation in the context of FDI for inertial observations from low-cost sensors.

Indeed, outside the context of inertial navigation other methods were developed for FDI purposes that take into account the correlation and the variance differences between residuals. Such methods can be found in chemical (and geochemical) engineering [4, 23–25] or in the classical statistical literature [26–28]. These methods take into account the shape and size of multivariate data (quantified by the covariance matrix) and are based on Mahalanobis distances (or on the Hotelling T^2 statistic).

This article is organized as follows.

First, the optimal geometry of redundant IMU systems is discussed. The relative geometry plays a central role in such systems since it determines how the information is spread among the axes of the sensor/body frame. Hence, a clear understanding of the geometric influence on redundant systems is an indispensable prerequisite for designing efficient FDI approaches.

Second, the FDI schemes are presented. Alternatives that are more suitable for MEMS-type sensors than the conventional parity space method are introduced.

Third, the performance of the algorithms is illustrated based on experiments with reference signals from higher grade inertial sensors.

II. GEOMETRY OF MULTIPLE IMU SYSTEMS

Since the early days of the inertial technology, redundant IMUs have been used in safety-critical operations. The redundant information was used to create fault-tolerant systems which were able to detect and isolate defective sensors. If two sensors are placed collinear to each other, it is possible to detect a fault that occurs in either one of the sensors. However, to isolate the erroneous device at least three sensors are required. Therefore, traditionally, three triple-axis IMUs were used (three sensors per axis) [5].

In 1974, A. Pejsa [29] developed a theory to optimally arrange any number of sensors. At the same time he showed that less than nine sensors (four in theory) are sufficient to correctly isolate faults in a three-dimensional space. This theory essentially considers situations where sensors are equally placed on a cone of half-angle α. In this case, the optimal half-angle α corresponds to the configuration in which the system variance is minimized.

A. Information Filters

Reference [5] proposes employing information filters to achieve optimal configuration by maximizing the amount of information from redundant systems. In short, this approach allows to analyze the selected configuration and provides a geometrical understanding of the distribution of information across axes. It is based on the information matrix $\mathbf{I}$ defined as

$$\mathbf{I} = \mathbf{H}^T\mathbf{S}^{-1}\mathbf{H} \tag{1}$$

where $\mathbf{S}$ is the covariance matrix of the observations and $\mathbf{H}$ is the design matrix of the system. Note that the matrix $\mathbf{I}$ is independent from the observations, since the observation vector is absent from (1). Thus, it provides a measure of certainty that is purely based on the system geometry. Since the determinant of the information matrix is directly related to the volume of the ellipsoid formed by the eigenvectors of $\mathbf{I}$ (which corresponds to the volume of information in the system), the optimization aims to maximize:

$$J = \det(\mathbf{I}). \tag{2}$$

It has been shown in [5] that optimal configurations adopt the shape of a sphere in the information space (i.e., equal amount of information is placed on each axis).

B. Partial Redundancy Method

Although the information filter approach relates the volume of information within the system to the determinability of the parameters, we do not believe that this represents an optimal criterion for FDI purposes. Indeed, reference [2] proposed another approach to determine optimal configurations for redundant inertial sensors that seems more adequate in this context. This method is typically applied in geodetic networks and has (to the authors' knowledge) never been applied in a context similar to that of this study. Indeed, redundant measurements in geodetic networks control each other. In order to assess the "amount" of controllability of each measurement, a z_i value can be computed and associated to every observation. This value, called partial redundancy, corresponds to the participation of the observation i to the global redundancy of the system. It varies from 0 to 1; an observation having an associated $z_i = 0$ is not controlled at all by others measurements. On the contrary, a $z_i = 1$ implies a "totally" controlled observation. In geodetic engineering, a network is typically considered good (i.e., technically and economically) if all measurements have an associated $z_i \in (0.25, 0.60)$ [30, 31]. In statistics, the z_i (called leverages) are used in the regression framework to control extreme values in the design matrix. Indeed, when z_i is small (near zero), an unexpected error in the response variable might not be reflected in the residuals [32].

Moreover, it can be demonstrated that the geometry of the system will influence z_i but summation of all z_i remains constant (for proof see [2, 30]). Indeed,

$$\sum_{i=1}^{m} z_i = m - p \tag{3}$$

where m is the number of observations and p the number of parameters of the considered system. Consequently, an optimal sensor configuration would minimize differences amongst measurements' controllability approaching a situation where observations are controlled equally. Mathematically, this is equivalent to

$$\min \sqrt{\mathbb{E}[(\mathbf{z} - \mathbb{E}[\mathbf{z}])^2]} \tag{4}$$

where $\mathbf{z}$ (i.e., the vector of z_i) is the diagonal of the matrix $\mathbf{Z}$, which is defined as

$$\mathbf{Z} = \mathbf{I}_m - \mathbf{H}(\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{S}^{-1}. \tag{5}$$

It was shown in [2] that the partial redundancies approach and the information filter approach give, in general, similar results. That said, in certain cases, the partial redundancies approach is a better indicator of the system's optimality (see Section IID).

C. Optimal Configuration with n Sensor Triads

Intuitively, skew-redundant configurations of sensors have higher information volume and are thus preferable. This strategy was adopted in [3], [5], [33]. However, as far as sensor triads are concerned, it will be demonstrated that their respective orientation is unimportant to configuration optimality with respect to the determinability of the parameters (information filters) and to the capacity of the model to detect and isolate errors (partial redundancies). Indeed, consider a situation of n sensors randomly oriented. The associated design matrix can be expressed as

$$\mathbf{H} = \begin{bmatrix} \mathbf{R}(\lambda_1, \theta_1) \\ \vdots \\ \mathbf{R}(\lambda_n, \theta_n) \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1 \\ \vdots \\ \mathbf{R}_n \end{bmatrix}$$

where $\mathbf{R}(\lambda_i, \theta_i)$ (i.e., $\mathbf{R}_i$) corresponds to the product of two rotation matrices and is defined as

$$\mathbf{R}_i = \begin{bmatrix} -\sin(\lambda_i)\cos(\theta_i) & -\sin(\theta_i) & -\cos(\lambda_i)\cos(\theta_i) \\ -\sin(\lambda_i)\sin(\theta_i) & \cos(\theta_i) & -\cos(\lambda_i)\sin(\theta_i) \\ \cos(\lambda_i) & 0 & -\sin(\lambda_i) \end{bmatrix} \tag{6}$$

with λ_i and θ_i denoting the rotation angles around the first and the second axes of the arbitrary sensor-frame (see Fig. 1).

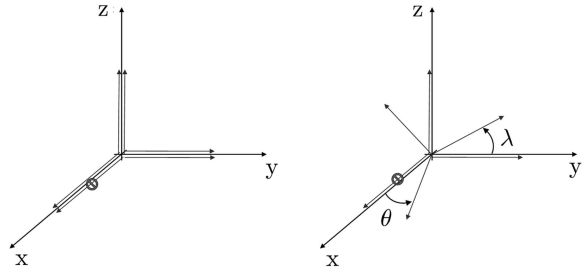

Fig. 1. Two sensor triads of equal precision are placed in two different configurations. Symbol ⊘ indicates sensors that will fail (superscript *).

Note that in our study we consider the individual sensors of an IMU (i.e., accelerometers and gyroscopes) to be sensitive along one principal axis and the placement along such axis to be arbitrary and not relevant to its output. For example, an accelerometer can be mounted along its sensitive axis under different roll angle without affecting its observation. Under such assumption, the rotation matrix describing the orientation of the individual sensors with respect to the body-frame is defined by two fundamental rotations with angles θ and λ, respectively (see (6)).

Note that, because $\mathbf{R}_i$ is an orthogonal matrix, it has the following property: $\mathbf{R}_i\mathbf{R}_i^T = \mathbf{I}_3$. Furthermore, assuming independent sensors triads each composed of 3 gyros (or accelerometers) of equal precisions, the covariance matrix of the observations $\mathbf{S}$ can be defined as follows:

$$\mathbf{S} = \mathbf{\Sigma}_S \otimes \mathbf{I}_3 \qquad \text{where} \quad \mathbf{\Sigma}_S = \begin{bmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

where $n = m/3$ is the number of IMUs employed. Then, using (5),

$$\begin{aligned} \mathbf{Z} &= \mathbf{I}_{3n} - \mathbf{H}(\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{S}^{-1} \\ &= \mathbf{I}_{3n} - \mathbf{H}(\mathbf{H}^T(\mathbf{\Sigma}_S^{-1} \otimes \mathbf{I}_3)\mathbf{H})^{-1}\mathbf{H}^T(\mathbf{\Sigma}_S^{-1} \otimes \mathbf{I}_3) \\ &= \mathbf{I}_{3n} - \mathbf{H}\left(\sum_{j=1}^{n} \frac{1}{\sigma_j^2}\mathbf{I}_3\right)^{-1} \mathbf{H}^T(\mathbf{\Sigma}_S^{-1} \otimes \mathbf{I}_3) \\ &= \mathbf{I}_{3n} - \frac{1}{\sum_{j=1}^{n} \frac{1}{\sigma_j^2}} \begin{bmatrix} \frac{1}{\sigma_1^2}\mathbf{I}_3 & \cdots & \mathbf{R}_1\mathbf{R}_n \\ \vdots & \ddots & \vdots \\ \mathbf{R}_n\mathbf{R}_1 & \cdots & \frac{1}{\sigma_n^2}\mathbf{I}_3 \end{bmatrix}. \end{aligned}$$

Thus, the ith element of the diagonal of $\mathbf{Z}$ can be expressed as

$$z_i = \frac{S_{ii}\sum_{j=1}^{n} \frac{1}{\sigma_j^2} - 1}{S_{ii}\sum_{j=1}^{n} \frac{1}{\sigma_j^2}} \tag{7}$$

where S_{ii} is the ith elements of the diagonal of the matrix $\mathbf{S}$. Moreover, it can be shown that $\sum_{i=1}^{3n} z_i = 3(n-1)$ with z_i defined in (7). Therefore the property defined in (3) holds.

The derived relation (7) also shows that the standard deviation of z_i is independent from the angles λ_i and θ_i ($\forall i = 1,\ldots,n$). This means that the relative orientation of the sensors is irrelevant to the system's optimality. The same result is also obtained using the information filter approach (for proof see [2]). Moreover, (7) implies that when sensors of different precisions are employed, measurements are not controlled in a similar manner and thus the system is not optimal (since we defined such system as one where all measurements are equally controlled). Nevertheless, if sensors of equal precision are only considered then $z_i = (n-1)/n$ ($\forall i = 1,\ldots,n$) and in this case the standard deviation of z_i is equal to 0.

D. Impact of Sensor Failures

In the previous section it was demonstrated that the orientation of sensor triads is irrelevant to the system's optimality. Now we investigate whether the failure of one of the sensors could alter this conclusion and therefore render the relative orientation of triads important. Consider the system presented in Fig. 1. Before the sensor marked with the symbol ⊘ fails, the information content of the two systems is strictly equivalent. Indeed, in the first (left) system the measurements along each axis are controlled by another one. This implies that all z_i have the same value ($z_{1,i} = \frac{1}{2}$, $\forall i = 1,\ldots,6$) and, therefore, the standard deviation of the vector $\mathbf{z}_1$ (i.e., the diagonal of $\mathbf{Z}_1$) is equal to 0. Consequently, the first (left) system has an optimal configuration.

For the second (right) system, (5) becomes (assuming for simplicity $\mathbf{S} = \mathbf{I}_5$):

$$\begin{aligned} \mathbf{Z}_2 &= \mathbf{I}_6 - \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T \\ &= \mathbf{I}_6 - \mathbf{H}\left([\mathbf{I}_3 \quad \mathbf{R}^T(\lambda,\theta)]\begin{bmatrix}\mathbf{I}_3 \\ \mathbf{R}(\lambda,\theta)\end{bmatrix}\right)^{-1}\mathbf{H}^T \\ &= \mathbf{I}_6 - \frac{1}{2}\begin{bmatrix}\mathbf{I}_3 \\ \mathbf{R}(\lambda,\theta)\end{bmatrix}[\mathbf{I}_3 \quad \mathbf{R}^T(\lambda,\theta)] \\ &= \mathbf{I}_6 - \frac{1}{2}\begin{bmatrix}\mathbf{I}_3 & \mathbf{R}^T(\lambda,\theta) \\ \mathbf{R}(\lambda,\theta) & \mathbf{I}_3\end{bmatrix}. \end{aligned}$$

Hence, it implies that the standard deviation of the vector $\mathbf{z}_2$ is also equal to 0 because $z_{2,i} = \frac{1}{2}$, $\forall i = 1,\ldots,6$.

However, after the presumed sensor failure, the standard deviations of the partial redundancies vary between both systems. In the first one, the remaining working sensor placed along the x-axis becomes uncontrolled and thus has associated z_i equal to 0. The other values remain unchanged. Therefore, $\sqrt{\mathbb{E}[(\mathbf{z}_1^\star - \mathbb{E}[\mathbf{z}_1^\star])^2]} \approx 0.224$. In the second (right)

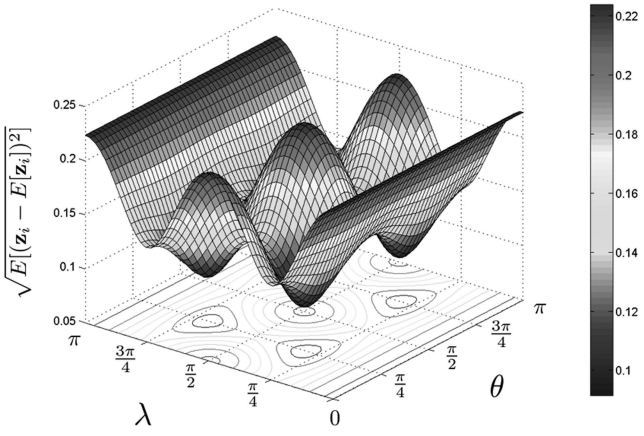

Fig. 2. Influence of λ and θ on standard deviation of $\mathbf{z}_2^\star$.

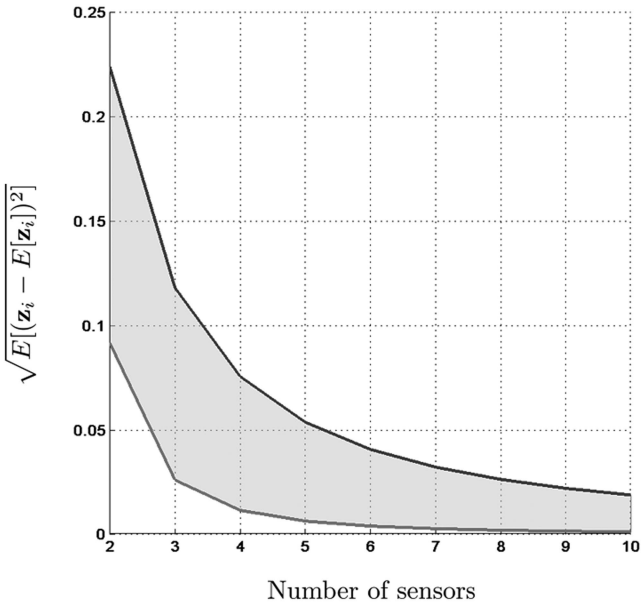

Fig. 3. The upper line corresponds to standard deviation of $\mathbf{z}_i$ in case of sensor failure for orthogonal configurations and lower line for optimal skewed configuration (with respect to sensor that fails). Area between the lines represents the definition interval of standard deviations of $\mathbf{z}_i$ in case of sensor failure for any configuration.

system, the standard deviation of $\mathbf{z}_2^\star$ can be expressed as a function of λ and θ. This relation is illustrated in Fig. 2. In consequence, it can be concluded that skew-configurations are preferable in the case of sensor failure, because the standard deviation of $\mathbf{z}_1^\star$ corresponds to the maximal of value of the standard deviation of $\mathbf{z}_2^\star$. This can be mathematically expressed as

$$\max_{\lambda,\theta}\left(\sqrt{\mathbb{E}[(\mathbf{z}_2^\star(\lambda,\theta)-\mathbb{E}[\mathbf{z}_2^\star(\lambda,\theta)])^2]}\right) = \sqrt{\mathbb{E}[(\mathbf{z}_1^\star-\mathbb{E}[\mathbf{z}_1^\star])^2]}.$$

From Fig. 2 it can be seen that one possible optimal configuration is, for example, $\hat{\lambda}=\pi/3$ and $\hat{\theta}=\pi/4$, with $\sqrt{\mathbb{E}[(\mathbf{z}_2^\star(\hat{\lambda},\hat{\theta})-\mathbb{E}[\mathbf{z}_2^\star(\hat{\lambda},\hat{\theta})])^2]}\approx 0.091$.

This example also shows the limitation of the configuration analyses by the information filter approach. In case of sensor failure this method gives identical results for both systems (i.e., $J_1^\star = J_2^\star = 4$, $\forall\lambda,\theta$). This is conceptually erroneous with respect to FDI purposes. Therefore, the method of partial redundancies is generally recommended to assess the optimality of IMU architectures although the two approaches can be considered complementary.

Finally, we study the impact of sensor failure on the optimality of orthogonal and skewed configuration as the number of redundant sensors increases. Fig. 3 shows that as the number of sensors increases the advantage of the skewed configurations decreases. Hence, for systems with large redundancy (i.e., 4 or more IMU triads), we recommend using the simplest configuration which is typically the orthogonal arrangement.

III. FAULT DETECTION AND ISOLATION WITH MEMS-IMUS

Based on the previously mentioned considerations, we present in this section two FDI approaches, namely the classical parity space method and a method based on Mahalanobis distances. Then, we experimentally investigate the performance of these methods when applied to MEMS-IMUs. Note that these methods perform statistical tests. Indeed, the fault detection can be viewed as a choice between two hypotheses concerning the absence (i.e., the null hypothesis H_0) or the presence of erroneous measurements (i.e., the alternative hypothesis H_1). Furthermore, the test statistic of the two methods all have, assuming multivariate normally distributed errors, a χ^2 distribution under H_0. The main difference between the parity space method and the approach based on the Mahalanobis distances is that the latter makes less assumptions regarding the form of the covariance matrix of the errors. Indeed, this matrix is directly estimated from the data while in the parity approach it is assumed that this matrix has the form $\sigma^2\mathbf{I}$. This method assumes the homoscedasticity and the independence of the errors. However, both approaches assume that the covariance structure is fixed in time.

A. Parity Space Method

Let $\mathbf{y}_t$ be the angular rate observation vector (or the linear acceleration vector) produced by n gyroscopes (or accelerometers) at time $t : t \in [1,\ldots,T]$. Assume for simplicity that only IMU triads are considered, therefore, $\mathbf{y}_t \in \Re^m$ where $m = 3\times n$. Let $\mathbf{x}_t \in \Re^p$ be the angular rates (or linear accelerations) to which the sensors are subjected at time t (i.e., the vector of unknown parameters we wish to estimate). Thus, we assume the following linear model between the observation vector $\mathbf{y}_t$ and the parameter vector $\mathbf{x}_t$ representing the nominal inertial signal in the body

frame:

$$\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{v}_t, \qquad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \sigma_t^2\mathbf{I})$$

where $\mathbf{H}_t \in \Re^{m\times p}$ is the system design matrix (i.e., the matrix that transforms data from the actual sensor axes to the three orthogonal axes of the predefined body frame).

Under these assumptions and if $(\mathbf{H}_t^T\mathbf{H}_t)^{-1}$ exists (i.e., if the parametrization is identifiable), then the maximum likelihood (ML) estimates of $\mathbf{x}_t$ and σ_t^2 are given by

$$\hat{\mathbf{x}}_t = (\mathbf{H}_t^T\mathbf{H}_t)^{-1}\mathbf{H}_t^T\mathbf{y}_t \tag{8}$$

$$\hat{\sigma}_t^2 = \frac{\|\mathbf{y}_t - \hat{\mathbf{y}}_t\|^2}{m} \tag{9}$$

where $\hat{\mathbf{y}}_t = \mathbf{H}_t\hat{\mathbf{x}}_t$ and $\|.\|$ represents the L^2 norm. Note that $\hat{\mathbf{y}}_t$ can also be written as

$$\hat{\mathbf{y}}_t = \mathbf{H}_t(\mathbf{H}_t^T\mathbf{H}_t)^{-1}\mathbf{H}_t^T\mathbf{y}_t = \mathbf{R}_t\mathbf{y}_t.$$

The matrix $\mathbf{R}_t$ is often call the "hat" matrix because it maps $\mathbf{y}_t$ into $\hat{\mathbf{y}}_t$. Note that the matrix $\mathbf{R}_t$ is symmetric and idempotent. Therefore, the residuals $\hat{\mathbf{v}}_t = \mathbf{y}_t - \hat{\mathbf{y}}_t$ are distributed as follows:

$$\hat{\mathbf{v}}_t \sim \mathcal{N}(\mathbf{0}, \sigma_t^2(\mathbf{I} - \mathbf{R}_t))$$

since

$$\mathbb{E}[\hat{\mathbf{v}}_t] = \mathbb{E}[\mathbf{v}_t] = \mathbf{0}$$

$$\mathrm{Var}[\hat{\mathbf{v}}_t] = (\mathbf{I} - \mathbf{R}_t)^T\mathrm{Var}[\mathbf{y}_t](\mathbf{I} - \mathbf{R}_t) = \sigma_t^2(\mathbf{I} - \mathbf{R}_t).$$

Let the vectors $\mathbf{y}$, $\mathbf{x}$, $\mathbf{v}$ and the matrix $\mathbf{H}$ be defined as

$$\mathbf{y} = \begin{bmatrix}\mathbf{y}_1\\ \vdots \\ \mathbf{y}_T\end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix}\mathbf{x}_1\\ \vdots \\ \mathbf{x}_T\end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix}\mathbf{v}_1\\ \vdots \\ \mathbf{v}_T\end{bmatrix}, \qquad \mathbf{H} = \begin{bmatrix}\mathbf{H}_1\\ \vdots \\ \mathbf{H}_T\end{bmatrix}.$$

The parity space method considers the statistic D_t which is defined as

$$D_t = \mathbf{v}_t^T\mathbf{v}_t. \tag{10}$$

A natural estimator of D_t is simply:

$$\hat{D}_t = \hat{\mathbf{v}}_t^T\hat{\mathbf{v}}_t. \tag{11}$$

Furthermore, it can be shown that

$$\frac{\hat{D}_t}{\sigma_t^2} \sim \chi^2(m-p).$$

Indeed, $\hat{D}_t$ can be written as

$$\hat{D}_t = \|(\mathbf{I} - \mathbf{R}_t)\mathbf{y}_t\|^2 = \|(\mathbf{I} - \mathbf{R}_t)\mathbf{v}_t\|^2 = \mathbf{v}_t^T(\mathbf{I} - \mathbf{R}_t)\mathbf{v}_t.$$

Let $\mathbf{z}_t = \mathbf{v}_t/\sigma_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, then,

$$\frac{\hat{D}_t}{\sigma^2} = \mathbf{z}_t^T(\mathbf{I} - \mathbf{R}_t)\mathbf{z}_t \sim \chi^2(m-p)$$

since the matrix $\mathbf{I} - \mathbf{R}_t$ is symmetric, idempotent, and has a rank of $m - p$ (see for example [34]).

However, in practice σ_t^2 is unknown and needs to be estimated from the observations. In the parity approach, it is assumed that the σ_t^2 remains constant through time, therefore, σ_t^2 is replaced by σ^2. A simple estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{\|\mathbf{y} - \hat{\mathbf{y}}\|^2}{mT}. \tag{12}$$

Note that when σ_t^2 is replaced by its estimate $\hat{\sigma}^2$, then the distribution of $\hat{D}_t/\hat{\sigma}^2$ holds only approximately. Let $\hat{\kappa}_t = \hat{D}_t/\hat{\sigma}^2$ be the (estimated) parity space test statistic, then:

$$\hat{\kappa}_t = \frac{\hat{D}_t}{\hat{\sigma}^2} \dot{\sim} \chi^2(m-p). \tag{13}$$

The isolation algorithm that is classically used in the parity space approach is based on ML as presented in [18]. In this algorithm the erroneous measurement is assumed to happen in the ηth sensor at time t where:

$$\eta = \arg\max_{i\in[1;m]} \frac{\hat{v}_i^2}{z_{t;ii}^p} \tag{14}$$

and where $z_{t;ii}^p$ corresponds to the diagonal of the matrix $\mathbf{Z}_t^p$ defined as

$$\mathbf{Z}_t^p = \mathbf{I}_m - \mathbf{H}_t(\mathbf{H}_t^T\mathbf{H}_t)^{-1}\mathbf{H}_t^T. \tag{15}$$

Therefore, the parity space methodology for FDI can be written in an algorithmic way as presented in Algorithm 1.

ALGORITHM 1 *FDI Based on Parity Approach*

$\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}$ *(equation 8)*
$\hat{\sigma}^2 = \frac{\|\mathbf{y} - \hat{\mathbf{y}}\|^2}{mT}$ *(equation 12)*
for $t \leftarrow 1$ **to** T **do**
  $\hat{\kappa}_t = \frac{\hat{D}_t}{\hat{\sigma}^2}$ *(equation 13)*
  Fault detection
  if $\hat{\kappa}_t > \chi^2_{1-\alpha}(m-p)$ **then**
    Fault isolation
    $\mathbf{Z}_t^p = \mathbf{I}_m - \mathbf{H}_t(\mathbf{H}_t^T\mathbf{H}_t)^{-1}\mathbf{H}_t^T$ *(equation 15)*
    Error in sensor ηth at t where $\eta = \arg\max_{i\in[1;m]} \frac{\hat{v}_i^2}{z_{t;ii}^p}$ *(equation 14)*
  end
end

B. Mahalanobis Distance

To extend the parity space method, we propose to apply a method that is based on the Mahalanobis distance and allows errors to be correlated and of different variances. Therefore, we wish to consider

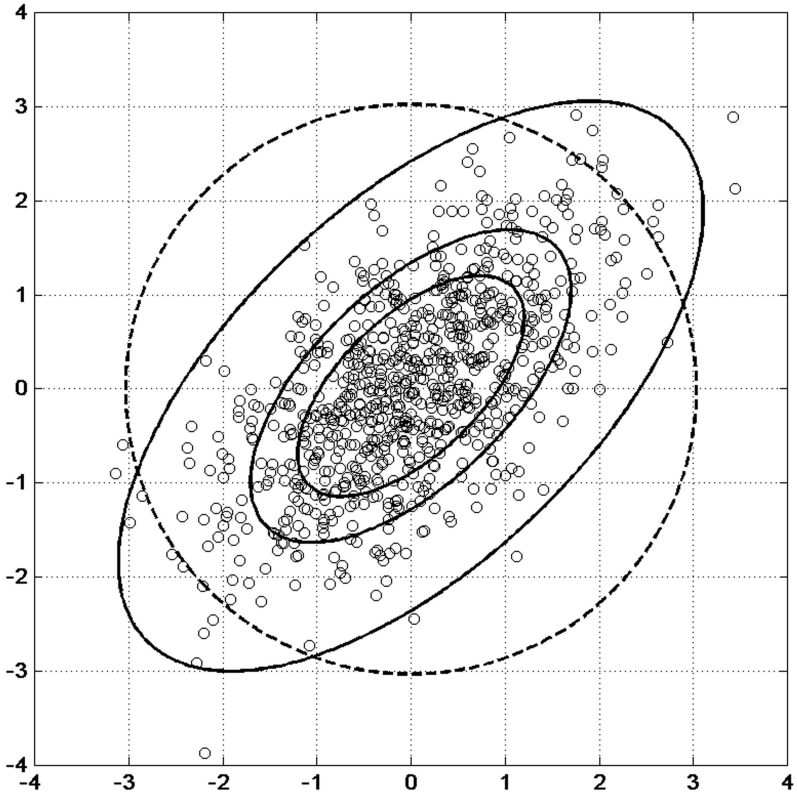

Fig. 4. Simulated standard bivariate normally distributed data with correlation of 0.60. Ellipses represent 0.50, 0.75, and 0.99 quantiles of $\chi^2(2)$ distribution. Dashed line corresponds 0.99 quantile of $\chi^2(2)$ distribution under assumption of uncorrelated data (as done in parity space method).

the following (extended) linear model:

$$\begin{aligned}
&\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{v}_t \\
&\mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}) \\
&\boldsymbol{\Omega} = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1m} \\ \vdots & \ddots & \vdots \\ \sigma_{1m} & \cdots & \sigma_m^2 \end{bmatrix}
\end{aligned}$$

where the variance-covariance matrix $\boldsymbol{\Omega}$ is unknown and assumed the same at each sampled time. Let $\boldsymbol{\lambda}$ be the vector that parametrized the matrix $\boldsymbol{\Omega}$, i.e.,

$$\boldsymbol{\lambda} = [\sigma_1^2 \cdots \sigma_m^2 \quad \sigma_{12} \cdots \sigma_{(m-1)m}]^T.$$

For a moment we assume that the vector $\mathbf{x}_t$ $(\forall t)$ and the variance-covariance matrix $\boldsymbol{\Omega}$ are known and that $\boldsymbol{\Omega}$ is positive definite, then, the Mahalanobis transformation is defined as

$$\mathbf{w}_t = \boldsymbol{\Omega}^{-1/2}\mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_m). \tag{16}$$

Equation (16) shows that the Mahalanobis transformation eliminates the correlation between the variables and standardizes the variance of each variable [35, 36]. By definition, it is therefore clear that $\mathbf{w}_i^T\mathbf{w}_i$ follows a $\chi^2(m)$ distribution. The (squared) Mahalanobis distances is related to a χ^2 distribution and is defined as

$$MD_t = \sqrt{\mathbf{v}_t^T\boldsymbol{\Omega}^{-1}\mathbf{v}_t^T}. \tag{17}$$

By setting the (squared) Mahalanobis distances equal to a certain constant, i.e., to a certain quantile of a $\chi^2(m)$ distribution, it is possible to define (hyper) ellipsoids with all the points having the same Mahalanobis distances from the centroid. Fig. 4 illustrates this concept with simulated bivariate normally distributed data. The ellipses represent the quantiles 0.50, 0.75, and 0.99 of a $\chi^2(2)$ distribution and, thus, corresponds to all the points having the same value of MD_t^2. Hence, this approach takes into account the shape of the data cloud. On the contrary, the parity space method assumes uncorrelated errors. In such case, all the points having the same D_t (as defined in (10)) values can be represented by circles (Fig. 4). Therefore, the parity space approach can produce unreliable results when residuals are correlated (or heteroscedastic).

When the vectors $\mathbf{x}_t$ and $\boldsymbol{\lambda}$ are unknown, which is the case in practice, their estimation can (most of the time) be realized by iteratively reweighted least squares (IRLS). A possible approach to realize this estimation is presented in Algorithm 2 (see also [37] for more details).

ALGORITHM 2 *Estimation of Extended Linear Models without Random Effects by IRLS*

$\hat{\mathbf{x}}^{(0)} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}$

$\hat{\boldsymbol{\Omega}}^{(0)} = \hat{\sigma}^2\mathbf{I}$ *where* $\hat{\sigma}^2 = \dfrac{\|\mathbf{y}-\hat{\mathbf{y}}\|^2}{mT}$

$\hat{\boldsymbol{\lambda}}^{(0)} = \arg\max_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} \ell(\boldsymbol{\lambda}\mid\mathbf{y})$ *where* $\ell(\boldsymbol{\lambda}\mid\mathbf{y})$

$$\propto m\log\|\mathbf{y}-\mathbf{H}\hat{\mathbf{x}}^{(0)}\| - \frac{\log|\boldsymbol{\Omega}^{(0)}|}{2}$$

$k \leftarrow 1$

repeat

$\mathbf{y}^* = (\hat{\boldsymbol{\Omega}}^{(k-1)})^{-T/2}\mathbf{y}$

$\mathbf{H}^* = (\hat{\boldsymbol{\Omega}}^{(k-1)})^{-T/2}\mathbf{H}$

$\mathbf{v}^* = (\hat{\boldsymbol{\Omega}}^{(k-1)})^{-T/2}\mathbf{v}$

$\hat{\mathbf{x}}^{(k)} = (\mathbf{H}^{*T}\mathbf{H}^*)^{-1}\mathbf{H}^{*T}\mathbf{y}^*$

$\hat{\boldsymbol{\lambda}}^{(k)} = \arg\max_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} \ell(\boldsymbol{\lambda}\mid\mathbf{y})$ *where* $\ell(\boldsymbol{\lambda}\mid\mathbf{y})$

$$\propto m\log\|\mathbf{y}-\mathbf{H}\hat{\mathbf{x}}^{(k)}\| - \frac{\log|\boldsymbol{\Omega}^{(k-1)}|}{2}$$

$k \leftarrow k+1$

until $\|\hat{\mathbf{x}}^{(k+1)} - \hat{\mathbf{x}}^{(k)}\| < \epsilon$

However, the estimation of $\mathbf{x}$ and $\boldsymbol{\lambda}$ is computationally very intensive for the large dataset we may encounter when dealing with inertial sensors. Indeed, in practice a trajectory of a few minutes duration yields a time vector of length $T = 10^4$ (assuming a sampling rate of 100 [Hz]). Standard statistical softwares which can be used to estimate such model such as the `R` function `gls()` of the well-known package `nlme` [37] requires a few minutes[1] with $T = 10^2$ and the estimation fails when $T > 10^3$. Therefore, we propose a simpler and considerably faster approach.[2] The vector $\hat{\mathbf{x}}$ is first obtained as in the parity space approach using (8), then, the elements of the matrix $\boldsymbol{\Omega}$ are estimated from $\hat{\mathbf{v}}$. Thus, our approach roughly corresponds to a "one step" IRLS estimation (see Algorithm 2).

As shown in (17) the calculation of the Mahalanobis distances require inversing the (estimated) covariance matrix of the residuals, i.e., $\hat{\boldsymbol{\Omega}}$. This operation is only possible if such matrix is full ranked. Since the computation of the residuals requires simultaneous estimation of 6 parameters (3 for the gyroscopes and 3 for the accelerometers), which can be understood as the measurement of a synthetic IMU, the residual covariance matrix is not full ranked. Practically, it implies that instead of all available residuals, only $m-3$ are used in the computation of the Mahalanobis distances. Note that the choice of the residuals used is unimportant since all the information is contained in any of the $m-3$ residuals. Consequently, the covariance matrix of $m-3$ of the residual $\boldsymbol{\Omega}^\star$ is positive definite. Moreover, as the estimate $\hat{\mathbf{x}}$ is computed without taking into account $\boldsymbol{\Omega}$, it will most likely be biased. This implies that $\hat{\mathbf{v}}$ (and $\hat{\mathbf{v}}^\star$) will not be centered in $\mathbf{0}$. Therefore, we propose employing the following estimator of MD_t:

$$\widehat{MD}_t^{\mathrm{ML}} = \sqrt{(\hat{\mathbf{v}}_t^\star - \bar{\mathbf{v}}_{\mathrm{ML}}^\star)^T(\hat{\boldsymbol{\Omega}}_{\mathrm{ML}}^\star)^{-1}(\hat{\mathbf{v}}_t^\star - \bar{\mathbf{v}}_{\mathrm{ML}}^\star)} \quad (18)$$

where $\bar{\mathbf{v}}_{\mathrm{ML}}^\star$ and $\hat{\boldsymbol{\Omega}}_{\mathrm{ML}}^\star$ are the ML estimators of the mean vector and of the covariance matrix, i.e.,

$$\bar{\mathbf{v}}_{\mathrm{ML}}^\star = \frac{1}{T}\sum_{t=1}^{T}\hat{\mathbf{v}}_t^\star \quad (19)$$

$$\hat{\boldsymbol{\Omega}}_{\mathrm{ML}}^\star = \frac{1}{T}\sum_{t=1}^{T}(\hat{\mathbf{v}}_t^\star - \bar{\mathbf{v}}_{\mathrm{ML}}^\star)(\hat{\mathbf{v}}_t^\star - \bar{\mathbf{v}}_{\mathrm{ML}}^\star)^T. \quad (20)$$

Let $\hat{\phi}_t^{\mathrm{ML}} = (\widehat{MD}_t^{\mathrm{ML}})^2$ be our test statistic, then, as for $\hat{\kappa}_t$ in (13), $\hat{\phi}_t^{\mathrm{ML}}$ follows (approximately) a χ^2 i.e.,

$$\hat{\phi}_t^{\mathrm{ML}} = (\widehat{MD}_t^{\mathrm{ML}})^2 \dot{\sim} \chi^2(m-p). \quad (21)$$

To isolate fault we propose a slightly modified version of the isolation algorithm defined by (14) and (15). This algorithm considers, like the parity space approach, that covariance matrix of the residuals is an identity matrix. Thus, we incorporated the estimated covariance matrix into the original formula. This modifies the definition of $\mathbf{Z}_t^p$ which becomes

$$\mathbf{Z}_t^{\mathrm{ML}} = \mathbf{I}_{m-3} - \mathbf{H}_t^\star(\mathbf{H}_t^{\star T}(\hat{\boldsymbol{\Omega}}_{\mathrm{ML}}^\star)^{-1}\mathbf{H}_t^\star)^{-1}\mathbf{H}_t^{\star T}(\hat{\boldsymbol{\Omega}}_{\mathrm{ML}}^\star)^{-1} \quad (22)$$

where $\mathbf{H}_t^\star$ is the design matrix associated with the chosen residuals $\hat{\mathbf{v}}_t^\star$. Note that this matrix needs to be computed twice with different $\hat{\mathbf{v}}_t^\star$ to obtain each of the diagonal elements $z_{t;ii}^{\mathrm{ML}}$. Then, we assume the fault occurs in the δth sensor at time t where δ is defined similarly to (14) as

$$\delta = \arg\max_{i\in[1;m]} \frac{\hat{v}_i^2}{z_{t;ii}^{\mathrm{ML}}}. \quad (23)$$

Our approach to FDI based on Mahalanobis distances can be summarized as in Algorithm 3.

ALGORITHM 3 *FDI Based on Mahalanobis Distances*

$\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}$ *(equation 8)*

$\bar{\mathbf{v}}_{\mathrm{ML}}^\star = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T}\hat{\mathbf{v}}_t^\star$ *(equation 19)*

$\hat{\boldsymbol{\Omega}}_{\mathrm{ML}}^\star = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T}(\hat{\mathbf{v}}_t^\star - \bar{\mathbf{v}}_{\mathrm{ML}}^\star)(\hat{\mathbf{v}}_t^\star - \bar{\mathbf{v}}_{\mathrm{ML}}^\star)^T$ *(equation 20)*

for $t \leftarrow 1$ **to** T **do**

$\hat{\phi}_t^{\mathrm{ML}} = (\mathbf{v}_t^\star - \bar{\mathbf{v}}_{\mathrm{ML}}^\star)^T(\hat{\boldsymbol{\Omega}}_{\mathrm{ML}}^\star)^{-1}(\mathbf{v}_t^\star - \bar{\mathbf{v}}_{\mathrm{ML}}^\star) \dot{\sim} \chi^2_{m-p}$ *(equation 21)*

Fault detection

if $\hat{\phi}_t^{\mathrm{ML}} > \chi^2_{1-\alpha}(m-p)$ **then**

Fault isolation

$\mathbf{Z}_t^{\mathrm{ML}} = \mathbf{I}_{m-3} - \mathbf{H}_t^\star(\mathbf{H}_t^{\star T}(\hat{\boldsymbol{\Omega}}_{\mathrm{ML}}^\star)^{-1}\mathbf{H}_t^\star)^{-1} \times \mathbf{H}_t^{\star T}(\hat{\boldsymbol{\Omega}}_{\mathrm{ML}}^\star)^{-1}$ *(equation 22)*

[1] 10 min and 32 s with a 2.93 GHz Quad-Core Intel Core i7 processor (with 8 MB level 3).

[2] Around 10 sec are needed for the estimation with $T = 7000$ on the same computer.

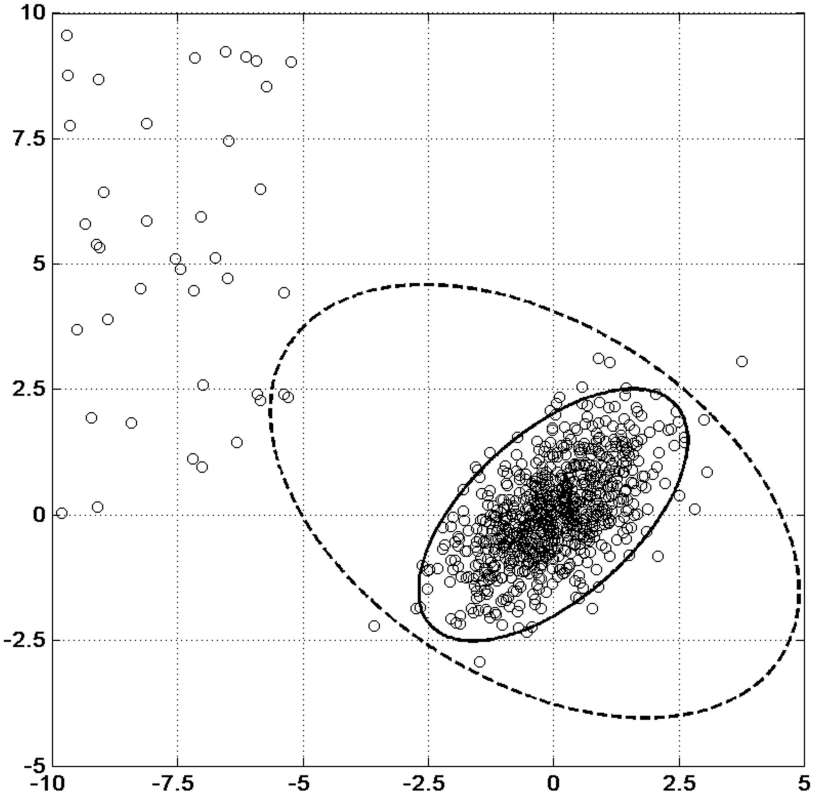

Fig. 5. Simulated standard bivariate normally distributed data with correlation of 0.60 and with 5% of perturbated data (outliers). Nonrobust estimation of covariance matrix is represented by dashed ellipse (0.975 quantile of $\chi^2(2)$) with associated correlation coefficient of -0.416. Robust estimation produces solid ellipse (0.975 quantile of $\chi^2(2)$) and estimates correlation of 0.597.

Error in sensor δth at t where

$\delta = \arg\max_{i\in[1;m]} \frac{\hat{\nu}_i^2}{z_{t;ii}^{\mathrm{ML}}}$ *(equation 23)*

end

end

IV. ROBUST APPROACH TO FDI

The FDI process corresponds to identifying the residuals lying outside of the circle (parity space) or the ellipse (Mahalanobis distance) for a predefined quantile of a $\chi^2(m-p)$ distribution.

However, the Mahalanobis distances (like D_t in (11)) are very sensitive to the presence of outliers [27]. A few extreme observations departing from the main data structure can have a severe influence on this distance measure (Fig. 5) due to the nonrobust estimation of the covariance matrix $\hat{\Omega}^\star$. Consequently, the Mahalanobis distances are heavily affected by the outliers it aims to detect. This is called the "masking effect," by which multiple outliers do not necessarily have large Mahalanobis distances [26, 27]. Hence, it is more appropriate to use distances based on robust rather than on ML estimators (as defined in (19) and (20)) for the multivariate location and scatter if the presence of multiple outliers is suspected (which is the case in the context of FDI). In practice, the minimum covariance determinant (MCD) estimator [38] is often employed although many other robust estimators have been introduced in the literature (for a review see [39]).

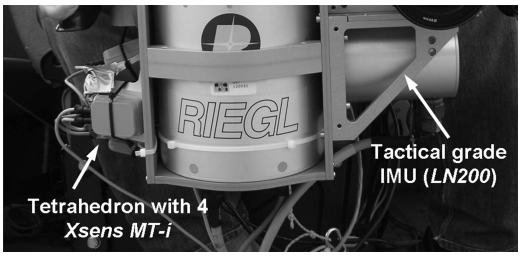

Fig. 6. Skew-redundant MEMS-IMUs placed in tetrahedron mounted together with tactical grade IMU (*LN200*).

The MCD estimator has the objective to find h observations (out of m) whose covariance matrix has the lowest determinant (i.e., the smallest (hyper) volume of the ellipses associated to the covariance matrix). The multivariate location and scatter are then estimated in a "standard" manner on these h points. It is generally recommended to use $h \approx 0.75$ m, as a compromise between robustness and efficiency [23, 40]. This will provide accurate results if the data set contains at most 25% of aberrant values, which is a very reasonable assumption in the case of inertial sensors. Therefore, the robust estimators of the mean

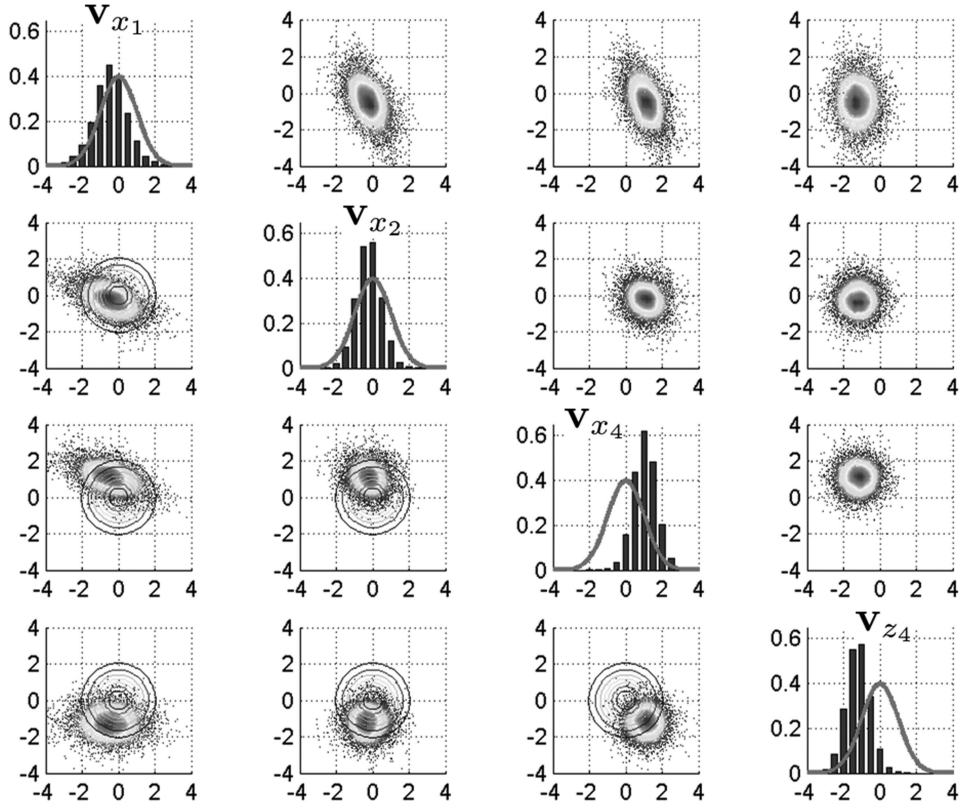

Fig. 7. Partial representation of covariance of residuals, diagonal elements compares histograms of residuals with distribution assumed in parity space method. Upper elements show scatter plots of residuals and lower elements compare scatter plots with assumed bivariate distribution.

vector and of the covariance matrix are defined as

$$\bar{\mathbf{v}}^{\bullet}_{\mathrm{MCD}} = \frac{1}{h}\sum_{t=1}^{h}\hat{\mathbf{v}}^{\bullet}_{t} \tag{24}$$

$$\hat{\Omega}^{\bullet}_{\mathrm{MCD}} = \frac{1}{h}\sum_{t=1}^{h}(\hat{\mathbf{v}}^{\bullet}_{t} - \bar{\mathbf{v}}^{\bullet})(\hat{\mathbf{v}}^{\bullet}_{t} - \bar{\mathbf{v}}^{\bullet})^{T} \tag{25}$$

where $\hat{\mathbf{v}}^{\bullet}_{t}$ corresponds to the h observations among $\hat{\mathbf{v}}^{\star}_{t}$ selected by the MCD algorithm. Based on these robust estimates, we propose the following robust estimator of MD_t:

$$\widehat{MD}_t^{\mathrm{MCD}} = \sqrt{(\hat{\mathbf{v}}^{\star}_{t} - \bar{\mathbf{v}}^{\bullet}_{\mathrm{MCD}})^{T}(\hat{\Omega}^{\bullet}_{\mathrm{MCD}})^{-1}(\hat{\mathbf{v}}^{\star}_{t} - \bar{\mathbf{v}}^{\bullet}_{\mathrm{MCD}})}. \tag{26}$$

When the MCD estimates are used to compute Mahalanobis distances (as defined in (26)), it leads to the so-called robust distances. Hence, our test statistic becomes

$$\hat{\phi}_t^{\mathrm{MCD}} = (\widehat{MD}_t^{\mathrm{MCD}})^2 \dot{\sim} \chi^2(m-p). \tag{27}$$

The isolation algorithm can also be slightly modified by using the MCD instead of the ML estimates. Therefore, the robust version of $\mathbf{Z}_t^{\mathrm{ML}}$ is simply:

$$\mathbf{Z}_t^{\mathrm{MCD}} = \mathbf{I}_{m-3} - \mathbf{H}_t^{\star}(\mathbf{H}_t^{\star T}(\hat{\Omega}^{\bullet}_{\mathrm{MCD}})^{-1}\mathbf{H}_t^{\star})^{-1}\mathbf{H}_t^{\star T}(\hat{\Omega}^{\bullet}_{\mathrm{MCD}})^{-1}. \tag{28}$$

The fault is assumed to occur in the ξth sensor at time t where ξ is defined as

$$\xi = \arg\max_{i\in[1;m]} \frac{\hat{\nu}_i^2}{z_{t;ii}^{\mathrm{MCD}}}. \tag{29}$$

The robust version of the Mahalanobis approach to FDI is identical to Algorithm 3 when replacing $\bar{\mathbf{v}}^{\star}_{\mathrm{ML}}$, $\hat{\Omega}^{\star}_{\mathrm{ML}}$, ϕ_t^{ML}, $\mathbf{Z}_t^{\mathrm{ML}}$ and δ by $\bar{\mathbf{v}}^{\bullet}_{\mathrm{MCD}}$, $\hat{\Omega}^{\bullet}_{\mathrm{MCD}}$, ϕ_t^{MCD}, $\mathbf{Z}_t^{\mathrm{MCD}}$, and ξ, respectively.

V. EXPERIMENTAL RESULTS

The results presented here are based on an experiment realized in a vehicle (Fig. 6). A regular tetrahedron consisting of 4 *Xsens MT-i* MEMS-IMUs was mounted on a rigid structure together with an IMU of reference (*LN200*). We decided to utilize this skew-redundant configuration which optimizes the controllability of measurements in case of sensor failure.

After computing the least squares residuals $\hat{\mathbf{v}}_t$ for the gyros (as defined in Section IIIA), we explore the structure of the associated covariance matrix. Fig. 7

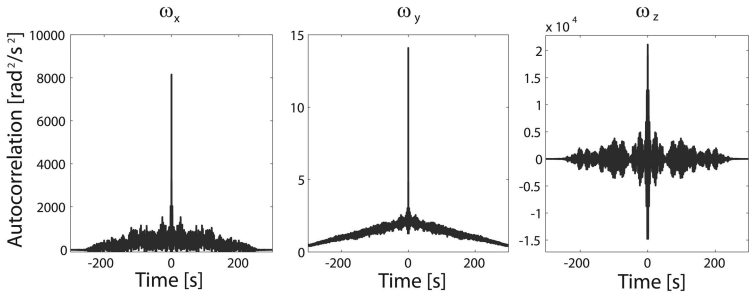

Fig. 8. Autocorrelation of MEMS-gyros residuals.

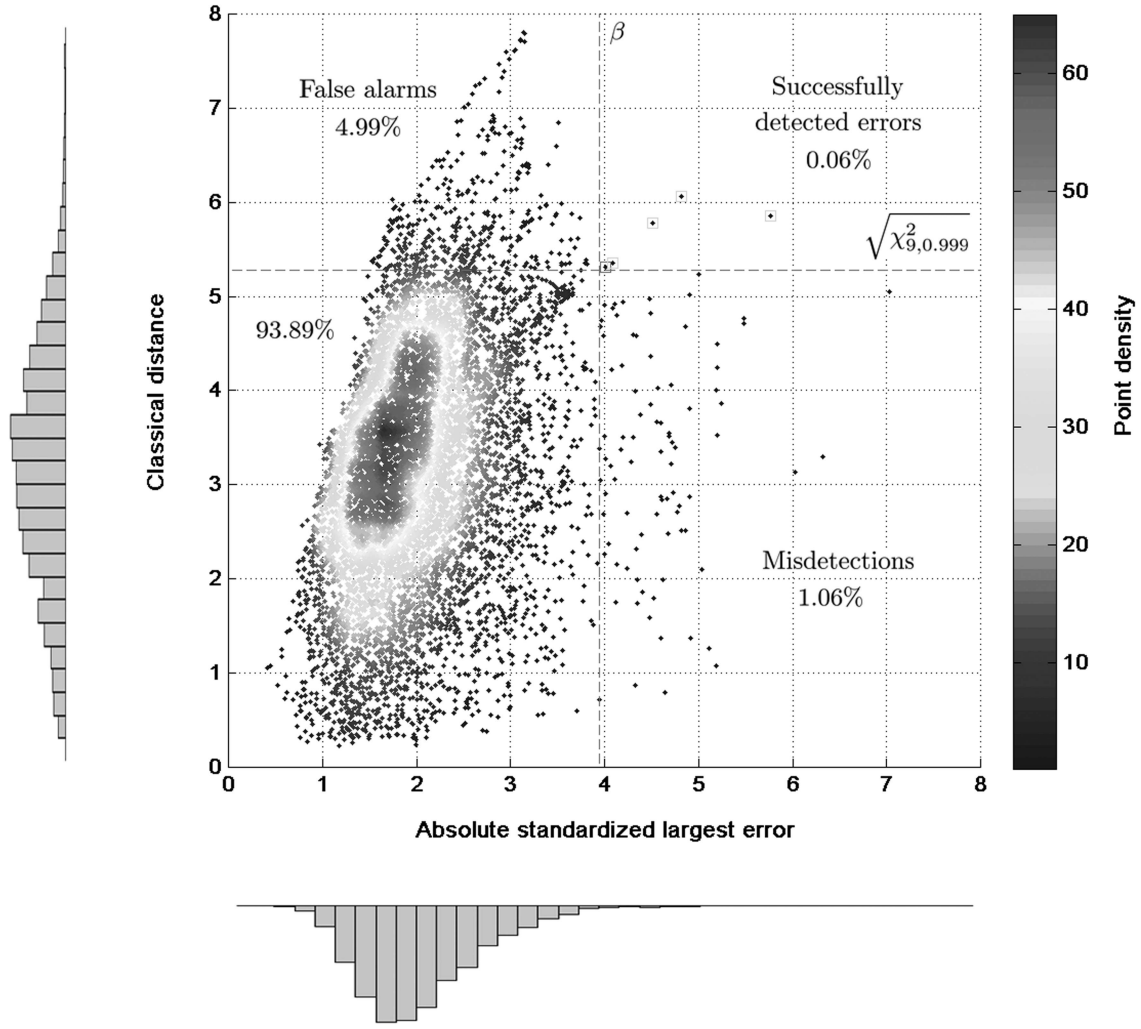

Fig. 9. Standford plot of parity space results (applied to MEMS gyros), squares indicate performance of isolation algorithm: success (light gray), error (dark gray).

shows a partial representation of this matrix (which is from $\Re^{12\times 12}$). It illustrates that the assumptions of the parity space method are not satisfied since the residuals are correlated (e.g. $\text{corr}(V_{x_1}, V_{x_2}) \approx -0.60$), have different variances (e.g. $\sigma^2_{V_{x_1}} \approx 2\sigma^2_{V_{x_4}}$) and are not centered in zero (e.g. V_{x_4}). Additionally, the autocorrelation plots of the MEMS gyros residuals reveal that a large proportion (of the nonfaulty noise observations) is composed of colored noise (Fig. 8). Indeed, the autocorrelations suggest that the errors are mainly composed of white noise and of random walk. Hence, the assumption of strictly Gaussian white noise errors in the parity space approach is unrealistic with MEMS-IMUs.

In order to assess the performance of each technique, we first used the reference IMU to determine the "true" errors. Then, we constructed $Z = \max(|e_1| \cdots |e_{12}|)$ (i.e., absolute standardized largest residuals) where e_i are the standardized residuals of the ith sensor. Furthermore, it can reasonably

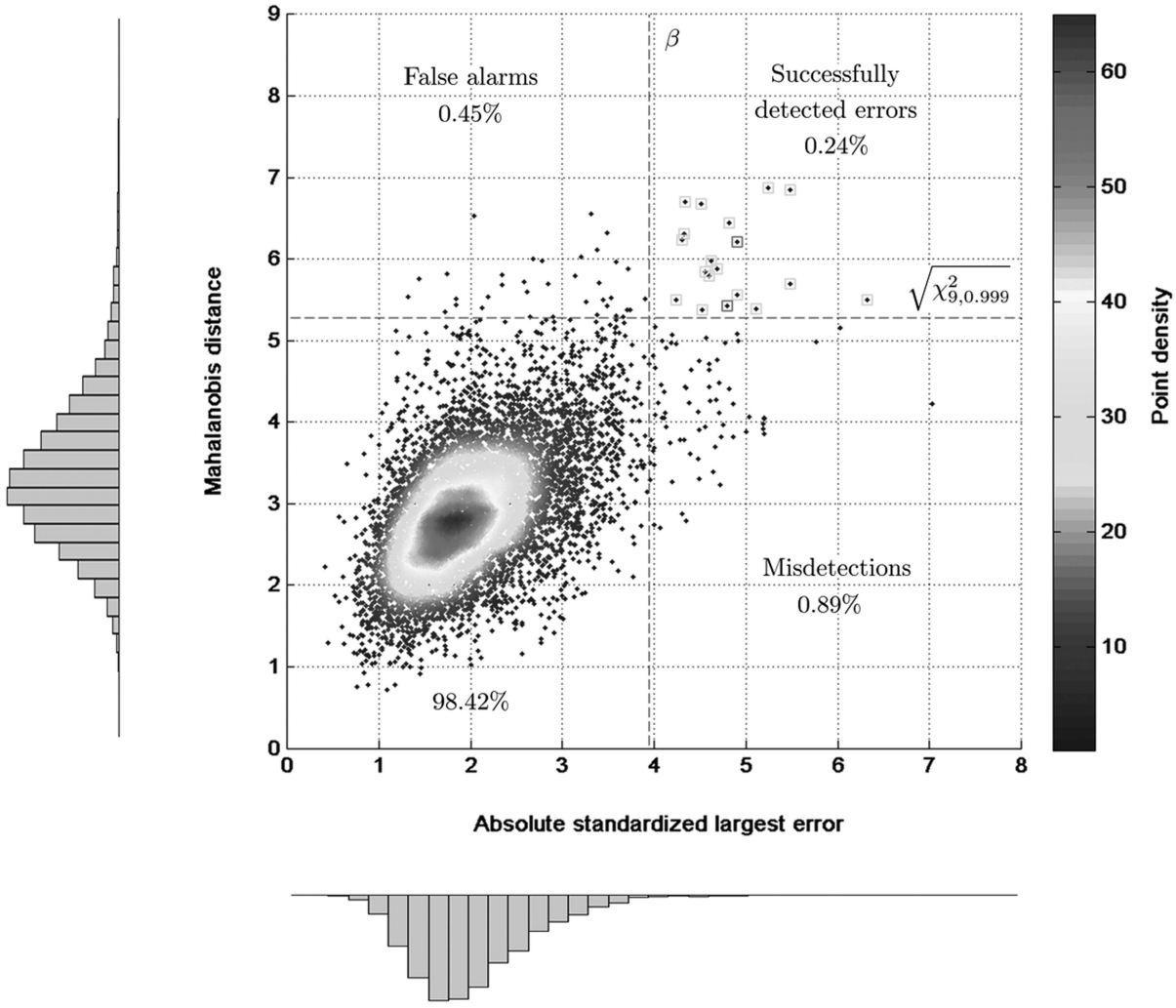

Fig. 10. Standford plot of Mahalanobis distances' based method (applied to MEMS gyros). Squares indicate performance of isolation algorithm: success (light gray), error (dark gray).

be assumed that these standardized residuals are independently and identically normally distributed (note that this assumption was experimentally verified). Then, the distribution of Z can be derived. $Y_i = |e_i|$ follows, by definition, a folded normal distribution or more precisely a half-normal distribution since $\mathbf{E}(e_i) = 0$. As the variance of e_i is standardized, it implies that the cumulative distribution function of Y_i is given by

$$F_Y(y) = \sqrt{\frac{2}{\pi}} \int_0^y \exp\left(-\frac{x^2}{2}\right) dx.$$

Further, $Z = \max(Y_i)$, thus, it has a cumulative distribution function given by

$$F_Z(z) = P[Y_1 \leq z \cap Y_2 \leq z \cap \cdots \cap Y_{12} \leq z] = (F_Y(z))^{12}.$$

Since we are not in the context of safety-critical applications, an arbitrary very small type I error value was chosen, i.e., $\alpha = 0.2\%$. In other words, a fault is considered to occur when $Z > \beta = 3.93$ since $F_z(\beta) = 1 - \alpha/2$.

The results for the parity space method are presented in Fig. 9. As can be observed the performance of this method is very poor. Indeed, approximately 95% of the outliers are not detected and the false-alarm-to-fault ratio is roughly 99%. The performance of the method is $1 - (p_{\text{FA}} + p_{\text{MD}}) = 93.95\%$.

Fig. 10 presents the results of the method based on Mahalanobis distances. Clearly, this provides better results with a performance of 98.66%. Nonetheless, the proportion of undetected error and the false-alarm-to-fault ratio remains relatively high (respectively 79% and 65%). Regarding the performance of the isolation algorithm, 91% of faults are isolated correctly. Fig. 11 allows graphically comparing the results of the two methods. It clearly shows that the Mahalanobis presents a very strong correlation with the known absolute standarized largest errors while this is not so clearly the case with the parity space method.

In the assumptions of the three FDI algorithms, all test statistics are theoretically following the same

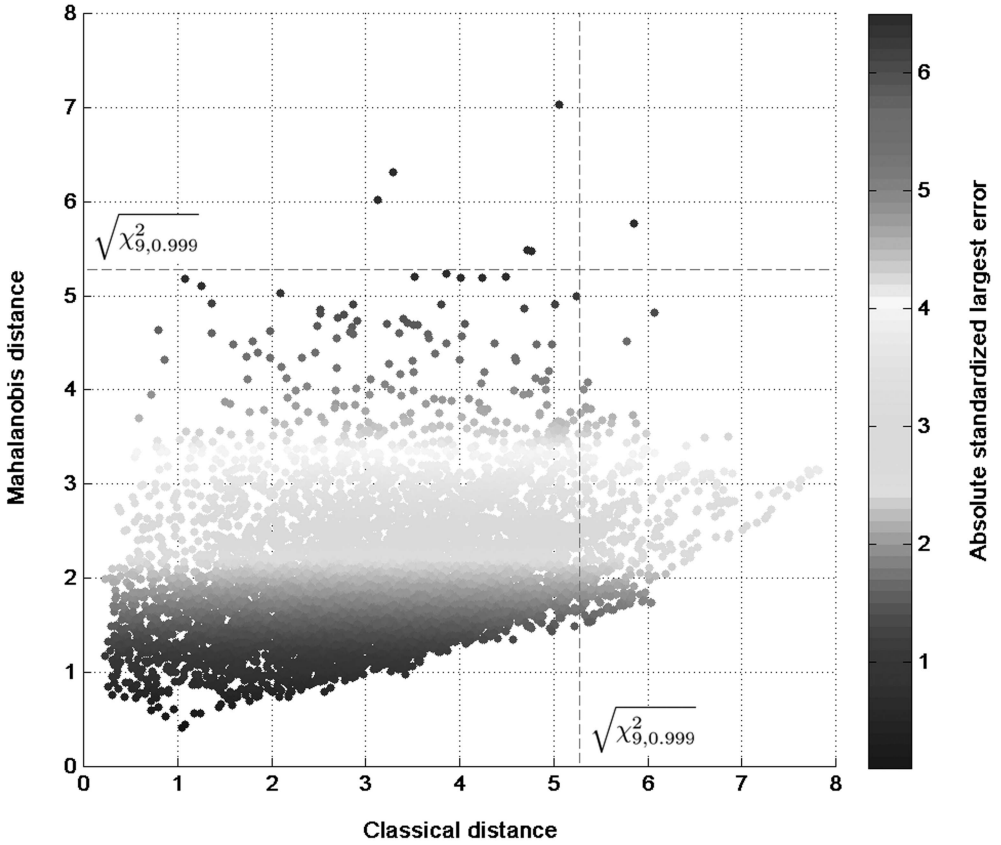

Fig. 11. Graphical comparison of performances between parity space and Mahalanobis-based method.

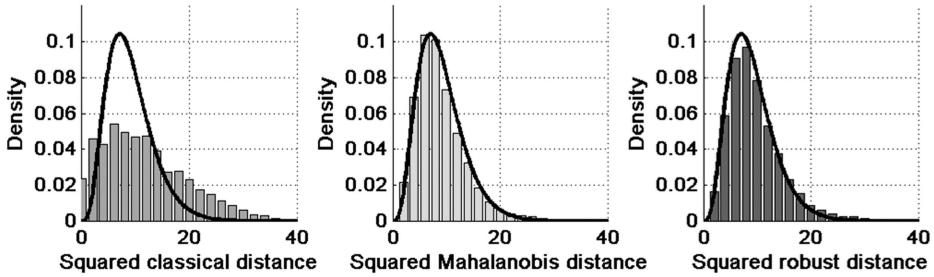

Fig. 12. Comparison of empirical and theoretical distributions of three FDI model test statistics.

$\chi^2(9)$ distribution. Fig. 12 compares the empirical and the theoretical distributions of the test statistics. It shows that the classical distances (i.e., parity space test statistics) do not follow $\chi^2(9)$ while the Mahalanobis and robust distances apparently do.

Finally, we compare the performances of the classical isolation algorithm with the approach presented in this article. The results are summarized in Table I.

The results of the robust method and the method based on Mahalanobis distances are very similar. This can be explained by the small proportion of outliers in this data set. Consequently, the ML and the robust estimators yield similar results, thus, $\hat{\Omega}^{\star}_{\mathrm{ML}} \approx \hat{\Omega}^{\bullet}_{\mathrm{MCD}}$ and $\bar{\mathbf{v}}^{\star}_{\mathrm{ML}} \approx \bar{\mathbf{v}}^{\bullet}_{\mathrm{MCD}}$. In order to compare the performances of the classical and robust estimation, we added perturbations to the original data. Indeed, 20% of the measurements of one of first IMU's gyros (on the x-axis) were randomly modified to yield to erroneous measurements. Then, we compared the performances of the method based on Mahalanobis distances and its robust version. The performance of the robust approach remained approximately

TABLE I
Performance Comparison

	FA	MD	SDE	IA	MIA
Classic	4.99%	1.06%	0.06%	80%	80%
Mahalanobis	0.45%	0.89%	0.24%	78%	91%
Robust	1.01%	0.95%	0.17%	75%	92%

Note: FA = False alarm, MD = misdetections and SDE = successfully detected errors, IA = performance of classical isolation algorithm, MIA = performance of modified isolation algorithm.

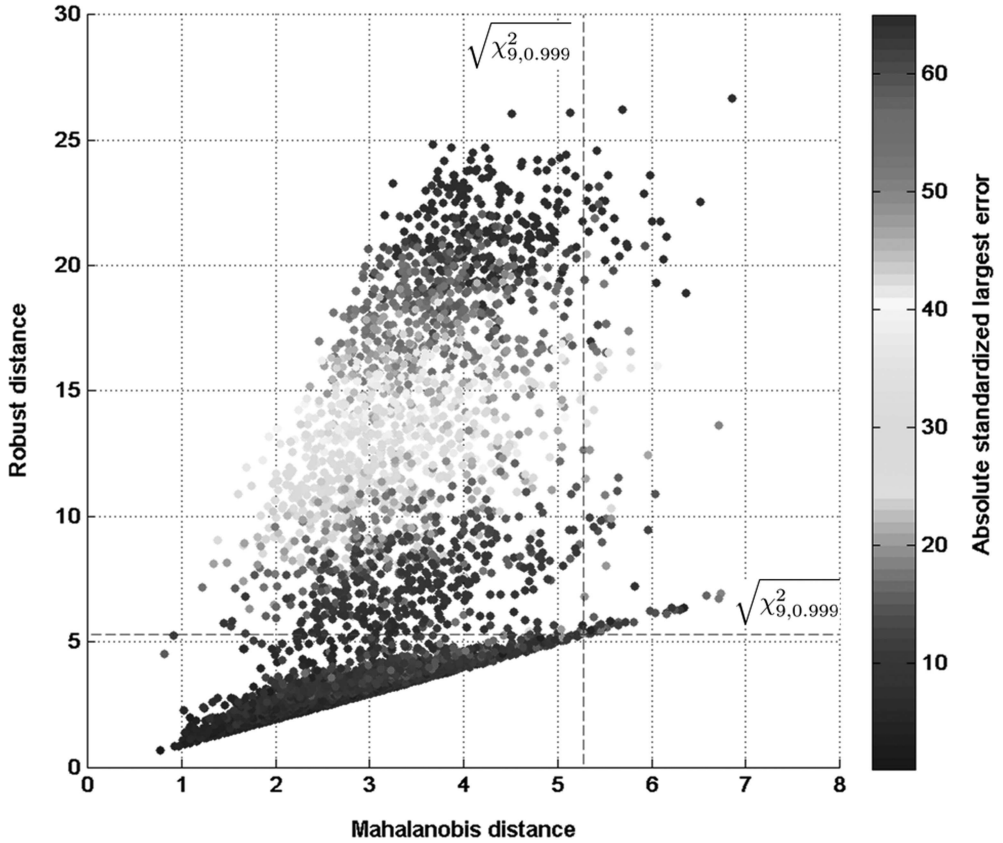

Fig. 13. Graphical comparison between methods based on Mahalanobis and robust distances with perturbed data.

TABLE II
Performance Comparison with Perturbed Data

	FA	MD	SDE
Mahalanobis	0.29%	21.17%	1.01%
Robust	2.71%	1.46%	20.72%

Note: FA = False alarm, MD = misdetections and SDE = successfully detected errors).

the same as in the previous case (95.82% instead of 96.73%). On the other hand, the method based on the Mahalanobis method provided in this case unrealistic results (Fig. 13). This illustrates the "masking effect" described in [26], [27], by which multivariate outliers do not necessarily yield to abnormally large Mahalanobis distances (or to large parity space statistic D). Table II summarizes these results.

VI. CONCLUSION

In the first part of the article, we proposed a new method based on partial redundancies to assess the optimality of redundant IMU configurations. With the help of this method (and also with the support of the information filter approach) we demonstrated that the relative sensor orientation is unimportant to the system's optimality as long as IMU triads are considered and no sensor fails. Indeed, the orientation of sensors does not change the distribution of information among axes. This conclusion has practical implications. It shows that complicated geometries based on platonic solids (which were traditionally employed in such systems) can be avoided if sensor failures are not critical. In addition, as the number of sensors increases, the advantages of the skewed configurations decreases. Hence, for systems with large redundancy (i.e., 4 or more IMU triads) we recommend the simplest configuration, which is typically the orthogonal arrangement.

In the rest of the paper, we focused on three FDI algorithms. We have shown that FDI algorithms based on Mahalanobis distances or robust distances give better results than the parity space method when applied to MEMS-IMUs. We also demonstrated that robust distances are more reliable when the number of failures increases. Hence, such FDI approaches are considered efficient and valuable for detecting gross errors, which in turn enhances the navigation performance and improves the stability of the filters used in the GPS/INS integration. Moreover, we also experimentally demonstrated that the performance of the isolation algorithm can be slightly improved by incorporating the covariance matrix into the classical method. Nevertheless, the percentages of undetected errors, as well as the level of false alarms, remains relatively high and, consequently, shows the

need for more complex FDI models. In addition, the FDI algorithms presented in this article would most likely perform poorly in the case of a slowly growing bias. However, the authors believe that this problem is better to be dealt with by other estimation approaches (e.g. GPS/INS integration employing filter banks with various stochastic assumptions [41]) because slowly growing bias can be in principal modeled and estimated by external aiding (e.g. GPS observations). On the other hand, one of the methods that could potentially cope with a slowly growing bias is the generalized autoregressive conditional heteroskedasticity (GARCH) approach presented in [42]. With this method the estimated variance of the sensor affected by a bias will grow in time, hence, the importance given to this sensor in the processing will decrease in time. In the future, we plan to investigate an alternative that employs multivariate GARCH models to estimate the variation of covariance matrix $\mathbf{\Omega}_t$ during the processing.

REFERENCES

[1] El-Sheimy, N. and Niu, X.
The promise of MEMS to the navigation community.
Inside GNSS, (Mar./Apr. 2007).

[2] Guerrier, S.
Integration of skew-redundant MEMS-IMU with GPS for improved navigation performance.
Master's thesis, Swiss Federal Institute of Technology Lausanne (EPFL), 2008. URL: http://www.hec.unige.ch/www/hec/m2/CorpsEnseignant/Assistants/GuerrierStephane/Publications.html.

[3] Waegli, A., Guerrier, S., and Skaloud, J.
Redundant MEMS-IMU integrated with GPS for performance assessment in sports.
In *Proceedings of the IEEE/ION PLANS 2008*, Monterey, CA, 2008.

[4] Gertler, J. and Singer, D.
A new structual framework for parity equation-based failure detection and isolation.
Automatica, **26**, 2 (1990), 381–388.

[5] Sukkarieh, S., et al.
A low-cost redundant inertial measurement unit for unmanned air vehicles.
The International Journal of Robotics Research, **19**, 11 (2000), 1089–1103.

[6] Krogmann, U.
Artificial neural networks for inertial sensor fault diagnosis.
In *Proceedings of the Symposium of Gyro Technology*, Stuttgart, Germany, 1995.

[7] Evans, F. and Wilcox, J.
Experimental strapdown redundant sensor inertial navigation system.
Journal of Spacecraft and Rockets, **7** (1970), 1070.

[8] Gilmore, J. and McKern, R.
A redundant strapdown inertial reference unit (SIRU).
Journal of Spacecraft and Rockets, **9**, 39 (1972).

[9] Potter, J. and Deckert, J.
Minimax failure detection and identification in redundant gyro and accelerometer systems.
Journal of Spacecraft and Rockets, **10** (1973), 236–243.

[10] Wilcox, J.
Competitive evaluation of failure detection algorithms for strapdown redundant inertial instruments.
In *Proceedings of the American Institute of Aeronautics and Astronautics, Guidance and Control Conference*, Key Biscayne, FL, 1973, 11.

[11] Willsky, A. and Jones, H.
A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems.
IEEE Transactions on Automatic Control, **AC-21**, 1 (1976), 108–112.

[12] Gai, E., Harrison, J., and Daly, K.
Generalized likelihood test for FDI in redundant sensor configurations.
Journal of Guidance and Control, **2**, 1 (1979), 9–17.

[13] Gai, E., Harrison, J., and Daly, K.
Performance of two redundant sensor configurations.
IEEE Transactions on Aerospace and Electronic Systems, **AES-15**, 3 (1979), 405–413.

[14] Hall, S., et al.
In-flight parity vector compensation for FDI.
IEEE Transactions on Aerospace and Electronic Systems, **AES-19** (1983), 668–676.

[15] Ray, A. and Desai, M.
A redundancy management procedure for fault detection and isolation.
Journal of Dynamic Systems, Measurement, and Control, **108**, 248 (1986).

[16] Sudano, J., Preisig, J., and Pokotylo, J.
Improved fault detection using a selected grouping of parity equations for advanced flight control systems.
In *Proceedings of the National IEEE Aerospace and Electronics Conference* (NAECON), 1988, 1565–1569.

[17] Sturza, M.
Skewed axis inertial sensor geometry for optimal performance.
In *Proceedings of the 8th AIAA/IEEE Digital Avionics Systems Conference*, San Jose, CA, Oct. 17–20, 1988; also Technical Papers, Part 1, A89-18051 05-06, AIAA, Washington, D.C., 1988, 128–135.

[18] Sturza, M.
Navigation system integrity monitoring using redundant measurements.
Journal of the Institute of Navigation, **35**, 4 (1988), 69–87.

[19] Guerrier, S.
Improving accuracy with multiple sensors: Study of redundant MEMS-IMU/GPS configurations.
In *Proceedings of the ION GNS 2009*, Savannah, GA, 2009.

[20] Kim, S., Kim, Y., and Park, C.
Failure diagnosis of skew-configured aircraft inertial sensors using wavelet decomposition.
IET Control Theory & Applications, **1**, 5 (2007), 1390–1397.

[21] Dong, Y. and Hongyue, Z.
Optimal design of robust analytical redundancy for a redundant strapdown inertial navigation system.
Control Engineering Practice, **4**, 12 (1996), 1747–1752.

[22] Jia, H.
Data fusion methodologies for multisensor aircraft navigation systems.
School of Engineering,
Cranfield University, Master's Thesis, 2004. Available: http://dspace.lib.cranfield.ac.uk/handle/1826/781.

[23] Filzmoser, P., Garrett, R., and Reimann, C.
Multivariate outlier detection in exploration geochemistry.
Computers & Geosciences, **31** (2005), 579–587.

[24] Gertler, J.
Failure dection and isolation in complex plants.
In *Proceedings of the IEEE Conference on Decision Control*, Athens, Greece, 1990.

[25] Goodlin, B., et al.
Simultaneous fault detection and classification for semiconductor manufacturing tools.
Journal of the Electrochemical Society, **150**, 12 (2003).

[26] Rousseeuw, P. and Van Driessen, K.
Fast algorithm for the minimum covariance determinant estimator.
Technometrics, **41**, 3 (1999).

[27] Rousseeuw, P. and Van Zomeren, B.
Unmasking multivariate outliers and leverage points.
Journal of the American Statistical Association, **85** (1990), 633–639.

[28] Beckman, R. and Cook, R.
Outlier.....s.
Technometrics, **25**, 2 (1983), 119–149.

[29] Pejsa, A.
Optimum skewed redundant inertial navigators.
AIAA Journal, **12** (1974), 899–902.

[30] Bjerhammar, A.
Theory of Errors and Generalized Matrix Inverses.
Amsterdam: Elsevier, 1973.

[31] Dupraz, H. and Stahl, M.
Ecole Polytechnique fédérale de Lausanne.
Théorie d'erreurs III, 1994.

[32] Heritier, S., et al.
Robust Methods in Biostatistics (Series in Probability and Statistics).
Hoboken, NJ: Wiley, 2009.

[33] Colomina, I., et al.
Redundant IMUs for precise trajectory determination.
In *Proceedings of the XXth ISPRS Congress*, Working Group I/5, Istanbul, Turkey, 2004.

[34] Rao, C. and Toutenburg, H.
Linear Models: Least Squares and Alternatives.
New York: Springer, 1995.

[35] Mahalanobis, P.
On the generalised distance in statistics.
In *Proceedings of the National Institute of Sciences of India*, 1936.

[36] Kanti V., Mardia, J. T. K., and Bibby, J. M.
Multivariate Analysis.
London: Academic Press, Elsevier Science, 1979.

[37] Pinheiro, J. and Bates, D.
Mixed-Effects Models in S and S-PLUS.
New York: Springer, 2009.

[38] Rousseeuw, P.
Multivariate estimation with high breakdown point.
Mathematical statistics and applications B.
Akadémiai Kiadó, (1985), 283–297.

[39] Maronna, R. and Yohai, V.
Robust estimation of multivariate location and scatter.
Encyclopedia of Statistical Sciences Update, **2** (1998), 589–596.

[40] Verboven, S. and Hubert, M.
LIBRA: A MATLAB library for robust analysis.
Chemometrics and Intelligent Laboratory Systems, **75** (2005), 127–136.

[41] Maybeck, P.
Stochastic Models, estimation and Control.
New York: Academic Press, 1982.

[42] Waegli, A., et al.
Noise reduction and estimation in multiple micro-electro-mechanical inertial systems.
Measurement Science and Technology, **21** (2010), 201.

Stéphane Guerrier obtained in 2008 an M.Sc. in geomatics engineering from the Swiss Federal Institute of Technology Lausanne (EPFL) for his work on the integration of redundant MEMS-IMUs with GPS.

He is now working as a Ph.D. student and teaching assistant in the Research Center for Statistics of the University of Geneva. His research concentrates on model selection in mixed linear models and on stochastic processes.

Adrian Waegli holds a Ph.D. in computer, communication and information sciences and an M.Sc. in geomatics engineering from the Swiss Federal Institute of Technology Lausanne (EPFL).

Jan Skaloud holds Ph.D. and M.Sc. degrees in geomatics engineering from Canada, The University of Calgary, Canada, and Dipl. Ing. degree from CVUT, Prague, Czech Republic.

He is the 2008–2012 Chair of the ISPRS (International Society for Photogrammetry and Remote Sensing) Working Group on Integrated Systems for Mobile Mapping. In 2009 he was recognized by the journal *GPS World* as one of the 50 world's most influential scientists in the field of satellite navigation. He is a lecturer and senior scientist at the Swiss Federal Institute of Technology Lausanne (EPFL).

Maria-Pia Victoria-Feser holds a Ph.D. in econometrics and statistics from the University of Geneva.

She is Professor of Statistics at the University of Geneva. She has been lecturer at the Statistics Department of the London School of Economics (1993–1996) and assistant professor at the Faculty of Psychology and Educational sciences of the University of Geneva (2000–2004).

Her research and publications are in statistics fields (with special emphazis in robust statistics) related to economics (income distribution, inequality and poverty), finance (extreme value modelling, portfolio optimisation), and of social sciences (generalised linear models, models with latent variables).