Attitude Estimation and Optimal Relative Orientation of Redundant Inertial Measurement Units

Corey L. Marcus *
The University of Texas at Austin, Austin, TX, 78712

Renato Zanetti † and Joel W. Getchius ‡
Blue Origin, Kent, WA, 98032

Navigation systems for human-rated space vehicles are typically required to be two-fault tolerant. This paper presents an optimal relative orientation of multiple orthogonal three-axis inertial measurement units in order to maximize fault detection capabilities. The optimality criteria is the entry-wise one norm of the matrix mapping measurement or estimate errors into the fault detection vector. This choice of norm maximizes the projection of each measurement or state error into all components of the fault detection vector. Other matrix norms based on spectral decomposition are shown to be inadequate. A new formulation of the Multiplicative Extended Kalman filter is derived and used to validate the optimal arrangements. The filter uses measurements from one three-axis gyroscope for model propagation and fuses measurements from additional redundant gyros and a star tracker with Kalman updates. Fault detection and isolation is performed with a Cramer-von Mises goodness of fit test using a buffer of stored prefit measurement residuals. It's shown that the faulty gyro axis can be isolated with greater certainty when the optimal arrangement is used.

I. Introduction

High reliability space systems, such as human-rated vehicles, typically employ redundant sensors to design a navigation solution capable of fault detection and isolation (FDI) [1]. Most aerospace navigation systems are based on an aided inertial solution [2], which relies on inertial measurement units (IMUs) to propagate the state in model-replacement mode [3] and use external measurements, such as GPS [4] or optical sensors [5] to compensate for the IMU errors.

Space vehicles such as Orion [6] and the Space Shuttle [7] use multiple inertial measurement units (IMUs). By comparing the redundant information, the system is able to identify mismatches (fault detection). In the presence of enough redundancy, the faulty measurement can also be isolated. The Space Shuttle used a mid-value select scheme to avoid faulty measurements even in the presence of small discrepancies. In the context of three sensors, measurements

*Graduate Student, Aerospace Engineering and Engineering Mechanics, Corresponding Author (corey.marcus@gmail.com)
†Navigation Engineer - Lunar Transportation
‡Navigation Manager - Lunar Transportation

from a failed IMU disagree with the other two, hence mid-value selection avoids the faulty measurement. The Space Shuttle also used an error threshold to identify larger discrepancies and to remove failed IMUs from the candidate set [8].

Sturza [9] proposed the parity space FDI technique [10], where linear redundant measurements are stacked together and a linear transformation of variables is used to map the measurement space into a full-rank component and a null-space component, called the parity space. From the parity matrix one can calculate a parity vector and a fault vector. Thresholds are used to detect faults based on the magnitude of the fault vector.

The key to effective FDI is the overlap of measurements. In the context of inertial strapdown sensors, two approaches are typically employed to achieve redundancy. The first is using so-called Redundant IMUs (RIMUs) where each IMU box has more than three sensing axis [11]. The second approach is the use of multiple three-axis orthogonal IMUs [12].

Jafari [13] proposes an approach to determine optimal configuration of RIMUs, i.e., how to orient the redundant axes relative to each other to maximize overall sensor accuracy. The analysis is repeated with and without single-axis failures. The approach in [13] maximizes a geometric dilution of precision (GDOP) performance index similar to what is used in GPS.

In the context of multiple three-axis IMUs (which are typically encountered in human-rated space vehicles) two FDI approaches are possible: axis-level checks and box-level checks. Axis-level checks compare each of the three orthogonal measurements of the IMUs against the measurements of all other IMUs. For example, in the case of three IMUs, a total of nine scalar measurements of acceleration and nine of angular velocity are tested. The advantage of this methodology is that two single-axis faults can be detected and one isolated with just two IMU boxes. The difficulty of this approach is that tight calibration of mounting errors between sensors is needed in order to detect axis-level faults accurately and effectively.

Box-level checks compare magnitudes of IMU measurements (acceleration and angular rate) which are insensitive to relative orientation errors of the IMU boxes. This approach requires three IMU boxes in order to detect two single-axis faults and to isolate one.

Reference [14] documents the system for NASA's Orion capsule which operates with three IMUs. The raw measurements are compared directly. If any two measurements differ for long enough, a gyro fault is declared. Reference [15] develops a FDI system for three orthogonal single axis gyros plus one redundant single axis gyro based on independent component analysis. They do not consider gyro biases or misalignments due to manufacturing defects or changing environmental conditions. Reference [16] uses a sliding window generalized likelihood ratio test to determine if a set of measurements are likely to have arisen from a sensor failure. They do not specifically compare redundant measurements, but instead compare measurements to a priori knowledge of system measurement and process noise.

IMU FDI is typically performed prior to the filtering step and does not employ the navigation estimates of the repeatable IMU errors. Vehicles with multiple IMUs typically use them in a model replacement mode. Usually, one IMU is selected to propagate the vehicle's state estimate while the others are used only for fault-detection.

Reference [17] proposes a scheme where six low grade gyros have their measurements averaged and fed into an multiplicative extended Kalman filter (MEKF). They show that this method produces similar performance to using a single tactical grade gyro. [18] shows very similar results also with an averaging scheme. These authors argue that averaging gyro measurements produces a single measurement with better noise characteristics. There are several drawbacks with this methodology; notably the biases of each gyros are not observable. Also, the low cost devices utilized often have significant bias drift rates and identifying failure may prove difficult.

Reference [19] takes an unorthodox approach and use a bank of nonlinear observers; each with their own gyro. This choice is motivated by convergence guarantees for the observers which do not exist for traditional EKFs or MEKFs. Other authors use the more traditional tactic and fuse the outputs of several different filters. The work of [20] compares several different filtering schemes. They compare measurement averaging, a single EKF fusing all measurements, a bank of EKFs for each gyro, and a bank of EKFs for each gyro with a reset of the covariance of each element in the bank to match the covariance of the output. The implementations all require dynamics models for propagation and no optimization of the relative orientations is performed. Reference [21] also compares several different methodologies; a single EKF fusing all measurements, a bank of EKFs for each measurement, and a bank of EKFs with physics-informed geometric constraints to assist in fusing their outputs. They point out that when each filter in the bank has its absolute attitude updated with the same measurement, the errors within the banks become correlated. This significantly complicates fusing their outputs. Like the previous work, all of their implementations require knowledge of vehicle dynamics and no sensor optimization is performed.

This paper contributes to IMU FDI in two significant ways. First, it derives the best configuration of multiple three-axis IMU sensors in order to aid axis-level FDI. Assume, for example, that two IMUs were oriented parallel to each other. One failure of one axis would make faults in the parallel axis of the other box undetectable. While the system has three redundant measurements, it can only detect one failure and isolate none. We seek skewed relative orientations between sensors to maximize the detection of multiple failures.

The second contribution concerns an IMU's angular rate measurements. We propose a methodology for using measurements from any number of three-axis gyros in a single Kalman filter and present a methodology to perform gyro FDI inside the filter. Our methodology centers on a modified version of the MEKF. One gyro is used in model replacement mode and the remaining are fused in a newly derived Kalman update. This approach allows fault detection of the gyro measurements compensated with the estimates of their repeatable errors, while crucially taking into account their correlation. Assume for example a single star tracker was used to estimate the biases of three gyros. Clearly, the biases estimation error of the three gyros would be correlated and this correlation needs to be accounted for. Although this filter is restricted to angular rate measurements and attitude estimation the methodology and techniques used for its formulation could easily be extended to include acceleration measurements so that all IMU outputs are used.

The remainder of this document is organized as follows. We first review the parity space technique in section II, we

then analytically show how to calculate the best orientation of a single redundant axis added to a three-axis orthogonal IMU in section III.A. Next, we expand this analytic solution to two three-axis orthogonal IMUs and derive conditions of optimality in section III.B. We leverage these optimality conditions to demonstrate our choice of relative orientations for three orthogonal IMUs is indeed optimal in section III.C. We leverage parity space to justify our choice of performance index and draw some conclusions in section III. The remainder of this paper focuses mostly on gyros, but all results can be also applied to accelerometers. Section IV describes a new form of the MEKF for redundant three-axis gyros and an accompanying FDI framework.

II. Parity Space

In this section we briefly review the fault vector and parity space methodology developed by Sturza [9], we follow the derivation of [12]. While parity space is not the only means of performing FDI, its definition of fault vector aids in interpreting our proposed optimality criteria.

Assume a linear measurement of the n dimensional state $\boldsymbol{x}$, i.e. $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is the measurement error. Assume also that the measurement vector is redundant, i.e. $\boldsymbol{y} \in \mathbb{R}^m$, with $m > n$ and $\boldsymbol{H}$ is full rank. Following Sturza's derivation we define the generalized inverse $\boldsymbol{H}^*$ and the parity matrix $\boldsymbol{P}$ as

$$\boldsymbol{H}^* = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}} \tag{1}$$

$$\boldsymbol{P} = null(\boldsymbol{H}^{\mathrm{T}})^{\mathrm{T}} \tag{2}$$

The parity matrix, $\boldsymbol{P}$, has rank, $m - n$ dimensions of $m - n \times m$ and has the following properties:

$$\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{I}, \qquad \boldsymbol{P}\boldsymbol{H} = \boldsymbol{0}$$

which follow immediately from equations (1) and (2). We define matrix $\boldsymbol{A}$ to map the measurement space into the state space and the parity space, and is given by

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{H}^* \\ \\ \boldsymbol{P} \end{bmatrix}$$

The inverse map is easily obtained since

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{H} & \boldsymbol{P}^{\mathrm{T}} \end{bmatrix}$$

The parity vector $\boldsymbol{p}$ is defined as $\boldsymbol{p} = \boldsymbol{P}\boldsymbol{y}$. The fault vector $\boldsymbol{\epsilon}$ is defined as the parity vector transformed back to the

measurement space

$$\boldsymbol{\epsilon} = \boldsymbol{A}^{-1} \begin{bmatrix} 0 \\ \boldsymbol{p} \end{bmatrix} = \boldsymbol{P}^{\mathrm{T}} \boldsymbol{P} \boldsymbol{y} = (\boldsymbol{I} - \boldsymbol{H}\boldsymbol{H}^{*}) \boldsymbol{y} = \boldsymbol{S}\boldsymbol{y} = \boldsymbol{S}\boldsymbol{\eta}$$

where matrix $\boldsymbol{S}$ is symmetric and the last equality holds because $\boldsymbol{S}\boldsymbol{H} = \boldsymbol{O}$. The fault vector has the useful property of being independent of the state. Notice that no singular value decomposition is needed since null space computations are unnecessary, this is due to $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P} = \boldsymbol{I} - \boldsymbol{H}\boldsymbol{H}^{*}$. Notice that

$$\mathbb{E}\left\{\epsilon\right\} = \boldsymbol{S}\,\mathbb{E}\left\{\boldsymbol{\eta}\right\} \tag{3}$$

$$\mathbb{E}\left\{\epsilon\epsilon^{\mathrm{T}}\right\} = \boldsymbol{S}\,\mathbb{E}\left\{\boldsymbol{\eta}\boldsymbol{\eta}^{\mathrm{T}}\right\}\boldsymbol{S}^{\mathrm{T}} \tag{4}$$

As derived by Sturza [9], fault detection is based on the decision scalar $d = \boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{\epsilon} = \boldsymbol{p}^{\mathrm{T}}\boldsymbol{p}$. Assuming $\boldsymbol{\eta}$ to be zero-mean and Gaussian with covariance matrix $\sigma^2\boldsymbol{I}$, then d/σ^2 is χ^2 distributed with $m - n$ degrees of freedom. A fault is detected when d/σ^2 exceeds a user-defined threshold, which is set based on the desired probability of false alarm and missed detection [9]. For $m - n \geq 2$, the fault can also be isolated and the faulty measurement is associated to the maximum value of ϵ_i^2/S_{ii} [9]. Alternatively, it is possible to first check the maximum value of ϵ_i^2/S_{ii}, and declare it faulty if it exceeds a threshold $\epsilon_i^2/S_{ii} > \tau\sigma^2$, where $\tau > 0$ is a tuning parameter. Finally, notice that $\boldsymbol{S}$ is idempotent, i.e., $\boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} = \boldsymbol{S}^2 = \boldsymbol{S}$.

III. Optimal IMU Orientations

Assume N orthogonal three-axis IMUs, stacking all measurements together, for example the gyro measurements, results in

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{T}_A \\ \boldsymbol{T}_B \\ \boldsymbol{T}_C \\ \vdots \end{bmatrix} \boldsymbol{\omega} + \boldsymbol{\eta} = \boldsymbol{H}\boldsymbol{\omega} + \boldsymbol{\eta} \tag{5}$$

where the measurement sensitivity matrix $\boldsymbol{H}$ is composed of the direction cosine matrices expressing the orientation of each IMU.

Under these conditions, we have that

$$\boldsymbol{S} = \boldsymbol{I} - \boldsymbol{H}(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{I} - \frac{1}{N}\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}} \tag{6}$$

with $\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$ having all 3×3 diagonal blocks equal to the identity matrix. Recall from the prior section that matrix $\boldsymbol{S}$

maps the measurement vector into the fault vector. We propose to maximize matrix $\boldsymbol{S}$ in order to maximize our ability to detect faults.

Many matrix norms exist, we choose to maximizing the matrix $\boldsymbol{L}_1$ norm of $\boldsymbol{S}$. The matrix L_1 norm defined as the sum of the absolute values of the matrix entries

$$\|\boldsymbol{M}\|_{L_1} = \sum_i \sum_j |M_{ij}| \tag{7}$$

Maximizing the the matrix $\boldsymbol{L}_1$ norm of $\boldsymbol{S}$ is equivalent to maximizing the $\boldsymbol{L}_1$ norm of $\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$, which is a matrix whose entries are the projection of all sensing axes into each other. It therefore makes intuitive sense to maximize it for fault detection capabilities. An entry-wise one-norm makes the most sense because it weights each projection equally. In contrast, an entry-wise two norm is equivalent to the Frobenious norm and produces:

$$\|\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\|_{L_2} = \sqrt{3}N \tag{8}$$

Clearly this is not a viable performance index because it is independent from the relative orientation of the IMUs.

Matrices $\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$ and $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$ share the same non-zero eigenvalues and singular values (all three equal to N regardless of the relative orientation of the N sensors). Therefore, any performance index based on the spectral decomposition of either $\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$ or $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$ is not suitable for the purposes of selecting an optimal relative orientation of multiple orthogonal IMUs.

Ref. [13] proposes the optimal relative orientation of sensing axis in a redundant IMU, i.e. an IMU with more than three sensing axes. They select the optimal orientation by optimizing one of two performance indexes related to the following matrix

$$\boldsymbol{Q} = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1} \tag{9}$$

The two performance indices are

$$\mathcal{J}_4 = \sqrt{\operatorname{trace} \boldsymbol{Q}} \tag{10}$$

$$\mathcal{J}_5 = \sqrt{\det \boldsymbol{Q}} \tag{11}$$

The goal of Ref. [13] is to maximize the accuracy of the solution, and indeed matrix $\boldsymbol{Q}$ is used in solving the normal equations [22]. However, $\boldsymbol{Q}$ is also used for parity space computations, see Eq. (1). Both $\mathcal{J}_4$ and $\mathcal{J}_5$ are functions of the eigenvalues of $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$, and hence are not appropriate for our problem.

A. One IMU and one additional sensing axis

First, we derive the optimal orientation of a single additional sensing axis in the presence of one IMU with three orthogonal axes. We denote the IMU as IMU A, and without any loss of generality we assume its axes are aligned with the canonical bases: $\begin{bmatrix} \boldsymbol{x}_a & \boldsymbol{y}_a & \boldsymbol{z}_a \end{bmatrix} = \boldsymbol{I}_{3\times3}$, where $\boldsymbol{I}_{3\times3}$ is the 3-by-3 identity matrix. Let's denote the direction of the additional sensing axis with the unit-vector $\boldsymbol{v}$, hence the stacked measurements $\boldsymbol{y} \in \mathbb{R}^4$ of angular velocity are:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{\omega} + \boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{I}_{3\times3} \\ \boldsymbol{v}^{\mathrm{T}} \end{bmatrix} \boldsymbol{\omega} + \boldsymbol{\eta} \tag{12}$$

where $\boldsymbol{H} \in \mathbb{R}^{4\times3}$ and $\boldsymbol{\eta}$ is the measurement error. After some manipulation, the generalized inverse $\boldsymbol{H}^* \in \mathbb{R}^{3\times4}$ and the parity matrix $\boldsymbol{P} \in \mathbb{R}^{1\times4}$ are given by

$$\boldsymbol{H}^* = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}} = \frac{1}{2}\begin{bmatrix} 2\boldsymbol{I}_{3\times3} - \boldsymbol{v}\boldsymbol{v}^{\mathrm{T}} & \boldsymbol{v} \end{bmatrix} \tag{13}$$

$$\boldsymbol{P} = null(\boldsymbol{H}^{\mathrm{T}})^{\mathrm{T}} = \frac{1}{\sqrt{2}}\begin{bmatrix} \boldsymbol{v}^{\mathrm{T}} & -1 \end{bmatrix} \tag{14}$$

$$\boldsymbol{S} = \boldsymbol{I}_{4\times4} - \boldsymbol{H}\boldsymbol{H}^* = \frac{1}{2}\left(\boldsymbol{I}_{4\times4} - \begin{bmatrix} [\boldsymbol{v}\times]^2 & \boldsymbol{v} \\ \boldsymbol{v}^{\mathrm{T}} & 0 \end{bmatrix} \right) \tag{15}$$

where the cross-vector skew-symmetric matrix $[\cdot\times]$ and the fact that $\boldsymbol{v}$ has unit norm were used. Therefore maximizing $\|\boldsymbol{S}\|_{L_1}$ is equivalent to maximizing the vector one-norm $\|\boldsymbol{v}\|_1$:

$$\max_{\boldsymbol{v}} \mathcal{J}_1 = \|\boldsymbol{v}\|_1, \qquad \text{subject to } \|\boldsymbol{v}\|_2 = 1 \tag{16}$$

Without any loss of generality we look for a vector $\boldsymbol{v}$ having all positive coordinates; other solutions in the other seven octants are trivial extensions of this one. The performance index implies that we want to maximize the projection of $\boldsymbol{v}$ into all axes of IMU A.

Define

$$\boldsymbol{v} = \begin{bmatrix} a \\ b \\ \sqrt{1 - a^2 - b^2} \end{bmatrix} \tag{17}$$

hence

$$\mathcal{J}_1 = \|\boldsymbol{v}\|_1 = a + b + \sqrt{1 - a^2 - b^2} \tag{18}$$

$$\frac{\partial \mathcal{J}_1}{\partial a} = 1 - \frac{a}{\sqrt{1 - a^2 - b^2}} \tag{19}$$

$$\frac{\partial \mathcal{J}_1}{\partial b} = 1 - \frac{b}{\sqrt{1 - a^2 - b^2}} \tag{20}$$

Setting the partials to zero we obtain the necessary conditions

$$\sqrt{1 - a^2 - b^2} = a \tag{21}$$

$$\sqrt{1 - a^2 - b^2} = b \tag{22}$$

Hence our optimally condition is $v_1 = v_2 = v_3 = \sqrt{3}/3$ and the optimal performance index is $\mathcal{J}_1^{opt} = \sqrt{3} \approx 1.73$. Intuitively this answer makes perfect sense, the bisector vector is exactly what we expected the answer should be, validating our intuition for the performance index.

The second order derivatives are

$$\frac{\partial^2 \mathcal{J}_1}{\partial a^2} = -\frac{1}{\sqrt{1 - a^2 - b^2}} - \frac{a^2}{(1 - a^2 - b^2)^{3/2}} = -\frac{1 - b^2}{(1 - a^2 - b^2)^{3/2}} \tag{23}$$

$$\frac{\partial^2 \mathcal{J}_1}{\partial b^2} = -\frac{1 - a^2}{(1 - a^2 - b^2)^{3/2}} \tag{24}$$

$$\frac{\partial^2 \mathcal{J}_1}{\partial b \partial a} = -\frac{ab}{(1 - a^2 - b^2)^{3/2}} \tag{25}$$

The sufficient condition is

$$-(\sqrt{3})^{3/2} \det \begin{bmatrix} 2\sqrt{3}/3 & 1/3 \\ 1/3 & 2\sqrt{3}/3 \end{bmatrix} = -\frac{11(\sqrt{3})^{3/2}}{3} < 0 \tag{26}$$

which is satisfied; ensuring the solution is indeed a maximum.

B. Two IMUs

We now extended the prior solution to two IMUs, denoted by A and B. Once again we choose A aligned with the canonical Euclidean ($\boldsymbol{T}_A = \boldsymbol{I}$) base while B has orthogonal axes $\begin{bmatrix} \boldsymbol{x}_b & \boldsymbol{y}_b & \boldsymbol{z}_b \end{bmatrix}$. For simplicity in this section we drop the B index $\boldsymbol{T}_B = \boldsymbol{T}$.

The redundant measurement $\boldsymbol{y} \in \mathbb{R}^6$ of angular velocity is given by:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{\omega} + \boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{I}_{3\times3} \\ \boldsymbol{T} \end{bmatrix} \boldsymbol{\omega} + \boldsymbol{\eta} \tag{27}$$

where $\boldsymbol{H} \in \mathbb{R}^{6\times3}$ and $\boldsymbol{\eta}$ is the measurement error. Following the parity space steps of Section II, the generalized inverse $\boldsymbol{H}^* \in \mathbb{R}^{3\times6}$ and the parity matrix $\boldsymbol{P} \in \mathbb{R}^{3\times6}$ are given by

$$\boldsymbol{H}^* = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}} = \frac{1}{2}\boldsymbol{H}^{\mathrm{T}} \tag{28}$$

$$\boldsymbol{P} = null(\boldsymbol{H}^{\mathrm{T}})^{\mathrm{T}} = \frac{1}{\sqrt{2}}\begin{bmatrix} -\boldsymbol{T} & \boldsymbol{I}_{3\times3} \end{bmatrix} \tag{29}$$

$$\boldsymbol{S} = \boldsymbol{I} - \boldsymbol{H}\boldsymbol{H}^* = \frac{1}{2}\begin{bmatrix} \boldsymbol{I}_{3\times3} & -\boldsymbol{T}^{\mathrm{T}} \\ -\boldsymbol{T} & \boldsymbol{I}_{3\times3} \end{bmatrix} \tag{30}$$

Therefore maximizing $\|\boldsymbol{S}\|_{L_1}$ is equivalent to maximizing $\|\boldsymbol{T}\|_{L_1}$. Then our performance index is given by

$$\mathcal{J}_2 = \|\boldsymbol{T}_A^{\mathrm{T}}\boldsymbol{T}_B\|_{L_1} = \|\boldsymbol{T}\|_{L_1} = -3 + \frac{1}{2}\left\|\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}\right\|_{L_1} \tag{31}$$

Since we are now trying to optimize three axes rather than one (albeit three axes constrained to be orthogonal to each other), we must have that

$$\mathcal{J}_2^{opt} \leq 3\mathcal{J}_1^{opt} = 3\sqrt{3} \approx 5.1962 \tag{32}$$

it turns out that $\mathcal{J}_2^{opt} = 5$ and the optimal solution is derived in Ref. [23] and given by:

$$\boldsymbol{T} = \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \tag{33}$$

Permutations of the coordinates of IMU B's axes and/or change of coordinate signs that result in an orthogonal matrix $\boldsymbol{T}_B$ will also produce an optimal solution with $\mathcal{J}_2 = 5$.

Even after two isolated faults in IMU B, the remaining axis of B is able to detect faults in all three axis of IMU A with a performance index of $\mathcal{J}_1 = 5/3 \approx 1.66$ only slightly below the optimal value of $\sqrt{3} \approx 1.73$.

1. First Order Optimality Conditions

A methodology to checking the first order conditions to be an extremum of $\mathcal{J}_2$ is now outlined. Define the entries of $\boldsymbol{T}$ as follows:

$$\boldsymbol{T} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \tag{34}$$

The first order differential of an orthogonal matrix is given by

$$d\boldsymbol{T} = [d\boldsymbol{\alpha}\times]\,\boldsymbol{T} = \begin{bmatrix} 0 & -d\alpha_3 & d\alpha_2 \\ d\alpha_3 & 0 & -d\alpha_1 \\ -d\alpha_2 & d\alpha_1 & 0 \end{bmatrix} \boldsymbol{T} \tag{35}$$

$$= \begin{bmatrix} T_{31}d\alpha_2 - T_{21}d\alpha_3 & T_{32}d\alpha_2 - T_{22}d\alpha_3 & T_{33}d\alpha_2 - T_{23}d\alpha_3 \\ T_{11}d\alpha_3 - T_{31}d\alpha_1 & T_{12}d\alpha_3 - T_{32}d\alpha_1 & T_{13}d\alpha_3 - T_{33}d\alpha_1 \\ T_{21}d\alpha_1 - T_{11}d\alpha_2 & T_{22}d\alpha_1 - T_{12}d\alpha_2 & T_{23}d\alpha_1 - T_{13}d\alpha_2 \end{bmatrix} \tag{36}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ -T_{31} & -T_{32} & -T_{33} \\ T_{21} & T_{22} & T_{23} \end{bmatrix} d\alpha_1 + \begin{bmatrix} T_{31} & T_{32} & T_{33} \\ 0 & 0 & 0 \\ -T_{11} & -T_{12} & -T_{13} \end{bmatrix} d\alpha_2 + \begin{bmatrix} -T_{21} & -T_{22} & -T_{23} \\ T_{11} & T_{12} & T_{13} \\ 0 & 0 & 0 \end{bmatrix} d\alpha_3 \tag{37}$$

Define the element-wise absolute value operator

$$|\boldsymbol{T}| = \begin{bmatrix} |T_{11}| & |T_{12}| & |T_{13}| \\ |T_{21}| & |T_{22}| & |T_{23}| \\ |T_{31}| & |T_{32}| & |T_{33}| \end{bmatrix} \tag{38}$$

then

$$d|\boldsymbol{T}| = |\boldsymbol{T} + d\boldsymbol{T}| - |\boldsymbol{T}|$$

$$= \begin{bmatrix} \text{sgn}\,(T_{11})\,(T_{31}d\alpha_2 - T_{21}d\alpha_3) & \text{sgn}\,(T_{12})\,(T_{32}d\alpha_2 - T_{22}d\alpha_3) & \text{sgn}\,(T_{13})\,(T_{33}d\alpha_2 - T_{23}d\alpha_3) \\ \text{sgn}\,(T_{21})\,(T_{11}d\alpha_3 - T_{31}d\alpha_1) & \text{sgn}\,(T_{22})\,(T_{12}d\alpha_3 - T_{32}d\alpha_1) & \text{sgn}\,(T_{23})\,(T_{13}d\alpha_3 - T_{33}d\alpha_1) \\ \text{sgn}\,(T_{31})\,(T_{21}d\alpha_1 - T_{11}d\alpha_2) & \text{sgn}\,(T_{32})\,(T_{22}d\alpha_1 - T_{12}d\alpha_2) & \text{sgn}\,(T_{33})\,(T_{23}d\alpha_1 - T_{13}d\alpha_2) \end{bmatrix} \tag{39}$$

Assuming that $\boldsymbol{T}$ has the signs as in Eq. (34)

$$\text{sign}\,\boldsymbol{T} = \begin{bmatrix} + & + & + \\ - & - & + \\ + & - & + \end{bmatrix} \tag{40}$$

we have that the first order differential of the performance index is

$$\begin{aligned} d\mathcal{J}_2 = \sum_{i=1}^{3} \frac{\partial \mathcal{J}_2}{\partial \alpha_i} d\alpha_i &= (T_{31} + T_{32} - T_{33} + T_{21} - T_{22} + T_{23})d\alpha_1 \\ &+ (T_{31} + T_{32} + T_{33} - T_{11} + T_{12} - T_{13})d\alpha_2 + (-T_{21} - T_{22} - T_{23} - T_{11} - T_{12} + T_{13})d\alpha_3 \end{aligned} \tag{41}$$

Therefore, the conditions for optimality are:

$$\frac{\partial \mathcal{J}_2}{\partial \alpha_1} = T_{31} + T_{32} - T_{33} + T_{21} - T_{22} + T_{23} = 0 \tag{42}$$

$$\frac{\partial \mathcal{J}_2}{\partial \alpha_2} = T_{31} + T_{32} + T_{33} - T_{11} + T_{12} - T_{13} = 0 \tag{43}$$

$$\frac{\partial \mathcal{J}_2}{\partial \alpha_3} = -T_{21} - T_{22} - T_{23} - T_{11} - T_{12} + T_{13} = 0 \tag{44}$$

And it is trivial to verify that the optimal matrix outlined above satisfies these conditions

$$\boldsymbol{T} = \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \tag{45}$$

Notice that one can also compute the first order differential by post-multiplying $\boldsymbol{T}$ by a skew symmetric matrix differential

$$d\boldsymbol{T} = \boldsymbol{T}\,[d\boldsymbol{\alpha}\times] = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \begin{bmatrix} 0 & -d\alpha_3 & d\alpha_2 \\ d\alpha_3 & 0 & -d\alpha_1 \\ -d\alpha_2 & d\alpha_1 & 0 \end{bmatrix} \tag{46}$$

$$= \begin{bmatrix} T_{12}d\alpha_3 - T_{13}d\alpha_2 & T_{13}d\alpha_1 - T_{11}d\alpha_3 & T_{11}d\alpha_2 - T_{12}d\alpha_1 \\ T_{22}d\alpha_3 - T_{23}d\alpha_2 & T_{23}d\alpha_1 - T_{21}d\alpha_3 & T_{21}d\alpha_2 - T_{22}d\alpha_1 \\ T_{32}d\alpha_3 - T_{33}d\alpha_2 & T_{33}d\alpha_1 - T_{31}d\alpha_3 & T_{31}d\alpha_2 - T_{32}d\alpha_1 \end{bmatrix} \tag{47}$$

$$= \begin{bmatrix} 0 & T_{13} & -T_{12} \\ 0 & T_{23} & -T_{22} \\ 0 & T_{33} & -T_{32} \end{bmatrix} d\alpha_1 + \begin{bmatrix} -T_{13} & 0 & T_{11} \\ -T_{23} & 0 & T_{21} \\ -T_{33} & 0 & T_{31} \end{bmatrix} d\alpha_2 + \begin{bmatrix} T_{12} & -T_{11} & 0 \\ T_{22} & -T_{21} & 0 \\ T_{32} & -T_{31} & 0 \end{bmatrix} d\alpha_3 \tag{48}$$

C. Three IMUs

Consider the case of three IMUs, A, B, and C; once again A is aligned with the canonical Euclidean base, B has axes $\boldsymbol{T}_B = \begin{bmatrix} \boldsymbol{x}_b & \boldsymbol{y}_b & \boldsymbol{z}_b \end{bmatrix}$, and C has axes $\boldsymbol{T}_C = \begin{bmatrix} \boldsymbol{x}_c & \boldsymbol{y}_c & \boldsymbol{z}_c \end{bmatrix}$. The stacked redundant measurement $\boldsymbol{y} \in \mathbb{R}^9$ of angular velocity is:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{\omega} + \boldsymbol{\eta} = \begin{bmatrix} \boldsymbol{I}_{3\times3} \\ \boldsymbol{T}_B \\ \boldsymbol{T}_C \end{bmatrix} \boldsymbol{\omega} + \boldsymbol{\eta} \tag{49}$$

where $\boldsymbol{H} \in \mathbb{R}^{9\times3}$ and once again $\boldsymbol{\eta}$ is the measurement error. The generalized inverse $\boldsymbol{H}^* \in \mathbb{R}^{3\times9}$ is

$$\boldsymbol{H}^* = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}} = \frac{1}{3}\boldsymbol{H}^{\mathrm{T}} \tag{50}$$

Matrix $\boldsymbol{S}$ is

$$\boldsymbol{S} = \boldsymbol{I} - \boldsymbol{H}\boldsymbol{H}^* = \frac{1}{3}\begin{bmatrix} 2\boldsymbol{I}_{3\times3} & -\boldsymbol{T}_B^{\mathrm{T}} & -\boldsymbol{T}_C^{\mathrm{T}} \\ -\boldsymbol{T}_B & 2\boldsymbol{I}_{3\times3} & -\boldsymbol{T}_B\boldsymbol{T}_C^{\mathrm{T}} \\ -\boldsymbol{T}_C & -\boldsymbol{T}_C\boldsymbol{T}_B^{\mathrm{T}} & 2\boldsymbol{I}_{3\times3} \end{bmatrix} \boldsymbol{\eta} \tag{51}$$

Therefore maximizing the L_1 norm of matrix $\boldsymbol{S}$ is equivalent to maximizing

$$\mathcal{J}_3 = \|\boldsymbol{T}_A^{\mathrm{T}}\boldsymbol{T}_B\|_{L_1} + \|\boldsymbol{T}_A^{\mathrm{T}}\boldsymbol{T}_C\|_{L_1} + \|\boldsymbol{T}_C^{\mathrm{T}}\boldsymbol{T}_B\|_{L_1} \tag{52}$$

$$= \|\boldsymbol{T}_B\|_{L_1} + \|\boldsymbol{T}_C\|_{L_1} + \|\boldsymbol{T}_C^{\mathrm{T}}\boldsymbol{T}_B\|_{L_1} \tag{53}$$

Because we are now optimizing the sum of the norm of three orthogonal matrices, we have that

$$\mathcal{J}_3^{opt} \leq 3\mathcal{J}_2^{opt} = 15 \tag{54}$$

it turns out that $\mathcal{J}_3^{opt} = 3\sqrt{6} + 15/2 \approx 14.8485$ and one possible optimal solutions is

$$\boldsymbol{x}_a = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{y}_a = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \boldsymbol{z}_a = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{55}$$

$$\boldsymbol{x}_b = \begin{bmatrix} 3/4 \\ \sqrt{6}/4 \\ -1/4 \end{bmatrix} \qquad \boldsymbol{y}_b = \begin{bmatrix} 1/4 \\ -\sqrt{6}/4 \\ -3/4 \end{bmatrix} \qquad \boldsymbol{z}_b = \begin{bmatrix} -\sqrt{6}/4 \\ 1/2 \\ -\sqrt{6}/4 \end{bmatrix} \tag{56}$$

$$\boldsymbol{x}_c = \begin{bmatrix} 3/4 \\ -\sqrt{6}/4 \\ -1/4 \end{bmatrix} \qquad \boldsymbol{y}_c = \begin{bmatrix} 1/4 \\ \sqrt{6}/4 \\ -3/4 \end{bmatrix} \qquad \boldsymbol{c}_c = \begin{bmatrix} \sqrt{6}/4 \\ 1/2 \\ \sqrt{6}/4 \end{bmatrix} \tag{57}$$

Once again, many optimal solutions exist, and they are permutations of the values above.

The optimal matrices are

$$\boldsymbol{T}_B = \begin{bmatrix} 3/4 & 1/4 & -\sqrt{6}/4 \\ \sqrt{6}/4 & -\sqrt{6}/4 & 1/2 \\ -1/4 & -3/4 & -\sqrt{6}/4 \end{bmatrix} \tag{58}$$

$$\boldsymbol{T}_C = \begin{bmatrix} 3/4 & 1/4 & \sqrt{6}/4 \\ -\sqrt{6}/4 & \sqrt{6}/4 & 1/2 \\ -1/4 & -3/4 & \sqrt{6}/4 \end{bmatrix} \tag{59}$$

Let's define their first order differentials as

$$d\boldsymbol{T}_B = \boldsymbol{T}_B[d\boldsymbol{\beta}\times] \qquad\qquad d\boldsymbol{T}_C = \boldsymbol{T}_C\ [d\boldsymbol{\gamma}\times] \tag{60}$$

hence

$$d(\boldsymbol{T}_C^{\mathrm{T}}\boldsymbol{T}_B) = \boldsymbol{T}_C^{\mathrm{T}}\boldsymbol{T}_B[d\boldsymbol{\beta}\times] - [d\boldsymbol{\gamma}\times]\,\boldsymbol{T}_C^{\mathrm{T}}\boldsymbol{T}_B \tag{61}$$

where

$$\boldsymbol{T}_C^{\mathrm{T}}\boldsymbol{T}_B = \begin{bmatrix} 1/4 & 3/4 & -\sqrt{6}/4 \\ 3/4 & 1/4 & \sqrt{6}/4 \\ \sqrt{6}/4 & -\sqrt{6}/4 & -1/2 \end{bmatrix} \tag{62}$$

with the above values of $\boldsymbol{T}_B$ and $\boldsymbol{T}_C$, the first order conditions are

$$\frac{\partial \mathcal{J}_3}{\partial \beta_1} = \sum \left(\begin{bmatrix} 0 & -\sqrt{6}/4 & 1/4 \\ 0 & -1/2 & \sqrt{6}/4 \\ 0 & \sqrt{6}/4 & -3/4 \end{bmatrix} + \begin{bmatrix} 0 & -\sqrt{6}/4 & 3/4 \\ 0 & \sqrt{6}/4 & -1/4 \\ 0 & 1/2 & -\sqrt{6}/4 \end{bmatrix} \right) = 0 \tag{63}$$

$$\frac{\partial \mathcal{J}_3}{\partial \beta_2} = \sum \left(\begin{bmatrix} \sqrt{6}/4 & 0 & -3/4 \\ -1/2 & 0 & \sqrt{6}/4 \\ -\sqrt{6}/4 & 0 & 1/4 \end{bmatrix} + \begin{bmatrix} \sqrt{6}/4 & 0 & -1/4 \\ -\sqrt{6}/4 & 0 & 3/4 \\ 1/2 & 0 & -\sqrt{6}/4 \end{bmatrix} \right) = 0 \tag{64}$$

$$\frac{\partial \mathcal{J}_3}{\partial \beta_3} = \sum \left(\begin{bmatrix} 1/4 & -3/4 & 0 \\ -\sqrt{6}/4 & \sqrt{6}/4 & 0 \\ 3/4 & -1/4 & 0 \end{bmatrix} + \begin{bmatrix} 3/4 & -1/4 & 0 \\ 1/4 & -3/4 & 0 \\ -\sqrt{6}/4 & \sqrt{6}/4 & 0 \end{bmatrix} \right) = 0 \tag{65}$$

$$\frac{\partial \mathcal{J}_3}{\partial \gamma_1} = \sum \left(\begin{bmatrix} 0 & \sqrt{6}/4 & -1/4 \\ 0 & 1/2 & -\sqrt{6}/4 \\ 0 & -\sqrt{6}/4 & 3/4 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ -\sqrt{6}/4 & \sqrt{6}/4 & 1/2 \\ 3/4 & -1/4 & -\sqrt{6}/4 \end{bmatrix} \right) = 0 \tag{66}$$

$$\frac{\partial \mathcal{J}_3}{\partial \gamma_2} = \sum \left(\begin{bmatrix} -\sqrt{6}/4 & 0 & 3/4 \\ 1/2 & 0 & -\sqrt{6}/4 \\ \sqrt{6}/4 & 0 & -1/4 \end{bmatrix} - \begin{bmatrix} \sqrt{6}/4 & -\sqrt{6}/4 & 1/2 \\ 0 & 0 & 0 \\ -1/4 & 3/4 & -\sqrt{6}/4 \end{bmatrix} \right) = 0 \tag{67}$$

$$\frac{\partial \mathcal{J}_3}{\partial \gamma_3} = \sum \left(\begin{bmatrix} 1/4 & -3/4 & 0 \\ -\sqrt{6}/4 & \sqrt{6}/2 & 0 \\ 3/4 & -1/4 & 0 \end{bmatrix} - \begin{bmatrix} -3/4 & -1/4 & \sqrt{6}/4 \\ 1/4 & 3/4 & -\sqrt{6}/4 \\ 0 & 0 & 0 \end{bmatrix} \right) = 0 \tag{68}$$

where the summation is over the matrices entries. The above equations show that the solution is indeed an extremum, it was numerically evaluated to be confirmed a maximum.

1. Three IMUs Parity Space

The fault vector $\boldsymbol{\epsilon}$ is given by

$$\boldsymbol{\epsilon} = \boldsymbol{S}\boldsymbol{y} = \boldsymbol{S}\boldsymbol{\eta} = \frac{1}{3} \begin{bmatrix} 2\boldsymbol{I}_{3\times3} & -\boldsymbol{T}_B^{\mathrm{T}} & -\boldsymbol{T}_C^{\mathrm{T}} \\ -\boldsymbol{T}_B & 2\boldsymbol{I}_{3\times3} & -\boldsymbol{T}_B\boldsymbol{T}_C^{\mathrm{T}} \\ -\boldsymbol{T}_C & -\boldsymbol{T}_C\boldsymbol{T}_B^{\mathrm{T}} & 2\boldsymbol{I}_{3\times3} \end{bmatrix} \boldsymbol{\eta} \tag{69}$$

Assume $\boldsymbol{\eta}$ is zero mean and has covariance matrix given by $\sigma^2\boldsymbol{I}_{9\times9}$. As reviewed in Section II, a fault is declared when $\boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{\epsilon}$ exceeds a predetermined threshold. When this happens, the fault is associated with the maximum value of $|\epsilon_i|$.

Assume that the first axis of IMU A experience a fault ϕ, i.e.

$$\boldsymbol{y}_a = \boldsymbol{H}\boldsymbol{\omega} + \boldsymbol{\eta} + \begin{bmatrix} \phi \\ \boldsymbol{0}_{8\times1} \end{bmatrix} \tag{70}$$

hence

$$\boldsymbol{\epsilon} = \boldsymbol{S}\boldsymbol{\eta} + \frac{\phi}{3} \begin{bmatrix} 2\boldsymbol{x}_a \\ -\boldsymbol{x}_b \\ -\boldsymbol{x}_c \end{bmatrix} \tag{71}$$

if the axis of the three IMUs are aligned the contribution of the fault to ϵ_1 is $2\phi/3$ while its contribution to $\epsilon_{4/7}$ is $-\phi/3$, hence the separation between the two absolute values is $\phi/3$. With our proposed solution the fault contribution to $\epsilon_{4/7}$ is $-\phi/4$ making the separation between the two values $5\phi/12$ which is a 25% improvement over the axes being aligned.

IV. Attitude Estimation with Redundant Gyroscopes

This section focuses on only one portion of the IMU output, angular rate measurements, and shows how multiple IMU measurements can be used simultaneously in a single navigation filter for attitude estimation and fault detection. The techniques and methods used here can be easily extended to include the IMU's accelerometer outputs as well so that

spacecraft position and velocity are estimated together with accelerometer fault detection.

A new MEKF formulation is proposed that use all available gyro measurements for both estimation and FDI. The new formulation employs the IMUs orientation derived above and creates measurements residuals based on their optimized fault detection vectors. In the presence of two IMUs, the filter fuses measurements from both gyros and one star tracker into an estimate of vehicle attitude. This formulation can be extended to any number of IMUs, and the three-IMU case is used to demonstrate FDI capabilities. Attitude is parameterized with a quaternion representing a transformation from an inertial frame into a body-fixed navigation frame $\boldsymbol{q}_{\text{inert}}^{\text{nav}}$. A quaternion's unit-norm constraint does not lend well to direct estimation in a Kalman filter. Instead, an unconstrained three vector parameterization $\boldsymbol{\theta}$ of attitude error between a nominal estimated attitude and the true attitude is estimated.

$$\boldsymbol{q}_{\text{inert}}^{\text{nav}} = \delta\boldsymbol{q}\left(\boldsymbol{\theta}\right) \otimes \left(\boldsymbol{q}_{\text{inert}}^{\text{nav}}\right)^{\text{nom}} \tag{72}$$

where $\otimes$ is the Shuster quaternion product (Equation (73) [24]) and the quaternions are represented as a four-element column-vector with a three dimensional vector $\boldsymbol{q}_v$ and scalar q_s component. We choose a scalar last representation.

$$\boldsymbol{q}_1 \otimes \boldsymbol{q}_2 = \begin{bmatrix} q_{s,1}\boldsymbol{q}_{v,2} + q_{s,2}\boldsymbol{q}_{v,1} - \boldsymbol{q}_{v,1} \times \boldsymbol{q}_{v,2} \\ q_{s,1}q_{s,2} - \boldsymbol{q}_{v,1}^T\boldsymbol{q}_{v,2} \end{bmatrix} \tag{73}$$

$\delta\boldsymbol{q}\left(\boldsymbol{\theta}\right)$ is defined as

$$\delta\boldsymbol{q}\left(\boldsymbol{\theta}\right) = \frac{1}{\sqrt{\frac{1}{4}\boldsymbol{\theta}^T\boldsymbol{\theta} + 1}} \begin{bmatrix} \frac{1}{2}\boldsymbol{\theta} \\ 1 \end{bmatrix}. \tag{74}$$

An approximate inverse operation (valid to first order in the error) is available as

$$\boldsymbol{\theta}(\delta\boldsymbol{q}) \approx 2\delta\boldsymbol{q}_v \tag{75}$$

where $\delta\boldsymbol{q}_v$ is the quaternion's vector component. $\boldsymbol{q}^*$ denotes the inverse of $\boldsymbol{q}$ such that $\boldsymbol{q}^* \otimes \boldsymbol{q}$ and $\boldsymbol{q} \otimes \boldsymbol{q}^*$ both produce the unit quaternion $\boldsymbol{q} \otimes \boldsymbol{q}^* = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$.

For ease of expression, the filter is derived as an estimate of a deviation $\delta\boldsymbol{x}$ from a nominal state $\boldsymbol{x}^{\text{nom}}$ for attitude and the remaining additive states $\boldsymbol{x}^a$.

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{q}_{\text{inert}}^{\text{nav}} \\ \boldsymbol{x}^a \end{bmatrix} \qquad \delta\boldsymbol{x} = \begin{bmatrix} \boldsymbol{\theta} \\ \delta\boldsymbol{x}^a \end{bmatrix} \tag{76}$$

$\delta \boldsymbol{x}$ perturbs $\boldsymbol{x}^{\text{nom}}$ with the $\boxplus$ operator defined as

$$\boldsymbol{x} = \delta \boldsymbol{x} \boxplus \boldsymbol{x}^{\text{nom}} \tag{77}$$

$$\boldsymbol{x} = \begin{bmatrix} \delta \boldsymbol{q}\left(\boldsymbol{\theta}\right) \otimes \left(\hat{\boldsymbol{q}}_{\text{inert}}^{\text{nav}}\right)^{\text{nom}} \\ \delta \boldsymbol{x}^{a} + \left(\boldsymbol{x}^{a}\right)^{\text{nom}} \end{bmatrix}. \tag{78}$$

After all measurements at a given time-step are processed the filter updates the nominal state with the estimated delta state.

$$\boldsymbol{x}^{\text{nom}} \leftarrow \delta \boldsymbol{x} \boxplus \boldsymbol{x}^{\text{nom}} \tag{79}$$

$$\delta \boldsymbol{x} \leftarrow \boldsymbol{0}_{27\times 1} \tag{80}$$

This work occasionally extracts only the additive states from $\boldsymbol{x}^{\text{nom}}$. This can be done with matrix $\boldsymbol{\Lambda}$.

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{0}_{3\times 4} & \boldsymbol{0}_{3\times 24} \\ \boldsymbol{0}_{24\times 4} & \boldsymbol{I}_{24\times 24} \end{bmatrix} \tag{81}$$

$$\begin{bmatrix} \boldsymbol{0}_{3\times 1} \\ \boldsymbol{x}^{a} \end{bmatrix} = \boldsymbol{\Lambda} \boldsymbol{x}^{\text{nom}} \tag{82}$$

The quantity 24 results from the 24 additive states considered in Equation (90).

Each gyro is modeled as having a bias $\boldsymbol{b}$, misalignments $\boldsymbol{\delta}$, scale factors $\boldsymbol{s}$, and additive Gaussian white noise $\boldsymbol{\eta}^{\omega}$ which corrupt the measure of angular rate. Some authors choose to differentiate between three IMU box level mounting misalignments and three internal manufacturing misalignments (also known as non-orthogonality parameters). For the purposes of this work, we refer to all of these simply as misalignments. The misalignment and scale factor terms are combined into a distortion matrix $\boldsymbol{M}^{\Delta}$.

$$\boldsymbol{M}^{\Delta} = \begin{bmatrix} s_x & \delta_{xy} & \delta_{xz} \\ \delta_{yx} & s_y & \delta_{yz} \\ \delta_{zx} & \delta_{zy} & s_z \end{bmatrix} \tag{83}$$

$\boldsymbol{M}_{\Delta}$ will be estimated for each gyro and so it must be vectorized for inclusion in the state space as

$$\boldsymbol{\Delta} = \begin{bmatrix} s_x & \delta_{xy} & \delta_{xz} & \delta_{yx} & s_y & \delta_{yz} & \delta_{zx} & \delta_{zy} & s_z \end{bmatrix}^{T} \tag{84}$$

The operator $[(\boldsymbol{a})\equiv]$ for $\boldsymbol{a}\in\mathbb{R}^3$ is defined as

$$[(\boldsymbol{a})\equiv] = \begin{bmatrix} \boldsymbol{a}^T & \boldsymbol{0}_{1\times3} & \boldsymbol{0}_{1\times3} \\ \boldsymbol{0}_{1\times3} & \boldsymbol{a}^T & \boldsymbol{0}_{1\times3} \\ \boldsymbol{0}_{1\times3} & \boldsymbol{0}_{1\times3} & \boldsymbol{a}^T \end{bmatrix} \tag{85}$$

With this new operator it is possible to write

$$\boldsymbol{M}^{\Delta}\boldsymbol{a} = [(\boldsymbol{a})\equiv]\,\boldsymbol{\Delta}. \tag{86}$$

The i^{th} gyro's continuous time sensed angular rate is

$$\boldsymbol{\omega}^{g_i}(t) = \left(\boldsymbol{I}_{3\times3} + \boldsymbol{M}^{\Delta_i}(t)\right)\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}^{\text{nav}}(t) + \boldsymbol{b}^{g_i}(t) + \boldsymbol{\eta}^{\omega_i}(t) \tag{87}$$

$\boldsymbol{T}_{\text{nav}}^{g_i}$ is a transformation matrix from the navigation frame to the i^{th} gyro frame, $\boldsymbol{\omega}^{\text{nav}}(t)$ is the vehicle's true angular rate in the navigation frame, and $\boldsymbol{\eta}^{\omega_i}(t)$ is a continuous, zero-mean, Gaussian, white process with the following properties.

$$\mathbb{E}\left\{\boldsymbol{\eta}^{\omega_i}(t)\right\} = \boldsymbol{0}_{3\times1} \tag{88}$$

$$\mathbb{E}\left\{\boldsymbol{\eta}^{\omega_i}(t)\boldsymbol{\eta}^{\omega_i}(\tau)^T\right\} = \sigma_\omega^2\delta(t-\tau)\boldsymbol{I}_{3\times3} \tag{89}$$

Here $\delta(a)$ is the dirac delta function.

The filter's delta state contains $\boldsymbol{\theta}$, biases, and distortion vectors.

$$\delta\boldsymbol{x} = \begin{bmatrix} \boldsymbol{\theta} \\ \delta\boldsymbol{b}^{g_1} \\ \delta\boldsymbol{b}^{g_2} \\ \delta\boldsymbol{\Delta}^{g_1} \\ \delta\boldsymbol{\Delta}^{g_2} \end{bmatrix} \tag{90}$$

Biases and distortion vectors are modeled as random walks

$$\dot{\boldsymbol{b}}^{g_i} = \boldsymbol{\eta}^{b_i} \tag{91}$$

$$\dot{\boldsymbol{\Delta}}^{g_i} = \boldsymbol{\eta}^{\Delta_i} \tag{92}$$

These are independent, zero-mean, Gaussian, white-noise processes with

$$\mathbb{E}\left\{\boldsymbol{\eta}^{b_i}(t)\boldsymbol{\eta}^{b_i}(\tau)^T\right\} = \sigma_b^2\delta(t-\tau)\boldsymbol{I}_{3\times3} \tag{93}$$

$$\mathbb{E}\left\{\boldsymbol{\eta}^{\Delta_i}(t)\boldsymbol{\eta}^{\Delta_i}(\tau)^T\right\} = \sigma_\Delta^2\delta(t-\tau)\boldsymbol{I}_{9\times9} \tag{94}$$

A. Faults

This work considers two seperate gyro failure modes. The first is as an unmodeled and constant drift in gyro bias applied to gyro one $\boldsymbol{\epsilon}^{g_1}$.

$$\dot{\boldsymbol{b}}^{g_1}(t) = \boldsymbol{\eta}^{b_1}(t) + \boldsymbol{\epsilon}^{g_1} \tag{95}$$

It's assumed that at filter initialization $\boldsymbol{\epsilon}^{g_1} = \boldsymbol{0}_{3\times1}$. It switches to a nonzero value at $T_{\boldsymbol{\epsilon}}$.

The second failure mode is an increase in the second gyro's angular random walk with matrix A^{g_2}.

$$\boldsymbol{\omega}^{g_2}(t) = \left(I + \boldsymbol{M}_{\Delta_2}(t)\right)\boldsymbol{T}_{\text{nav}}^{g_2}\boldsymbol{\omega}^{\text{nav}}(t) + \boldsymbol{b}^{g_2}(t) + \left(\boldsymbol{I}_{3\times3} + \boldsymbol{A}^{g_2}\right)\boldsymbol{\eta}^{\omega_2}(t) \tag{96}$$

Again it's assumed that when the filter is initialized $\boldsymbol{A}^{g_2} = \boldsymbol{0}_{3\times3}$. It will switch to a non-zero value at T_A.

B. Discretization

The continuous time models in Equations (87), (91), and (92) are discretized from time t_{k-1} to $t_k = t_{k-1} + \Delta t$. The bias evolution can be discretized as

$$\boldsymbol{b}_k^{g_i} = \boldsymbol{b}_{k-1}^{g_i} + \int_{t_{k-1}}^{t_k}\boldsymbol{\eta}^{b_i}(\tau)d\tau. \tag{97}$$

It can be shown that the integral term is a discrete zero-mean Gaussian white-noise process $\boldsymbol{\nu}^{b_i}$ with

$$\boldsymbol{\nu}_{k-1}^{b_i} = \int_{t_{k-1}}^{t_k}\boldsymbol{\eta}^{b_i}(\tau)d\tau \tag{98}$$

$$\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{b_i}\right\} = \boldsymbol{0}_{3\times1} \tag{99}$$

$$\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{b_i}\boldsymbol{\nu}_{k-1}^{b_i\,T}\right\} = \Delta t\sigma_b^2\boldsymbol{I}_{3\times3} \tag{100}$$

Similarly,

$$\boldsymbol{\Delta}_k^{g_i} = \boldsymbol{\Delta}_{k-1}^{g_i} + \int_{t_{k-1}}^{t_k}\boldsymbol{\eta}^{\Delta_i}(\tau)d\tau \tag{101}$$

$$\boldsymbol{\nu}_{k-1}^{\Delta_i} = \int_{t_{k-1}}^{t_k}\boldsymbol{\eta}^{\Delta_i}(\tau)d\tau \tag{102}$$

$$\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{\Delta_i}\right\} = \boldsymbol{0}_{9\times1} \tag{103}$$

$$\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{\Delta_i}\boldsymbol{\nu}_{k-1}^{\Delta_i\,T}\right\} = \Delta t\sigma_{\Delta}^2\boldsymbol{I}_{9\times 9}. \tag{104}$$

The continuous time angular rate from Equation (87) is discretized into a measurement of the average angular rate over the sample period.

$$\boldsymbol{\omega}_k^{g_i} = \frac{1}{\Delta t}\int_{t_{k-1}}^{t_{k-1}+\Delta t}\left(\boldsymbol{I}_{3\times 3} + \boldsymbol{M}^{\Delta_i}(\tau)\right)\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}^{\text{nav}}(\tau) + \boldsymbol{b}^{g_i}(\tau) + \boldsymbol{\eta}^{\omega_i}(\tau)d\tau \tag{105}$$

It is assumed that $\boldsymbol{\omega}^{\text{nav}}(\tau)$ is constant over the sample interval $t \in (t_{k-1},\, t_k]$. By using the $[(\cdot)\equiv]$ operator and replacing $\boldsymbol{b}^{g_i}(\tau)$ and $\boldsymbol{\Delta}^{g_i}(\tau)$ their discrete time solutions we obtain

$$\boldsymbol{\omega}_k^{g_i} = \boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}_k^{\text{nav}} + \left[\left(\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}_k^{\text{nav}}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_i} + \boldsymbol{b}_{k-1}^{g_i} + \boldsymbol{\nu}_k^{\omega_i}. \tag{106}$$

Equation (106) may be equivalently expressed as a function of the discrete-time distortion matrix as

$$\boldsymbol{\omega}_k^{g_i} = \left(\boldsymbol{I}_{3\times 3} + \boldsymbol{M}_{k-1}^{g_i}\right)\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}_k^{\text{nav}} + \boldsymbol{b}_{k-1}^{g_i} + \boldsymbol{\nu}_k^{\omega_i}. \tag{107}$$

$\boldsymbol{\nu}^{\omega_i}$ is a zero-mean discrete Gaussian white-noise process defined as

$$\begin{aligned}\boldsymbol{\nu}_k^{\omega_i} = {} & \left[\left(\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}_k^{\text{nav}}\right)\equiv\right]\frac{1}{\Delta t}\int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{s}\boldsymbol{\eta}^{\Delta_i}(\tau)d\tau ds \\ & + \frac{1}{\Delta t}\int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{s}\boldsymbol{\eta}^{b_i}(\tau)d\tau ds + \frac{1}{\Delta t}\int_{t_{k-1}}^{t_k}\boldsymbol{\eta}^{\omega_i}ds.\end{aligned} \tag{108}$$

This process's covariance and correlation with the other noise processes are required for simulation. It has been shown in [3] that the covariance of $\boldsymbol{\nu}^{\omega_i}$ is

$$\mathbb{E}\left\{\boldsymbol{\nu}_k^{\omega_i}\boldsymbol{\nu}_k^{\omega_i T}\right\} = \frac{\sigma_\omega^2}{\Delta t}\boldsymbol{I}_{3\times 3} + \frac{\sigma_b^2\Delta t}{12}\boldsymbol{I}_{3\times 3} + \frac{\sigma_\Delta^2\Delta t}{12}\left[\left(\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}^{\text{nav}}\right)\equiv\right]\left[\left(\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}^{\text{nav}}\right)\equiv\right]^T. \tag{109}$$

$\boldsymbol{\nu}_k^{\omega_i}$ is correlated with $\boldsymbol{\nu}_{k-1}^{b_i}$ as

$$\mathbb{E}\left\{\boldsymbol{\nu}_k^{\omega_i}\boldsymbol{\nu}_{k-1}^{b_i\,T}\right\} = \frac{\sigma_b^2\Delta t}{2}\boldsymbol{I}_{3\times 3} \tag{110}$$

and $\boldsymbol{\nu}_{k-1}^{\Delta_i}$ as

$$\mathbb{E}\left\{\boldsymbol{\nu}_k^{\omega_i}\boldsymbol{\nu}_{k-1}^{\Delta_i\,T}\right\} = \left[\left(\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}^{\text{nav}}\right)\equiv\right]\frac{\sigma_\Delta^2\Delta t}{2}\boldsymbol{I}_{3\times 3} \tag{111}$$

1. Simulation of Angular Rate Measurements

When simulating angular rate measurements in a later section it will be necessary to account for the correlations between bias/distortion evolution and measurement. To simplify our implementation we introduce a model for $\boldsymbol{\omega}_k^{g_i}$ which we denote $\tilde{\boldsymbol{\omega}}_k^{g_i}$.

$$\tilde{\boldsymbol{\omega}}_k^{g_i} = \boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}_k^{\text{nav}} + \frac{1}{2}\left[\left(\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}_k^{\text{nav}}\right)\equiv\right]\left(\boldsymbol{\Delta}_{k-1}^{g_i} + \boldsymbol{\Delta}_k^{g_i}\right) + \frac{1}{2}\left(\boldsymbol{b}_{k-1}^{g_i} + \boldsymbol{b}_k^{g_i}\right) + \tilde{\boldsymbol{\nu}}_k^{\omega_i} \tag{112}$$

The model is equivalent to the true measurement. It's validity may be checked by evaluating $\mathbb{E}\left\{\tilde{\boldsymbol{\omega}}_k^{g_i}\boldsymbol{b}_k^{g_i T}\right\}$ and $\mathbb{E}\left\{\tilde{\boldsymbol{\omega}}_k^{g_i}\boldsymbol{\Delta}_k^{g_i T}\right\}$.

The variable $\tilde{\boldsymbol{\nu}}_k^{\omega_i}$ is a zero-mean, discrete, Gaussian random variable independent from $\boldsymbol{\nu}_k^{b_i}$ or $\boldsymbol{\nu}_k^{\Delta_i}$ with covariance $\boldsymbol{R}_k^{\omega_i}$ as shown in Equation (113). The derivation details may be found in [3].

$$\boldsymbol{R}_k^{\omega_i} = \frac{\sigma_\omega^2}{\Delta t}\boldsymbol{I}_{3\times3} + \frac{\sigma_b^2\Delta t}{12}\boldsymbol{I}_{3\times3} + \frac{\sigma_\Delta^2\Delta t}{12}\left[\left(\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}_k^{\text{nav}}\right)\equiv\right]\left[\left(\boldsymbol{T}_{\text{nav}}^{g_i}\boldsymbol{\omega}_k^{\text{nav}}\right)\equiv\right]^T \tag{113}$$

C. Propagation with $\omega_k^{g_1}$

A model-replacement strategy is used for the filter's propagation phase from an estimate at time t_{k-1} to a prediction at time t_k. Measurements from the first gyro are used to propagate the vehicle's attitude in lieu of an explicit model of the vehicle's body torques and inertial properties. The formulation for $\boldsymbol{\omega}^{g_1}(t)$ may be inverted to get an expression for $\boldsymbol{\omega}^{\text{nav}}(t)$. Higher order terms are neglected and it is assumed that $(\boldsymbol{I}_{3\times3} + \boldsymbol{M}^{\Delta_i}(t))^{-1} \approx \boldsymbol{I}_{3\times3} - \boldsymbol{M}^{\Delta_i}(t)$. This approximation can be validated by truncating the Taylor series expansion of $k(\boldsymbol{M}) = (\boldsymbol{I}_{3\times3} + \boldsymbol{M}^{\Delta_i}(t))^{-1}$ at first order and assuming $\boldsymbol{M}$ is small.

$$\boldsymbol{\omega}^{\text{nav}}(t) = \boldsymbol{T}_{g_1}^{\text{nav}}\omega^{g_1}(t) - \boldsymbol{T}_{g_1}^{\text{nav}}\boldsymbol{b}^{g_1}(t) - \boldsymbol{T}_{g_1}^{\text{nav}}\boldsymbol{\eta}^{\omega_1}(t) - \boldsymbol{T}_{g_1}^{\text{nav}}\left[\left(\boldsymbol{\omega}^{g_1}\right)\equiv\right]\boldsymbol{\Delta}^{g_1}(t) \tag{114}$$

An estimate of $\boldsymbol{\omega}^{\text{nav}}$ in continuous time is found as

$$\hat{\boldsymbol{\omega}}^{\text{nav}}(t) = \boldsymbol{T}_{g_1}^{\text{nav}}\boldsymbol{\omega}^{g_1}(t) - \boldsymbol{T}_{g_1}^{\text{nav}}\hat{\boldsymbol{b}}^{g_1}(t) - \boldsymbol{T}_{g_1}^{\text{nav}}\left[\left(\boldsymbol{\omega}^{g_1}\right)\equiv\right]\hat{\boldsymbol{\Delta}}^{g_1}(t). \tag{115}$$

The nominal estimate of attitude $\left(\boldsymbol{q}_{\text{inert}}^{\text{nav}}\right)^{\text{nom}}$ is propagated numerically with the standard kinematics formulation.

$$\left(\dot{\boldsymbol{q}}_{\text{inert}}^{\text{nav}}\right)^{\text{nom}} = \frac{1}{2}\begin{bmatrix}\hat{\boldsymbol{\omega}}^{\text{nav}}\\ 0\end{bmatrix}\otimes\left(\boldsymbol{q}_{\text{inert}}^{\text{nav}}\right)^{\text{nom}} \tag{116}$$

The evolution of state $\boldsymbol{\theta}$ can be described with Equation (117) [3].

$$\dot{\boldsymbol{\theta}}(t) = -\left[\hat{\boldsymbol{\omega}}^{\text{nav}}(t)\times\right]\boldsymbol{\theta}(t) + \delta\boldsymbol{\omega}^{\text{nav}}(t) \tag{117}$$

Here the δ prefix represents angular rate estimate error, i.e. $\delta\boldsymbol{\omega}^{\text{nav}} = \boldsymbol{\omega}^{\text{nav}} - \hat{\boldsymbol{\omega}}^{\text{nav}}$. $\delta\boldsymbol{\omega}^{\text{nav}}$ can be expressed in terms of the delta state and measurement noise as

$$\delta\boldsymbol{\omega}^{\text{nav}}(t) = -\boldsymbol{T}_{g_1}^{\text{nav}}\delta\boldsymbol{b}^{g_1}(t) - \boldsymbol{T}_{g_1}^{\text{nav}}\boldsymbol{\eta}^{\omega_1}(t) - \boldsymbol{T}_{g_1}^{\text{nav}}\left[\left(\boldsymbol{\omega}^{g_1}(t)\right)\equiv\right]\delta\boldsymbol{\Delta}^{g_1}(t) \tag{118}$$

So

$$\dot{\boldsymbol{\theta}}(t) = \begin{bmatrix} -\left[\hat{\boldsymbol{\omega}}^{\text{nav}}(t)\times\right] & -\boldsymbol{T}_{g_1}^{\text{nav}} & \boldsymbol{0}_{3\times3} & -\boldsymbol{T}_{g_1}^{\text{nav}}\left[\left(\boldsymbol{\omega}^{g_1}(t)\right)\equiv\right] & \boldsymbol{0}_{3\times9} \end{bmatrix}\delta\boldsymbol{x}(t) - \boldsymbol{T}_{g_1}^{\text{nav}}\boldsymbol{\eta}^{\omega_1}(t) \tag{119}$$

With this, a continuous time stochastic differential equation describing the evolution of the delta state is written.

$$\delta\dot{\boldsymbol{x}}(t) = \begin{bmatrix} -\left[\hat{\boldsymbol{\omega}}^{\text{nav}}(t)\times\right] & -\boldsymbol{T}_{g_1}^{\text{nav}} & \boldsymbol{0}_{3\times3} & -\boldsymbol{T}_{g_1}^{\text{nav}}\left[\left(\boldsymbol{\omega}^{g_1}(t)\right)\equiv\right] & \boldsymbol{0}_{3\times9} \\ \boldsymbol{0}_{24\times3} & \boldsymbol{0}_{24\times3} & \boldsymbol{0}_{24\times3} & \boldsymbol{0}_{24\times9} & \boldsymbol{0}_{24\times9} \end{bmatrix}\delta\boldsymbol{x}(t) + \begin{bmatrix} -\boldsymbol{T}_{g_1}^{\text{nav}} & \boldsymbol{0}_{3\times24} \\ \boldsymbol{0}_{24\times3} & \boldsymbol{I}_{24\times24} \end{bmatrix}\begin{bmatrix} \boldsymbol{\eta}^{\omega_1}(t) \\ \boldsymbol{\eta}^{b_1}(t) \\ \boldsymbol{\eta}^{b_2}(t) \\ \boldsymbol{\eta}^{\Delta_1}(t) \\ \boldsymbol{\eta}^{\Delta_2}(t) \end{bmatrix} \tag{120}$$

$$\delta\dot{\boldsymbol{x}}(t) = F(t)\delta\boldsymbol{x}(t) + Gw(t) \tag{121}$$

If it is assumed that $\hat{\boldsymbol{\omega}}^{\text{nav}}$ and $\boldsymbol{\omega}^{g_1}$ are constant over the sample interval the system becomes linear, time-invariant and an analytic expression for the state transition matrix is available. The measurements of angular rate over the interval $(t_{k-1},\, t_k]$ arrive at t_k. At this point, the best discrete time estimate for $\boldsymbol{\omega}_k^{\text{nav}}$ is given as a function of the previous nominal state because the delta state is always zero when new measurements arrive.

$$\bar{\boldsymbol{\omega}}_k^{\text{nav}} = \boldsymbol{T}_{g_1}^{\text{nav}}\boldsymbol{\omega}_k^{g_1} - \boldsymbol{T}_{g_1}^{\text{nav}}(\boldsymbol{b}_{k-1}^{g_1})^{\text{nom}} - \boldsymbol{T}_{g_1}^{\text{nav}}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right](\boldsymbol{\Delta}_{k-1}^{g_1})^{\text{nom}} \tag{122}$$

The state transition matrix is found as a function of $\bar{\omega}_k^{\text{nav}}$.

$$\boldsymbol{\Phi}(\Delta t) = e^{F\Delta t} = \begin{bmatrix} \boldsymbol{\Phi}_{11} & \boldsymbol{\Phi}_{12}\boldsymbol{T}_{g_1}^{\text{nav}} & \mathbf{0}_{3\times3} & \boldsymbol{\Phi}_{12}\boldsymbol{T}_{g_1}^{\text{nav}}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right] & \mathbf{0}_{3\times9} \\ \mathbf{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times9} & \mathbf{0}_{3\times9} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \boldsymbol{I}_{3\times3} & \mathbf{0}_{3\times9} & \mathbf{0}_{3\times9} \\ \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \boldsymbol{I}_{9\times9} & \mathbf{0}_{9\times9} \\ \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{0}_{9\times9} & \boldsymbol{I}_{9\times9} \end{bmatrix} \tag{123}$$

where

$$\boldsymbol{\Phi}_{11} = \boldsymbol{I}_{3\times3} - \left[\bar{\omega}_k^{\text{nav}}\times\right]\frac{\sin\left(\|\bar{\omega}_k^{\text{nav}}\|\Delta t\right)}{\|\bar{\omega}_k^{\text{nav}}\|} + \left[\bar{\omega}_k^{\text{nav}}\times\right]^2\frac{1-\cos\left(\|\bar{\omega}_k^{\text{nav}}\|\Delta t\right)}{\|\bar{\omega}_k^{\text{nav}}\|^2} \tag{124}$$

$$\begin{aligned}\boldsymbol{\Phi}_{12} = &-\Delta t\boldsymbol{I}_{3\times3} + \left[\bar{\omega}_k^{\text{nav}}\times\right]\frac{1-\cos\left(\|\bar{\omega}_k^{\text{nav}}\|\Delta t\right)}{\|\bar{\omega}_k^{\text{nav}}\|^2} \\ &- \left[\bar{\omega}_k^{\text{nav}}\times\right]^2\frac{\|\bar{\omega}_k^{\text{nav}}\|\Delta t - \sin\left(\|\bar{\omega}_k^{\text{nav}}\|\Delta t\right)}{\|\bar{\omega}_k^{\text{nav}}\|^3}\end{aligned} \tag{125}$$

With the assumption of linear time-invariance the discrete time solution to the linear system in Equation (121) is

$$\delta\boldsymbol{x}_k = \boldsymbol{\Phi}(\Delta t)\delta\boldsymbol{x}_{k-1} + \boldsymbol{\nu}_{k-1}^x \tag{126}$$

$$\boldsymbol{\nu}_{k-1}^x = \int_{t_{k-1}}^{t_k}\boldsymbol{\Phi}(\tau)\boldsymbol{G}\boldsymbol{w}(\tau)d\tau \tag{127}$$

$\boldsymbol{\nu}_{k-1}^x$ is a zero-mean Gaussian white process with covariance $\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^x{\boldsymbol{\nu}_{k-1}^x}^T\right\}$. This covariance can be approximately expressed with the following matrix $\boldsymbol{Q}$ when the angular rates are sufficiently low and constant over the sample period.

$$\boldsymbol{Q} = \begin{bmatrix} \left(\sigma_\omega^2\Delta t + \frac{\sigma_b^2\Delta t^3}{2} + \frac{\|\omega_k^{g_1}\|^2\Delta t^3\sigma_\Delta^2}{3}\right)\boldsymbol{I}_{3\times3} & \frac{-\sigma_b^2\Delta t^2}{2}\boldsymbol{T}_{g_1}^{\text{nav}} & \mathbf{0}_{3\times3} & \frac{-\sigma_\Delta^2\Delta t^2}{2}\boldsymbol{T}_{g_1}^{\text{nav}}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right] & \mathbf{0}_{3\times9} \\ \frac{-\sigma_b^2\Delta t^2}{2}\boldsymbol{T}_{\text{nav}}^{g_1} & \sigma_b^2\Delta t\boldsymbol{I}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times9} & \mathbf{0}_{3\times9} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \sigma_b^2\Delta t\boldsymbol{I}_{3\times3} & \mathbf{0}_{3\times9} & \mathbf{0}_{9\times9} \\ \frac{-\sigma_\Delta^2\Delta t^2}{2}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right]^T\boldsymbol{T}_{\text{nav}}^{g_1} & \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \sigma_\Delta^2\Delta t\boldsymbol{I}_{9\times9} & \mathbf{0}_{9\times9} \\ \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{0}_{9\times9} & \sigma_\Delta^2\Delta t\boldsymbol{I}_{9\times9} \end{bmatrix} \tag{128}$$

The delta state and its associated covariance $\hat{\boldsymbol{\Sigma}}$ are propagated from an estimate at t_{k-1} to a prediction at t_k. The propagation equations are provided for clarity below but it is noted that the estimate at t_{k-1} will always be zero after it is

used to refresh the nominal state at the previous time-step.

$$\delta\bar{\boldsymbol{x}}_k = \boldsymbol{\Phi}(\Delta t)\delta\hat{\boldsymbol{x}}_{k-1} \tag{129}$$

$$\bar{\boldsymbol{\Sigma}}_k = \boldsymbol{\Phi}(\Delta t)\hat{\boldsymbol{\Sigma}}_{k-1}\boldsymbol{\Phi}(\Delta t)^T + \boldsymbol{Q} \tag{130}$$

D. Update with $\omega_k^{g_2}$

Each measurement from the second gyro at the k^{th} time-step is used to update the state. The measurement model starts from the discrete time model for the gyro measurements in Equation (107). The expression for $\boldsymbol{\omega}_k^{g_1}$ is inverted to solve for $\boldsymbol{\omega}_k^{\text{nav}}$ which is substituted into the expression for $\boldsymbol{\omega}_k^{g_2}$.

$$\boldsymbol{\omega}_k^{g_2} = \left(\boldsymbol{I}_{3\times3} + \boldsymbol{M}_{k-1}^{g_2}\right)\boldsymbol{T}_{\text{nav}}^{g_2}\boldsymbol{T}_{g_1}^{\text{nav}}\left(\boldsymbol{I}_{3\times3} + \boldsymbol{M}_{k-1}^{g_1}\right)^{-1}\left(\boldsymbol{\omega}_k^{g_1} - \boldsymbol{b}_{k-1}^{g_1} - \boldsymbol{\nu}_k^{\omega_1}\right) + \boldsymbol{b}_{k-1}^{g_2} + \boldsymbol{\nu}_k^{\omega_2}. \tag{131}$$

Then by neglecting higher order terms and simplifying the transformation matrices

$$\boldsymbol{\omega}_k^{g_2} \approx \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1} - \boldsymbol{T}_{g_1}^{g_2}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_1} - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{b}_{k-1}^{g_1} - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1} + \left[\left(\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_2} + \boldsymbol{b}_{k-1}^{g_2} + \boldsymbol{\nu}_k^{\omega_2}. \tag{132}$$

This is written explicitly as a function of $\boldsymbol{x}_{k-1}^{\text{nom}}$ and $\delta\boldsymbol{x}_{k-1}$. The expression is simplified by defining the jacobian $H = \frac{\partial\boldsymbol{\omega}_k^{g_2}}{\partial\delta\boldsymbol{x}_{k-1}}$.

$$\boldsymbol{\omega}_k^{g_2} = \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1} + \left[\boldsymbol{0}_{3\times3} \quad -\boldsymbol{T}_{g_1}^{g_2} \quad \boldsymbol{I}_{3\times3} \quad -\boldsymbol{T}_{g_1}^{g_2}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right] \quad \left[\left(\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1}\right)\equiv\right]\right]\left(\boldsymbol{\Lambda}\boldsymbol{x}_{k-1}^{\text{nom}} + \delta\boldsymbol{x}_{k-1}\right) - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1} + \boldsymbol{\nu}_k^{\omega_2} \tag{133}$$

$$\boldsymbol{\omega}_k^{g_2} = \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1} + H\left(\boldsymbol{\Lambda}\boldsymbol{x}_{k-1}^{\text{nom}} + \delta\boldsymbol{x}_{k-1}\right) - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1} + \boldsymbol{\nu}_k^{\omega_2} \tag{134}$$

It is noted that the measurement is explicitly a function of the state at t_{k-1} and implicitly a function of the state at time t_k through inclusion of the noise terms $\boldsymbol{\nu}_k^{\omega_i}$. These terms encapsulate process and measurement noise as well as the correlations between them.

Making use of the fact that $\delta\bar{\boldsymbol{x}}$ is zero, the expected value of the linearized measurement model $\bar{\boldsymbol{\omega}}_k^{g_2}$ is given as

$$\bar{\boldsymbol{\omega}}_k^{g_2} = \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1} + \boldsymbol{H}\boldsymbol{\Lambda}\boldsymbol{x}_{k-1}^{\text{nom}} \tag{135}$$

The expected value of the nonlinear measurement model will be required for the Kalman update and is denoted

$$\bar{\hat{\boldsymbol{\omega}}}_k^{g_2} = \boldsymbol{T}_{\text{nav}}^{g_2}\boldsymbol{T}_{g_1}^{\text{nav}}\left(\boldsymbol{I}_{3\times3} + \hat{\boldsymbol{M}}_{k-1}^{g_2}\right)\left(\boldsymbol{I}_{3\times3} + \hat{\boldsymbol{M}}_{k-1}^{g_1}\right)^{-1}\left(\boldsymbol{\omega}_k^{g_1} - \hat{\boldsymbol{b}}_{k-1}^{g_1}\right) + \hat{\boldsymbol{b}}_{k-1}^{g_2}. \tag{136}$$

The Kalman gain $\boldsymbol{K}$ is given by

$$\boldsymbol{K} = \boldsymbol{\Sigma}_{\delta \boldsymbol{x}_k \omega_k^{g_2}} \boldsymbol{\Sigma}^{-1}_{\omega_k^{g_2} \omega_k^{g_2}} \tag{137}$$

where

$$\boldsymbol{\Sigma}_{\delta \boldsymbol{x}_k \omega_k^{g_2}} = \mathbb{E}\left\{ \left(\delta \boldsymbol{x}_k - \mathbb{E}\left\{\delta \boldsymbol{x}_k\right\}\right) \left(\boldsymbol{\omega}_k^{g_2} - \mathbb{E}\left\{\boldsymbol{\omega}_k^{g_2}\right\}\right)^T \right\} \tag{138}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\omega}_k^{g_2} \boldsymbol{\omega}_k^{g_2}} = \mathbb{E}\left\{ \left(\boldsymbol{\omega}_k^{g_2} - \mathbb{E}\left\{\boldsymbol{\omega}_k^{g_2}\right\}\right) \left(\boldsymbol{\omega}_k^{g_2} - \mathbb{E}\left\{\boldsymbol{\omega}_k^{g_2}\right\}\right)^T \right\} \tag{139}$$

$\boldsymbol{\Sigma}_{\delta \boldsymbol{x}_k \omega_k^{g_2}}$ is found through the following manipulations which make use of the fact that $\mathbb{E}\left\{\delta \boldsymbol{x}_k\right\} = \boldsymbol{0}_{27\times 1}$ and that the noise terms are independent of the state.

$$\begin{aligned}
\boldsymbol{\Sigma}_{\delta \boldsymbol{x}_k \omega_k^{g_2}} &= \mathbb{E}\left\{ \left(\delta \boldsymbol{x}_k - \mathbb{E}\left\{\delta \boldsymbol{x}_k\right\}\right) \left(\boldsymbol{\omega}_k^{g_2} - \mathbb{E}\left\{\boldsymbol{\omega}_k^{g_2}\right\}\right)^T \right\} && (140)\\
&= \mathbb{E}\left\{ \left(\boldsymbol{\Phi}(\Delta t)\delta \boldsymbol{x}_{k-1} + \boldsymbol{\nu}_{k-1}^{x}\right) \left(\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1} + \boldsymbol{H}\left(\boldsymbol{\Lambda}\boldsymbol{x}_{k-1}^{\text{nom}} + \delta \boldsymbol{x}_{k-1}\right) - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1} + \boldsymbol{\nu}_k^{\omega_2} - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1} - H\boldsymbol{\Lambda}\boldsymbol{x}_{k-1}^{\text{nom}}\right)^T \right\} && (141)\\
&= \mathbb{E}\left\{ \left(\boldsymbol{\Phi}(\Delta t)\delta \boldsymbol{x}_{k-1} + \boldsymbol{\nu}_{k-1}^{x}\right) \left(\boldsymbol{H}\delta \boldsymbol{x}_{k-1} - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1} + \boldsymbol{\nu}_k^{\omega_2}\right)^T \right\} && (142)\\
&= \boldsymbol{\Phi}(\Delta t)\hat{\boldsymbol{\Sigma}}_{k-1}\boldsymbol{H}^T - \mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{x}\boldsymbol{\nu}_k^{\omega_1 T}\right\}\boldsymbol{T}_{g_2}^{g_1} + \mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{x}\boldsymbol{\nu}_k^{\omega_2 T}\right\} && (143)
\end{aligned}$$

Differences between this formulation and the standard (non-correlated) Kalman filter formulation are immediately apparent and are due to the measurement's explicit inclusion of the state at the previous time-step and correlations between process and measurement noise. The correlation terms are given below.

$$\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{x}\boldsymbol{\nu}_k^{\omega_1 T}\right\} = \begin{bmatrix} -\boldsymbol{T}_{g_1}^{\text{nav}}\left(\left(\sigma_\omega^2 + \frac{\Delta t^2\sigma_b^2}{6}\right)\boldsymbol{I}_{3\times 3} + \frac{\Delta t^2\sigma_\Delta^2}{6}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right]\left[\left(\boldsymbol{T}_{\text{nav}}^{g_1}\bar{\boldsymbol{\omega}}_k^{\text{nav}}\right)\equiv\right]^T\right) \\ \frac{\Delta t\,\sigma_b^2}{2}\boldsymbol{I}_{3\times 3} \\ \boldsymbol{0}_{3\times 3} \\ \frac{\Delta t\,\sigma_\Delta^2}{2}\left[\left(\boldsymbol{T}_{\text{nav}}^{g_1}\boldsymbol{\omega}_k^{\text{nav}}\right)\equiv\right]^T \\ \boldsymbol{0}_{9\times 3} \end{bmatrix} \tag{144}$$

$$\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{x}\boldsymbol{\nu}_{k}^{\omega_2 T}\right\} = \begin{bmatrix} \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} \\ \frac{1}{2}\Delta t\sigma_b^2\boldsymbol{I}_{3\times3} \\ \mathbf{0}_{9\times3} \\ \frac{\Delta t}{2}\sigma_\Delta^2\left[\left(\boldsymbol{T}_{\text{nav}}^{g_2}\bar{\boldsymbol{\omega}}_k^{\text{nav}}\right)\equiv\right]^T \end{bmatrix} \tag{145}$$

$\boldsymbol{\Sigma}_{\omega_k^{g_2}\omega_k^{g_2}}$ is found in a similar manner; this time by making use of the fact that $\boldsymbol{\nu}_k^{\omega_1}$ is independent of $\boldsymbol{\nu}_k^{\omega_2}$.

$$\boldsymbol{\Sigma}_{\omega_k^{g_2}\omega_k^{g_2}} = \mathbb{E}\left\{\left(\boldsymbol{\omega}_k^{g_2} - \mathbb{E}\left\{\boldsymbol{\omega}_k^{g_2}\right\}\right)\left(\boldsymbol{\omega}_k^{g_2} - \mathbb{E}\left\{\boldsymbol{\omega}_k^{g_2}\right\}\right)^T\right\} \tag{146}$$

$$= \mathbb{E}\left\{\left(H\delta\boldsymbol{x}_{k-1} - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1} + \boldsymbol{\nu}_k^{\omega_2}\right)\left(H\delta\boldsymbol{x}_{k-1} - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1} + \boldsymbol{\nu}_k^{\omega_2}\right)^T\right\} \tag{147}$$

$$= H\hat{\boldsymbol{\Sigma}}_{k-1}\boldsymbol{H}^T + \boldsymbol{T}_{g_1}^{g_2}\,\mathbb{E}\left\{\boldsymbol{\nu}_k^{\omega_1}\boldsymbol{\nu}_k^{\omega_1 T}\right\}\boldsymbol{T}_{g_2}^{g_1} + \mathbb{E}\left\{\boldsymbol{\nu}_k^{\omega_2}\boldsymbol{\nu}_k^{\omega_2 T}\right\} \tag{148}$$

The covariance terms $\mathbb{E}\left\{\boldsymbol{\nu}_k^{\omega_i}\boldsymbol{\nu}_k^{\omega_i T}\right\}$ is given by Equation (109).

Since $\delta\bar{\boldsymbol{x}}_k$ is always zero it can be neglected from the calculation of the delta state update.

$$\delta\hat{\boldsymbol{x}}_k = \boldsymbol{K}\left(\boldsymbol{\omega}_k^{g_2} - \bar{\boldsymbol{\omega}}_k^{g_2}\right) \tag{149}$$

$$\hat{\boldsymbol{\Sigma}}_k = \left(\boldsymbol{I}_{27\times27} - \boldsymbol{K}\boldsymbol{H}\right)\bar{\boldsymbol{\Sigma}}_k \tag{150}$$

Finally, the delta state is used to update the nominal state and then reset to zero.

$$\boldsymbol{x}_k^{\text{nom}} \leftarrow \delta\hat{\boldsymbol{x}}_k \boxplus \boldsymbol{x}_k^{\text{nom}} \tag{151}$$

$$\delta\hat{\boldsymbol{x}}_k \leftarrow \mathbf{0}_{27\times1} \tag{152}$$

1. Observability

One possible alternative to this filter design is to average gyro measurements and use this average in a model-replacement framework. Consider a simple noiseless measurement of angular rate for a single axis of rotation system.

$$\omega_k^{g_i} = \omega_k^{\text{nav}} + b_k^{g_i} \tag{153}$$

In this simplified example it is assumed that each of the two scalar gyro measurements are free from scale factor or misalignment errors which enter nonlinearly and complicate analysis. The state contains only angular displacement and

bias $\boldsymbol{x} = [\theta,\ b^{g_1},\ b^{g_2}]^T$. The biases are constant in time.

$$b_k^{g_i} = b_{k-1}^{g_i} \tag{154}$$

There are two redundant measurements available $\omega_k^{g_1}$ and $\omega_k^{g_2}$. They can be averaged to produce a synthetic measurement $\omega_k^{g_s}$.

$$\omega_k^{g_s} = \frac{1}{2}\omega_k^{g_1} + \frac{1}{2}\omega_k^{g_2} \tag{155}$$

The model propagation can be written as a function of $\omega_k^{g_s}$.

$$\boldsymbol{x}_k = \left(\boldsymbol{I}_{3\times 3} + \boldsymbol{T}_k^s \boldsymbol{L}^s\right)\boldsymbol{x}_{k-1} \tag{156}$$

where

$$\boldsymbol{T}_k^s = \begin{bmatrix} \Delta t \omega_k^{g_s} \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{L}^s = \begin{bmatrix} 0 \\ -1/2 \\ -1/2 \end{bmatrix}^T . \tag{157}$$

Noiseless measurements of angular displacement are also available.

$$y_k^s = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}\boldsymbol{x}_k \tag{158}$$

$$= \boldsymbol{H}^s \boldsymbol{x}_k \tag{159}$$

A linear observability matrix can be constructed by defining the state transition matrix $\boldsymbol{\Phi}_k^s = \boldsymbol{I}_{3\times 3} + \boldsymbol{T}_k^s \boldsymbol{L}^s$.

$$\boldsymbol{O}^s = \begin{bmatrix} \boldsymbol{H}^s \\ \boldsymbol{\Phi}_k^s \boldsymbol{H}^s \\ \boldsymbol{\Phi}_k^s \boldsymbol{\Phi}_k^s \boldsymbol{H}^s \end{bmatrix} \tag{160}$$

$\boldsymbol{O}^s$ has rank two and so the system is unobservable when $\omega_k^{g_s}$ is used for model propagation.

If our methodology is followed and $\omega_k^{g_1}$ is used for propagation the time evolution of $\boldsymbol{x}$ becomes

$$\boldsymbol{x} = (\boldsymbol{I}_{3\times 3} + \boldsymbol{T}_k \boldsymbol{L})\,\boldsymbol{x}_{k-1} \tag{161}$$

where

$$\boldsymbol{T}_k = \begin{bmatrix} \Delta t \omega_k^{g_1} \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{L} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}^T . \tag{162}$$

The same measurement of angular displacement is stacked with $\omega_k^{g_2}$ to form a new $\boldsymbol{y}_k$.

$$\boldsymbol{y}_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix} \boldsymbol{x}_k \tag{163}$$

$$= \boldsymbol{H}\boldsymbol{x}_k \tag{164}$$

Again, a linear observability matrix can be constructed by defining the state transition matrix $\boldsymbol{\Phi}_k = \boldsymbol{I}_{3\times 3} + \boldsymbol{T}_k \boldsymbol{L}$.

$$\boldsymbol{O} = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{\Phi}_k \boldsymbol{H} \\ \boldsymbol{\Phi}_k \boldsymbol{\Phi}_k \boldsymbol{H} \end{bmatrix} \tag{165}$$

In this case the rank of $\boldsymbol{O}$ is three and so the system is observable . This highlights a key advantage of our estimation scheme over measurement averaging; the individual biases of each gyro become observable. The individual scale factors and misalignment terms also become observable when persistence of excitation conditions are satisfied [25].

So far, this section has considered purely deterministic systems while the work as a whole is interested in stochastic ones. Observability of deterministic systems implies that estimate error will asymptotically collapse to zero as measurements are received. In the presence of process and measurement noise estimate error no longer collapses to zero. We assume that these noise terms are small and estimate errors will not grow unbounded with time.

E. Extension to Three Gyros

The filter may be extended to three gyros. The first gyro is used for model replacement and the remaining two are used in a single Kalman update. The linearized measurement models in the three gyro case are shown in Equations (166) and (167). These measurement models introduce additional correlations between the second and third gyro states but provide FDI benefits as discussed in a later section.

$$\boldsymbol{\omega}_k^{g_2} = \frac{1}{2}\left(\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1} - \boldsymbol{T}_{g_1}^{g_2}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_1} - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{b}_{k-1}^{g_1} - \boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1}\right)$$

$$+\frac{1}{2}\left(\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\omega}_k^{g_3}-\boldsymbol{T}_{g_3}^{g_2}\left[\left(\boldsymbol{\omega}_k^{g_3}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_3}-\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{b}_{k-1}^{g_3}-\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\nu}_k^{\omega_3}\right)$$
$$+\left[\left(\frac{1}{2}\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}^{g_1}+\frac{1}{2}\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\omega}^{g_3}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_2}+\boldsymbol{b}_{k-1}^{g_2}+\boldsymbol{\nu}_k^{\omega_2} \tag{166}$$

and

$$\boldsymbol{\omega}_k^{g_3}=\frac{1}{2}\left(\boldsymbol{T}_{g_1}^{g_3}\boldsymbol{\omega}_k^{g_1}-\boldsymbol{T}_{g_1}^{g_3}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_1}-\boldsymbol{T}_{g_1}^{g_3}\boldsymbol{b}_{k-1}^{g_1}-\boldsymbol{T}_{g_1}^{g_3}\boldsymbol{\nu}_k^{\omega_1}\right)$$
$$+\frac{1}{2}\left(\boldsymbol{T}_{g_2}^{g_3}\boldsymbol{\omega}_k^{g_2}-\boldsymbol{T}_{g_2}^{g_3}\left[\left(\boldsymbol{\omega}_k^{g_2}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_2}-\boldsymbol{T}_{g_2}^{g_3}\boldsymbol{b}_{k-1}^{g_2}-\boldsymbol{T}_{g_2}^{g_3}\boldsymbol{\nu}_k^{\omega_2}\right)$$
$$+\left[\left(\frac{1}{2}\boldsymbol{T}_{g_1}^{g_3}\boldsymbol{\omega}^{g_1}+\frac{1}{2}\boldsymbol{T}_{g_2}^{g_3}\boldsymbol{\omega}^{g_2}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_2}+\boldsymbol{b}_{k-1}^{g_2}+\boldsymbol{\nu}_k^{\omega_2} \tag{167}$$

Next, the Jacobians $\boldsymbol{H}^{g_2}=\frac{\partial\boldsymbol{\omega}_k^{g_2}}{\partial\delta\boldsymbol{x}_{k-1}}$ and $\boldsymbol{H}^{g_3}=\frac{\partial\boldsymbol{\omega}_k^{g_3}}{\partial\delta\boldsymbol{x}_{k-1}}$ are defined.

$$\boldsymbol{H}^{g_2}=\begin{bmatrix}\boldsymbol{0}_{3\times3}\\-\frac{1}{2}\boldsymbol{T}_{g_1}^{g_2}\\\boldsymbol{I}_{3\times3}\\-\frac{1}{2}\boldsymbol{T}_{g_3}^{g_2}\\-\frac{1}{2}\boldsymbol{T}_{g_1}^{g_2}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right]\\\frac{1}{2}\left[\left(\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1}+\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\omega}_k^{g_3}\right)\equiv\right]\\-\frac{1}{2}\boldsymbol{T}_{g_3}^{g_2}\left[\left(\boldsymbol{\omega}_k^{g_3}\right)\equiv\right]\end{bmatrix}^T\qquad\boldsymbol{H}^{g_3}=\begin{bmatrix}\boldsymbol{0}_{3\times3}\\-\frac{1}{2}\boldsymbol{T}_{g_1}^{g_3}\\-\frac{1}{2}\boldsymbol{T}_{g_2}^{g_3}\\\boldsymbol{I}_{3\times3}\\-\frac{1}{2}\boldsymbol{T}_{g_1}^{g_3}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right]\\-\frac{1}{2}\boldsymbol{T}_{g_2}^{g_3}\left[\left(\boldsymbol{\omega}_k^{g_2}\right)\equiv\right]\\\frac{1}{2}\left[\left(\boldsymbol{T}_{g_1}^{g_3}\boldsymbol{\omega}_k^{g_1}+\boldsymbol{T}_{g_2}^{g_3}\boldsymbol{\omega}_k^{g_2}\right)\equiv\right]\end{bmatrix}^T \tag{168}$$

These are stacked to form a composite jacobian.

$$\boldsymbol{H}=\begin{bmatrix}\boldsymbol{H}^{g_2}\\\boldsymbol{H}^{g_3}\end{bmatrix} \tag{169}$$

The measurements from Equations (166) and (167) can be written as a function of $\boldsymbol{x}_{k-1}^{\text{nom}}$ and $\delta\boldsymbol{x}_{k-1}$.

$$\boldsymbol{\omega}_k^{g_2}=\frac{1}{2}\left(\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1}+\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\omega}_k^{g_3}\right)+\boldsymbol{H}^{g_2}\left(\boldsymbol{\Lambda}\boldsymbol{x}_{k-1}^{\text{nom}}+\delta\boldsymbol{x}_{k-1}\right)-\frac{1}{2}\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1}-\frac{1}{2}\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\nu}_k^{\omega_3}+\boldsymbol{\nu}_k^{\omega_2} \tag{170}$$
$$\boldsymbol{\omega}_k^{g_3}=\frac{1}{2}\left(\boldsymbol{T}_{g_1}^{g_3}\boldsymbol{\omega}_k^{g_1}+\boldsymbol{T}_{g_2}^{g_3}\boldsymbol{\omega}_k^{g_2}\right)+\boldsymbol{H}^{g_3}\left(\boldsymbol{\Lambda}\boldsymbol{x}_{k-1}^{\text{nom}}+\delta\boldsymbol{x}_{k-1}\right)-\frac{1}{2}\boldsymbol{T}_{g_1}^{g_3}\boldsymbol{\nu}_k^{\omega_1}-\frac{1}{2}\boldsymbol{T}_{g_2}^{g_3}\boldsymbol{\nu}_k^{\omega_2}+\boldsymbol{\nu}_k^{\omega_3} \tag{171}$$

The Kalman update requires the following set of covariances which are straightforward to derive.

$$\text{cov}\left(\boldsymbol{\omega}_k^{g_2},\,\boldsymbol{\omega}_k^{g_2}\right)=\boldsymbol{H}^{g_2}\hat{\boldsymbol{\Sigma}}_{k-1}(\boldsymbol{H}^{g_2})^T+\frac{1}{4}\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{R}_k^{\omega_1}\boldsymbol{T}_{g_2}^{g_1}+\frac{1}{4}\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{R}_k^{\omega_3}\boldsymbol{T}_{g_2}^{g_3}+\boldsymbol{R}_k^{\omega_2} \tag{172}$$

$$\text{cov}\left(\boldsymbol{\omega}_k^{g_3},\, \boldsymbol{\omega}_k^{g_3}\right) = \boldsymbol{H}^{g_3}\hat{\boldsymbol{\Sigma}}_{k-1}(\boldsymbol{H}^{g_3})^T + \frac{1}{4}\boldsymbol{T}_{g_1}^{g_3}\boldsymbol{R}_k^{\omega_1}\boldsymbol{T}_{g_3}^{g_1} + \frac{1}{4}\boldsymbol{T}_{g_2}^{g_3}\boldsymbol{R}_k^{\omega_2}\boldsymbol{T}_{g_3}^{g_2} + \boldsymbol{R}_k^{\omega_3} \tag{173}$$

$$\text{cov}\left(\boldsymbol{\omega}_k^{g_2},\, \boldsymbol{\omega}_k^{g_3}\right) = \boldsymbol{H}^{g_2}\hat{\boldsymbol{\Sigma}}_{k-1}(\boldsymbol{H}^{g_3})^T + \frac{1}{4}\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{R}_k^{\omega_1}\boldsymbol{T}_{g_3}^{g_1} - \frac{1}{2}\boldsymbol{R}_k^{\omega_2}\boldsymbol{T}_{g_3}^{g_2} - \frac{1}{2}\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{R}_k^{\omega_3} \tag{174}$$

Finding $\text{cov}\left(\delta\boldsymbol{x}_k,\, \boldsymbol{\omega}_k^{g_2}\right)$ is more involved. The delta state evolves as

$$\delta\boldsymbol{x}_k = \boldsymbol{\Phi}(\Delta t)\delta\boldsymbol{x}_{k-1} + \boldsymbol{\nu}_{k-1}^x. \tag{175}$$

As an intermediate step, the residual of $\boldsymbol{\omega}_k^{g_2}$ from its expectation is defined as $\boldsymbol{r}_k^{g_2}$. Some manipulations on $\boldsymbol{r}_k^{g_2}$ will be performed.

$$\boldsymbol{r}_k^{g_2} = \boldsymbol{\omega}_k^{g_2} - \mathbb{E}\left\{\boldsymbol{\omega}_k^{g_2}\right\} \tag{176}$$

$$\boldsymbol{r}_k^{g_2} = \frac{1}{2}\left(\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1} + \boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\omega}_k^{g_3}\right) + \boldsymbol{H}^{g_2}\left(\boldsymbol{\Lambda}\boldsymbol{x}_{k-1}^{\text{nom}} + \delta\boldsymbol{x}_{k-1}\right) - \frac{1}{2}\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1} - \frac{1}{2}\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\nu}_k^{\omega_3} + \boldsymbol{\nu}_k^{\omega_2} - \frac{1}{2}\left(\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\omega}_k^{g_1} + \boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\omega}_k^{g_3}\right) - \boldsymbol{H}^{g_2}\boldsymbol{\Lambda}\boldsymbol{x}_{k-1}^{\text{nom}} \tag{177}$$

$$\boldsymbol{r}_k^{g_2} = \boldsymbol{H}^{g_2}\delta\boldsymbol{x}_{k-1} - \frac{1}{2}\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1} - \frac{1}{2}\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\nu}_k^{\omega_3} + \boldsymbol{\nu}_k^{\omega_2} \tag{178}$$

This term is substituted into the covariance formulation.

$$\text{cov}\left(\delta\boldsymbol{x}_k,\, \boldsymbol{\omega}_k^{g_2}\right) = \mathbb{E}\left\{\left(\delta\boldsymbol{x}_k - \mathbb{E}\left\{\delta\boldsymbol{x}_k\right\}\right)\left(\boldsymbol{r}_k^{g_2}\right)^T\right\} \tag{179}$$

$$= \mathbb{E}\left\{\left(\boldsymbol{\Phi}(\Delta t)\delta\boldsymbol{x}_{k-1} + \boldsymbol{\nu}_{k-1}^x\right)\left(\boldsymbol{H}^{g_2}\delta\boldsymbol{x}_{k-1} - \frac{1}{2}\boldsymbol{T}_{g_1}^{g_2}\boldsymbol{\nu}_k^{\omega_1} - \frac{1}{2}\boldsymbol{T}_{g_3}^{g_2}\boldsymbol{\nu}_k^{\omega_3} + \boldsymbol{\nu}_k^{\omega_2}\right)^T\right\} \tag{180}$$

$$= \boldsymbol{\Phi}(\Delta t)\hat{\boldsymbol{\Sigma}}_{k-1}(\boldsymbol{H}^{g_2})^T - \frac{1}{2}\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^x\boldsymbol{\nu}_k^{\omega_1 T}\right\}\boldsymbol{T}_{g_2}^{g_1} - \frac{1}{2}\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^x\boldsymbol{\nu}_k^{\omega_3 T}\right\}\boldsymbol{T}_{g_2}^{g_3} + \mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^x\boldsymbol{\nu}_k^{\omega_2 T}\right\} \tag{181}$$

Similarly,

$$\text{cov}\left(\delta\boldsymbol{x}_k,\, \boldsymbol{\omega}_k^{g_3}\right) = \boldsymbol{\Phi}(\Delta t)\hat{\boldsymbol{\Sigma}}_{k-1}(\boldsymbol{H}^{g_3})^T - \frac{1}{2}\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^x\boldsymbol{\nu}_k^{\omega_1 T}\right\}\boldsymbol{T}_{g_3}^{g_1} - \frac{1}{2}\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^x\boldsymbol{\nu}_k^{\omega_2 T}\right\}\boldsymbol{T}_{g_3}^{g_2} + \mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^x\boldsymbol{\nu}_k^{\omega_3 T}\right\}. \tag{182}$$

The correlations between $\boldsymbol{\nu}_{k-1}^x$ and each $\boldsymbol{\nu}_k^{\omega_i}$ are similar in form to the two gyro case (Equations (144) and (145))

and are found in a similar manner. They are given by:

$$\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{x}\boldsymbol{\nu}_{k}^{\omega_1 T}\right\} = \begin{bmatrix} -\boldsymbol{T}_{g_1}^{\text{nav}}\left(\left(\sigma_{\omega}^{2}+\frac{\Delta t^2\sigma_b^2}{6}\right)\boldsymbol{I}_{3\times 3}+\frac{\Delta t^2\sigma_{\Delta}^2}{6}\left[\left(\boldsymbol{\omega}_k^{g_1}\right)\equiv\right]\left[\left(\boldsymbol{T}_{\text{nav}}^{g_1}\bar{\boldsymbol{\omega}}_k^{\text{nav}}\right)\equiv\right]^T\right) \\ \frac{\Delta t}{2}\sigma_b^2\boldsymbol{I}_{3\times 3} \\ \boldsymbol{0}_{3\times 3} \\ \boldsymbol{0}_{3\times 3} \\ \frac{\Delta t}{2}\sigma_{\Delta}^2\left[\left(\boldsymbol{T}_{\text{nav}}^{g_1}\bar{\boldsymbol{\omega}}_k^{\text{nav}}\right)\equiv\right]^T \\ \boldsymbol{0}_{9\times 3}\boldsymbol{0}_{9\times 3} \end{bmatrix} \tag{183}$$

$$\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{x}\boldsymbol{\nu}_{k}^{\omega_2 T}\right\} = \begin{bmatrix} \boldsymbol{0}_{3\times 3} \\ \boldsymbol{0}_{3\times 3} \\ \frac{1}{2}\Delta t\sigma_b^2\boldsymbol{I}_{3\times 3} \\ \boldsymbol{0}_{3\times 3} \\ \boldsymbol{0}_{9\times 3} \\ \frac{\Delta t}{2}\sigma_{\Delta}^2\left[\left(\boldsymbol{T}_{\text{nav}}^{g_2}\bar{\boldsymbol{\omega}}_k^{\text{nav}}\right)\equiv\right]^T \\ \boldsymbol{0}_{9\times 3} \end{bmatrix} \tag{184}$$

$$\mathbb{E}\left\{\boldsymbol{\nu}_{k-1}^{x}\boldsymbol{\nu}_{k}^{\omega_3 T}\right\} = \begin{bmatrix} \boldsymbol{0}_{3\times 3} \\ \boldsymbol{0}_{3\times 3} \\ \boldsymbol{0}_{3\times 3} \\ \frac{1}{2}\Delta t\sigma_b^2\boldsymbol{I}_{3\times 3} \\ \boldsymbol{0}_{9\times 3} \\ \boldsymbol{0}_{9\times 3} \\ \frac{\Delta t}{2}\sigma_{\Delta}^2\left[\left(\boldsymbol{T}_{\text{nav}}^{g_3}\bar{\boldsymbol{\omega}}_k^{\text{nav}}\right)\equiv\right]^T \end{bmatrix} \tag{185}$$

The measurements from gyros two and three are stacked in a single vector $\boldsymbol{y}_k$ to be processed simultaneously.

$$\boldsymbol{y}_k = \begin{bmatrix} \boldsymbol{\omega}_k^{g_2} \\ \boldsymbol{\omega}_k^{g_3} \end{bmatrix} \tag{186}$$

The Kalman gain $\boldsymbol{K}$ is found as

$$\boldsymbol{K} = \boldsymbol{\Sigma}_{\delta \boldsymbol{x}_k \boldsymbol{y}_k} \boldsymbol{\Sigma}_{\boldsymbol{y}_k \boldsymbol{y}_k}^{-1} \tag{187}$$

with

$$\boldsymbol{\Sigma}_{\delta \boldsymbol{x}_k \boldsymbol{y}_k} = \begin{bmatrix} \text{cov}\left(\delta \boldsymbol{x}_k,\, \boldsymbol{\omega}_k^{g_2}\right) & \text{cov}\left(\delta \boldsymbol{x}_k,\, \boldsymbol{\omega}_k^{g_3}\right) \end{bmatrix} \tag{188}$$

and

$$\boldsymbol{\Sigma}_{\delta \boldsymbol{x}_k \boldsymbol{y}_k} = \begin{bmatrix} \text{cov}\left(\boldsymbol{\omega}_k^{g_2},\, \boldsymbol{\omega}_k^{g_2}\right) & \text{cov}\left(\boldsymbol{\omega}_k^{g_2},\, \boldsymbol{\omega}_k^{g_3}\right) \\ \text{cov}\left(\boldsymbol{\omega}_k^{g_2},\, \boldsymbol{\omega}_k^{g_3}\right)^T & \text{cov}\left(\boldsymbol{\omega}_k^{g_3},\, \boldsymbol{\omega}_k^{g_3}\right) \end{bmatrix}. \tag{189}$$

As in the two gyro case, the delta state estimate is updated as:

$$\delta \hat{\boldsymbol{x}}_k = \boldsymbol{K}\left(\boldsymbol{y}_k - \bar{\boldsymbol{y}}_k\right) \tag{190}$$

$$\hat{\boldsymbol{\Sigma}}_k = \left(\boldsymbol{I}_{39\times 39} - \boldsymbol{K}\boldsymbol{H}\right)\bar{\boldsymbol{\Sigma}}_k \tag{191}$$

Again, the delta state is used to update the nominal state and then reset to zero.

$$\boldsymbol{x}_k^{\text{nom}} \leftarrow \delta \hat{\boldsymbol{x}}_k \boxplus \boldsymbol{x}_k^{\text{nom}} \tag{192}$$

$$\delta \hat{\boldsymbol{x}}_k \leftarrow \boldsymbol{0}_{39\times 1} \tag{193}$$

F. Update with Star Tracker Measurements

The spacecraft also receives star tracker measurements of the vehicle's attitude with respect to an inertial frame. It is assumed that these measurements occur at a slower rate than gyro measurements but are resonant such that a set of gyro measurements always arrives at the same time as a star tracker measurement. In this work the star tracker measurements are processed sequentially after the gyros measurements. If desired, it is trivial to modify the algorithm to process gyro and star tracker measurements simultaneously in a single update.

Star tracker measurements take the form of a quaternion and are modeled as

$$\tilde{\boldsymbol{q}}_{\text{inert}}^{\text{nav}} = \delta \boldsymbol{q}(\boldsymbol{\nu}_k^{\text{st}}) \otimes \delta \boldsymbol{q}(\boldsymbol{\theta}_k) \otimes \left(\boldsymbol{q}_{\text{inert}}^{\text{nav}}\right)^{\text{nom}} \tag{194}$$

$\boldsymbol{\nu}_k^{\text{st}} \in \mathbb{R}^3$ is a discrete, zero-mean, Gaussian noise term with covariance $\boldsymbol{\Sigma}_{\text{st}}$. The star tracker measurements are post-processed to find the quantity $\boldsymbol{z}_k^{\text{st}}$ defined as

$$\boldsymbol{z}_k^{\text{st}} = \boldsymbol{\theta}\left(\tilde{\boldsymbol{q}}_{\text{inert}}^{\text{nav}} \otimes \left(\boldsymbol{q}_{\text{inert}}^{\text{nav}}\right)^{\text{nom}*}\right). \tag{195}$$

When the magnitude of $\boldsymbol{\nu}_k^{\text{st}}$ and $\boldsymbol{\theta}_k$ is small the following model approximately holds.

$$\boldsymbol{z}_k^{\text{st}} = \boldsymbol{\nu}_k^{\text{st}} + \boldsymbol{\theta}_k \tag{196}$$

Equation (196) is used to provide a standard Kalman update to the delta state.

The Jacobian is easily found as

$$\boldsymbol{H}^{\text{st}} = \begin{bmatrix} \boldsymbol{I}_{3\times3} & \boldsymbol{0}_{3\times24} \end{bmatrix} \tag{197}$$

With this, the Kalman gain is found as

$$\boldsymbol{K}^{\text{st}} = \left(\hat{\boldsymbol{\Sigma}}_k \boldsymbol{H}^{\text{st}T}\right)\left(\boldsymbol{H}^{\text{st}}\hat{\boldsymbol{\Sigma}}_k \boldsymbol{H}^{\text{st}T} + \boldsymbol{\Sigma}_{\text{st}}\right)^{-1} \tag{198}$$

An update is then found with the following equations which have taken into consideration that the delta state is zero before the measurement is processed.

$$\delta\check{\boldsymbol{x}}_k = \boldsymbol{K}^{\text{st}}\boldsymbol{z}_k^{\text{st}} \tag{199}$$

$$\check{\boldsymbol{\Sigma}}_k = \left(\boldsymbol{I}_{27\times27} - \boldsymbol{K}^{\text{st}}\boldsymbol{H}^{\text{st}}\right)\hat{\boldsymbol{\Sigma}}_k \tag{200}$$

Finally, two reset operations are performed.

$$\hat{\boldsymbol{\Sigma}}_k \leftarrow \check{\boldsymbol{\Sigma}}_k \tag{201}$$

$$\boldsymbol{x}_k^{\text{nom}} \leftarrow \delta\check{\boldsymbol{x}}_k \boxplus \boldsymbol{x}_k^{\text{nom}} \tag{202}$$

G. Gyro Fault Detection and Isolation

The advantage of the proposed filter is that biases and misalignments from different IMUs can all be estimated, together with their correlations. This information can be used for fault detection and isolation. Rather than treating them separately, the gyro FDI methodology is incorporated into the proposed MEKF. The methodology works on a component level so that faults within individual gyro axes can be detected and isolated. It is efficient in that it makes use of several quantities which have already been calculated in the process of updating the filter state. Namely the prefit measurement residuals for the second and third gyros along with their associated covariances. An additional prefit measurement residual for the first gyro is also required for FDI purposes only.

This is a notable deviation from traditional FDI methodologies which leverage the previously discussed parity space methods. In those methods, a linear transformation is applied to a raw measurement resulting in a fault vector which is only a function of measurement error. This transformation is described in Equation (69) for a three gyro system. This

work takes an alternate path and seeks a FDI methodology which is more tightly integrated with the estimator. This choice allows us to explicitly consider estimated gyro states and their uncertainty in the FDI scheme. Later, we will argue that our prefit measurement residuals have obvious parallels to the fault vector from Equation (69) and so all previous analysis for parity space FDI is equally valid for our residual FDI technique.

In this work FDI will be performed using only the state estimates and the gyroscope measurements. Alternative FDI methodologies could certainly be devised which use additional measurement types (such as star trackers) to detect faults within the gyroscopes. We aim to show that this additional complexity is unnecessary. Furthermore; such methods may not work for all phases of flight. For example, star tracker measurements are generally unavailable when a vehicle is under thrust or within a planet's atmosphere.

The measurement prefit residuals $\boldsymbol{r}_k^{g_i}$ for each each gyro are defined as

$$\boldsymbol{r}_k^{g_i} = \boldsymbol{\omega}_k^{g_i} - \mathbb{E}\left\{\boldsymbol{\omega}_k^{g_i}\right\}. \tag{203}$$

A new model for the first gyro measurement is required in addition to the previously derived models for the second and third gyro measurements in Equations (166) and (167).

$$\begin{aligned}\boldsymbol{\omega}_k^{g_1} = {} & \frac{1}{2}\boldsymbol{T}_{g_2}^{g_1}\left(\boldsymbol{\omega}_k^{g_2} - \left[\left(\boldsymbol{\omega}_k^{g_2}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_2} - \boldsymbol{b}_{k-1}^{g_2} - \boldsymbol{\nu}_k^{\omega_2}\right) + \frac{1}{2}\boldsymbol{T}_{g_3}^{g_1}\left(\boldsymbol{\omega}_k^{g_3} - \left[\left(\boldsymbol{\omega}_k^{g_3}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_3} - \boldsymbol{b}_{k-1}^{g_3} - \boldsymbol{\nu}_k^{\omega_3}\right) \\ & + \left[\left(\frac{1}{2}\boldsymbol{T}_{g_2}^{g_1}\boldsymbol{\omega}^{g_2} + \frac{1}{2}\boldsymbol{T}_{g_3}^{g_1}\boldsymbol{\omega}^{g_3}\right)\equiv\right]\boldsymbol{\Delta}_{k-1}^{g_1} + \boldsymbol{b}_{k-1}^{g_1} + \boldsymbol{\nu}_k^{\omega_1}\end{aligned} \tag{204}$$

From this model a third residual can be formed. Stacking the three residuals together results in the scaled fault vector of Eq. (69), with the additional benefit that the measurements are compensated with the estimates of biases and misalignments for more accurate FDI. The stacked residuals will be defined more explicitly later in Equation (208).

The covariance of each residual is also required. Note that the covariance of each residual is equivalent to the covariance of the corresponding measurement. Therefore, covariances for residuals two and three are available in Equations (172) and (173). The covariance for the first residual may be found in a similar manner to the second and

third by first defining $\boldsymbol{H}^{g_1}$.

$$\boldsymbol{H}^{g_1} = \begin{bmatrix} \boldsymbol{0}_{3\times3} \\ \boldsymbol{I}_{3\times3} \\ -\frac{1}{2}\boldsymbol{T}_{g_2}^{g_1} \\ -\frac{1}{2}\boldsymbol{T}_{g_3}^{g_1} \\ \frac{1}{2}\left[\left(\boldsymbol{T}_{g_2}^{g_1}\boldsymbol{\omega}_k^{g_2} + \boldsymbol{T}_{g_3}^{g_1}\boldsymbol{\omega}_k^{g_3}\right)\equiv\right] \\ -\frac{1}{2}\boldsymbol{T}_{g_3}^{g_1}\left[\left(\boldsymbol{\omega}_k^{g_2}\right)\equiv\right] \\ -\frac{1}{2}\boldsymbol{T}_{g_3}^{g_1}\left[\left(\boldsymbol{\omega}_k^{g_3}\right)\equiv\right] \end{bmatrix}^T \tag{205}$$

The covariance for $\boldsymbol{r}_k^{g_1}$ can be expressed with $\boldsymbol{H}^{g_1}$.

$$\operatorname{cov}\left(\boldsymbol{r}_k^{g_1}, \boldsymbol{r}_k^{g_1}\right) = \boldsymbol{H}^{g_1}\hat{\boldsymbol{\Sigma}}_{k-1}(\boldsymbol{H}^{g_1})^T + \frac{1}{4}\boldsymbol{T}_{g_2}^{g_1}\boldsymbol{R}_k^{\omega_2}\boldsymbol{T}_{g_1}^{g_2} + \frac{1}{4}\boldsymbol{T}_{g_3}^{g_1}\boldsymbol{R}_k^{\omega_3}\boldsymbol{T}_{g_1}^{g_3} + \boldsymbol{R}_k^{\omega_1} \tag{206}$$

To elucidate the comparison between our FDI methodology and the parity space methods; consider the error in each gyro's estimated angular rate as a function of the delta state and measurement noise.

$$\delta\boldsymbol{\omega}_k^{g_i} = -\left[\left(\boldsymbol{\omega}_k^{g_i}\right)\equiv\right]\delta\boldsymbol{\Delta}_{k-1}^{g_i} - \delta\boldsymbol{b}_{k-1}^{g_i} - \boldsymbol{\nu}_k^{\omega_i} \tag{207}$$

The estimated angular rate errors and measurement residuals can be stacked as

$$\boldsymbol{r}_k^{\text{all}} = \begin{bmatrix} \boldsymbol{r}_k^{g_1} \\ \boldsymbol{r}_k^{g_2} \\ \boldsymbol{r}_k^{g_3} \end{bmatrix} \qquad \delta\boldsymbol{\omega}^{\text{all}} = \begin{bmatrix} \delta\boldsymbol{\omega}_k^{g_1} \\ \delta\boldsymbol{\omega}_k^{g_2} \\ \delta\boldsymbol{\omega}_k^{g_3} \end{bmatrix} \tag{208}$$

There is a linear transformation between $\delta\boldsymbol{\omega}^{\text{all}}$ and $\boldsymbol{r}_k^{\text{all}}$

$$\boldsymbol{r}_k^{\text{all}} = \begin{bmatrix} \boldsymbol{I}_{3\times3} & -\frac{1}{2}\boldsymbol{T}_{g_2}^{g_1} & -\frac{1}{2}\boldsymbol{T}_{g_3}^{g_1} \\ -\frac{1}{2}\boldsymbol{T}_{g_2}^{g_1} & \boldsymbol{I}_{3\times3} & -\frac{1}{2}\boldsymbol{T}_{g_3}^{g_1} \\ -\frac{1}{2}\boldsymbol{T}_{g_2}^{g_1} & -\frac{1}{2}\boldsymbol{T}_{g_3}^{g_1} & \boldsymbol{I}_{3\times3} \end{bmatrix}\delta\boldsymbol{\omega}^{\text{all}} \tag{209}$$

This transformation is identical to the one described in Equation (69) modulo a factor of $2/3$. Conceptually, the fault vector $\boldsymbol{\epsilon}$ and measurement error $\boldsymbol{\eta}$ have been replaced with $\boldsymbol{r}_k^{\text{all}}$ and $\delta\boldsymbol{\omega}^{\text{all}}$, respectively. From this, we claim that the previous analysis on optimal orientation of IMUs for parity space FDI also applies to our residual FDI technique.

The primary metric of interest in detecting failures is the Mahalanobis distance of each element of each $\boldsymbol{r}_k^{g_i}$. This is a measure of distance between a point $\boldsymbol{p}$ and distribution specified by a mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{R}$. Its formulation is provided below.

$$d = \sqrt{(\boldsymbol{p} - \boldsymbol{\mu})^T \boldsymbol{R}^{-1} (\boldsymbol{p} - \boldsymbol{\mu})} \tag{210}$$

If $\boldsymbol{\mu}$ and $\boldsymbol{R}$ describe a multivariate Gaussian distribution from which $\boldsymbol{p}$ is sampled then d^2 is chi-square distributed with ϕ degrees of freedom where ϕ is equal to the dimension of the multivariate Gaussian.

For systems which are linear, Gaussian, and have known dynamics, each $\boldsymbol{r}_k^{g_i}$ will follow a zero-mean Gaussian distribution with the associated covariance from Equations (172), (173), or (206). If these assumptions hold and no faults are present the squared Mahalanobis distance $\left(d_k^{g_i}(j)\right)^2$ for the j^{th} element of $\boldsymbol{r}_k^{g_i}$ will follow a chi-squared distribution with $\phi = 1$ degrees of freedom. Many samples of $\left(d_k^{g_i}(j)\right)^2$ will be collected and gyro failure will be declared if the samples are unlikely to come from a chi-squared distribution.

$$\left(d_k^{g_i}(j)\right)^2 = \frac{\left(\boldsymbol{r}_k^{g_i}(j)\right)^2}{\text{cov}\left(\boldsymbol{r}_k^{g_i}, \boldsymbol{r}_k^{g_i}\right)(j,\, j)} \tag{211}$$

Here $\text{cov}\left(\boldsymbol{r}_k^{g_i}, \boldsymbol{r}_k^{g_i}\right)(j,\, j)$ is the j_{th} diagonal element of $\text{cov}\left(\boldsymbol{r}_k^{g_i}, \boldsymbol{r}_k^{g_i}\right)$. It would also be possible to check that the residuals themselves follow a Gaussian distribution but that distribution's covariance varies with time. That makes it unsuitable for this work's fault detection test as it requires an independent and identically distributed batch of samples. The chi-squared distribution does not have this issue since ϕ is constant.

Another possibility would be to check if the Mahalanobis distance of the entire three dimension residual follows a chi-squared distribution with $\phi = 3$. This would allow the test to explicitly consider correlations between residual elements. However, once a fault was detected a second test would be required for isolation. Testing each component individually allows us to detect and isolate faults with a single set of tests.

To check if the chi-squared distribution is an accurate model a Cramer-von Mises Goodness of Fit Test [26, 27] (CVMT) is performed. The CVMT is a hypothesis test comparing a proposed cumulative distribution to an empirical cumulative distribution. To this end, the N_{CVMT} most recent $\left(d_k^{g_i}(j)\right)^2$ samples from each updating gyro are stored and assumed to be independent samples from an unknown distribution. In reality, the samples are correlated through time and are not truly independent unless the system is linear and free from modeling error.

CVMT test statistics $W_{g^i}^2(j)$ for each one of the gyro axes are found with the following [26].

$$W_{g_i}^2(j) = \frac{1}{12N_{\text{CVMT}}} + \sum_{k=1}^{N_{\text{CVMT}}} \left(\chi_c^2\left(\left(d_k^{g_i}(j)\right)^2,\, 1\right) - \frac{2k-1}{2N_{\text{CVMT}}}\right)^2 \tag{212}$$

Here, $\chi_c^2(z,\ \phi)$ is the chi-square cumulative distribution function evaluated at z and index k no longer represents time but rather represents the position of $(d_k^{g_i}(j))^2$ in a list of the N_{CVMT} most recent samples sorted in ascending order. If the distribution is correct $W_{g_i}^2(j)$ approaches 0 as the number of samples tends towards infinity.

If the number of samples is less than infinity but the distribution is still correct $W_{g_i}^2$ is a non-negative random variable with a known cumulative distribution function $W_c^2(z,\ N_{\text{CVMT}})$ [28]. From this a hypothesis test can be defined with null hypothesis that the $\left(d_k^{g_i}(j)\right)^2$ samples come from a chi-square distribution with $\phi = 1$. A detection threshold W_{fail}^2 is found by inverting the cumulative distribution function to solve the following equation.

$$1 - \alpha = W_c^2(W_{\text{fail}}^2,\ N_{\text{CVMT}}). \tag{213}$$

Whenever $W_{g_i}^2(j) > W_{\text{fail}}^2$ the null hypothesis is rejected and it is possible a gyro has failed with a probability of false alarm α. This work uses $\alpha = 0.01$. The inversion is numerically troublesome so T_{fail} is found via interpolation of the tables published by Csorgo and Faraway [28]. Over extended periods of time a false alarm is almost certain to occur. To avoid failing a gyro due to a false alarm gyro failure is declared only when $W_{g_i}^2(j) > W_{\text{fail}}^2$ consistently for a period longer than T_{fail} seconds.

Determining which gyro has failed is straightforward. If one of the tests exceeds the threshold for longer than T_{fail} a fault is detected. The fault is isolated as belonging to the gyro component with the largest $W_{g_i}^2(j)$. This choice is motivated by the fact that faults in one axis can cause non chi-squared behavior in other axes due to the filter fusion. It is theoretically possible but practically unlikely that the axis which fails first does not display the largest fault signature.

As mentioned in a previous section this FDI framework benefits heavily from the measurement models introduced in Equations (166), (167), and (204) and an optimal relative gyro alignment. Notably the expectation of the measurement in a given gyro is the average of the measurements in the other two gyros. This averaging improves a given fault signature's signal to noise ratio and helps to ensure that fault signatures are largest in their corresponding $\left(d_k^{g_i}(j)\right)^2$. An optimal alignment serves to further improve signal to noise ratio. Therefore, these measurement models combined with an optimal gyro alignment improve the FDI system's chances of correctly isolating the fault.

H. Monte Carlo Testing

A 1000 run Monte Carlo test validates performance of the proposed filter using measurements from two gyros running at Δt^{g_i}. As a baseline comparison a filter which averages measurements from the two gyros is used. The averaged measurement is assumed to have come from a synthetic gyro which is aligned with the navigation frame. The synthetic measurements $\omega_k^{g_s}$ will be used for filter propagation only. Both filters will use measurements from a star tracker running at Δt^{st} of spacecraft attitude with respect to an inertial frame. Inspection of the synthetic measurements

shows that the comparison filter can estimate a synthetic bias $\boldsymbol{b}_k^{g_s}$ and distortion vector $\boldsymbol{\Delta}_k^{g_s}$.

$$\boldsymbol{\omega}_k^{g_s} = \frac{1}{2}\sum_{i=1}^{2} \boldsymbol{T}_{g_i}^{\text{nav}} \boldsymbol{\omega}_k^{g_i} \tag{214}$$

$$= \frac{1}{2}\sum_{i=1}^{2} \boldsymbol{T}_{g_i}^{\text{nav}} \left(\boldsymbol{T}_{\text{nav}}^{g_i} \boldsymbol{\omega}_k^{\text{nav}} + M_{k-1}^{\Delta_i} \boldsymbol{T}_{\text{nav}}^{g_i} \boldsymbol{\omega}_k^{\text{nav}} + \boldsymbol{b}_{k-1}^{g_i} + \boldsymbol{\nu}_k^{\omega_i} \right) \tag{215}$$

$$= \boldsymbol{\omega}_k^{\text{nav}} + \frac{1}{2}\sum_{i=1}^{2} \boldsymbol{T}_{g_i}^{\text{nav}} M_{k-1}^{\Delta_i} \boldsymbol{T}_{\text{nav}}^{g_i} \boldsymbol{\omega}_k^{\text{nav}} + \frac{1}{2}\sum_{i=1}^{2} \boldsymbol{T}_{g_i}^{\text{nav}} \boldsymbol{b}_{k-1}^{g_i} + \frac{1}{2}\sum_{i=1}^{2} \boldsymbol{T}_{g_i}^{\text{nav}} \boldsymbol{\nu}_k^{\omega_i} \tag{216}$$

$$= \boldsymbol{\omega}_k^{\text{nav}} + M_{k-1}^{\Delta_s} \boldsymbol{\omega}_k^{\text{nav}} + \boldsymbol{b}_{k-1}^{g_s} + \boldsymbol{\nu}_k^{\omega_s} \tag{217}$$

If both gyros have the same properties it can be shown that the following noise characteristics hold for the synthetic gyro.

$$\mathbb{E}\left\{ \boldsymbol{\nu}_{k-1}^{b_s} \boldsymbol{\nu}_{k-1}^{b_s\,T} \right\} = \frac{1}{2}\mathbb{E}\left\{ \boldsymbol{\nu}_{k-1}^{b_1} \boldsymbol{\nu}_{k-1}^{b_1\,T} \right\} \tag{218}$$

$$\mathbb{E}\left\{ \boldsymbol{\nu}_{k-1}^{\Delta_s} \boldsymbol{\nu}_{k-1}^{\Delta_s\,T} \right\} = \frac{1}{2}\mathbb{E}\left\{ \boldsymbol{\nu}_{k-1}^{\Delta_1} \boldsymbol{\nu}_{k-1}^{\Delta_1\,T} \right\} \tag{219}$$

$$\mathbb{E}\left\{ \boldsymbol{\nu}_{k-1}^{\omega_s} \boldsymbol{\nu}_{k-1}^{\omega_s\,T} \right\} = \frac{1}{2}\mathbb{E}\left\{ \boldsymbol{\nu}_{k-1}^{\omega_1} \boldsymbol{\nu}_{k-1}^{\omega_1\,T} \right\} \tag{220}$$

Each gyro has identical noise characteristics. The gyro angular random walk ARW_g, bias instability BI_g, scale factor instability SFI_g, misalignment instability MI_g, and star tracker noise characteristics are shown in Table 1.

Table 1 Sensor noise parameters. Note that ppm is shorthand for parts per million.

ARW_g	$0.344\,\text{deg}/\sqrt{\text{h}}$
BI_g	$0.206\,\text{deg/h}$
SFI_g	$360\,\text{ppm/h}$
MI_g	$0.0206\,\text{deg/h}$
$\boldsymbol{\Sigma}_{\text{st}}$	$36^2\,\boldsymbol{I}_{3\times3}\,\text{arcsec}^2$

To ensure observability of all states the vehicle begins at rest and is subject to sinusoidal torque inputs in the body frame with magnitude T_{input} and frequency f_T^j. The filters have no knowledge of these inputs or the vehicle's inertia tensor $\boldsymbol{J}$. The two gyros are nominally misaligned with each other and the navigation frame. The star tracker's reference frame is aligned with the vehicle's navigation frame. The vehicle's properties and input torque characteristics are shown in Table 2. For brevity transformation matrices are expressed in terms of their equivalent scalar-last quaternion representations.

Each filter begins with the same initial estimate. Covariances for each estimated state are provided in Table 3. The initial uncertainty in distortion states is separated into $\hat{\boldsymbol{\Sigma}}_{\Delta_s}$ and $\hat{\boldsymbol{\Sigma}}_{\Delta_\delta}$ representing scale factor and misalignment initial uncertainty, respectively.

Table 2 Physical vehicle properties.

$\boldsymbol{J}$	$\text{blkdiag}\{2,\ 4,\ 1\}\ \text{kg m}^2$
$\boldsymbol{q}_{\text{nav}}^{g_1}$	$\begin{bmatrix}0.98 & 0.06 & -0.20 & 0.06\end{bmatrix}$
$\boldsymbol{q}_{\text{nav}}^{g_2}$	$\begin{bmatrix}0.89 & -0.33 & 0.11 & 0.29\end{bmatrix}$
Δt^{g_i}	0.01 s
Δt^{st}	1.0 s
T_{input}	0.05 N m
$\begin{bmatrix}f_T^x & f_T^y & f_T^z\end{bmatrix}$	$\begin{bmatrix}10.0 & 15.35 & 18.12\end{bmatrix}$ Hz

Table 3 Initial estimate uncertainty.

$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}$	$0.1^2\boldsymbol{I}_{3\times3}\ \text{deg}^2$
$\hat{\boldsymbol{\Sigma}}_{b}$	$100^2\ \boldsymbol{I}_{3\times3}\ \text{deg}^2/\text{h}^2$
$\hat{\boldsymbol{\Sigma}}_{\Delta_s}$	$1700^2\ \boldsymbol{I}_{3\times3}\ \text{ppm}^2$
$\hat{\boldsymbol{\Sigma}}_{\Delta_\delta}$	$0.1^2\ \boldsymbol{I}_{6\times6}\ \text{deg}^2$

Figure 1 compares attitude estimation performance for each filter. Both filters provide consistent and unbiased estimates of attitude with little difference in their performance. This means that the multi-gyro filter's orange plots closely overlap with all green plots for the filter with a synthetic gyro.

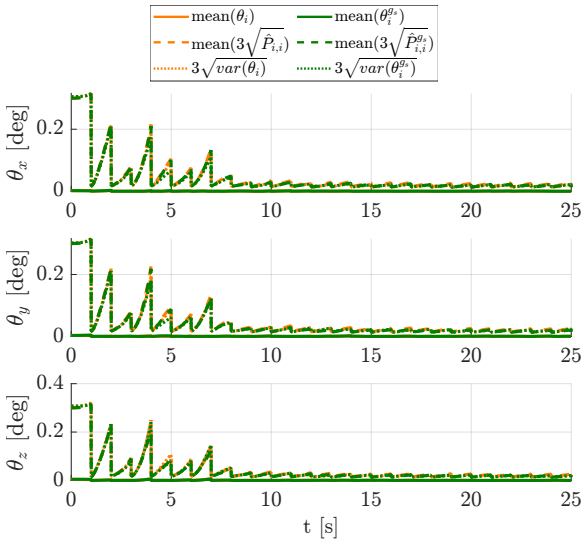

Fig. 1 The multi-gyro filter's attitude estimation error performance compared with the synthetic gyro performance.

Figures 2 and 3 show the multi-gyro filter's estimate errors for the first gyro's bias and distortion matrix, respectively. The multi-gyro filter consistently estimates all gyro parameters without filter bias. Results for the second gyro also show consistent and unbiased parameter estimation.

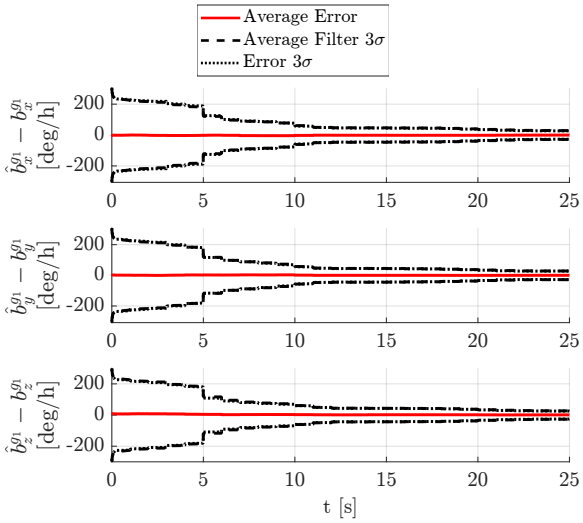

Fig. 2 The multi-gyro filter's bias estimation performance for the propagation gyro.

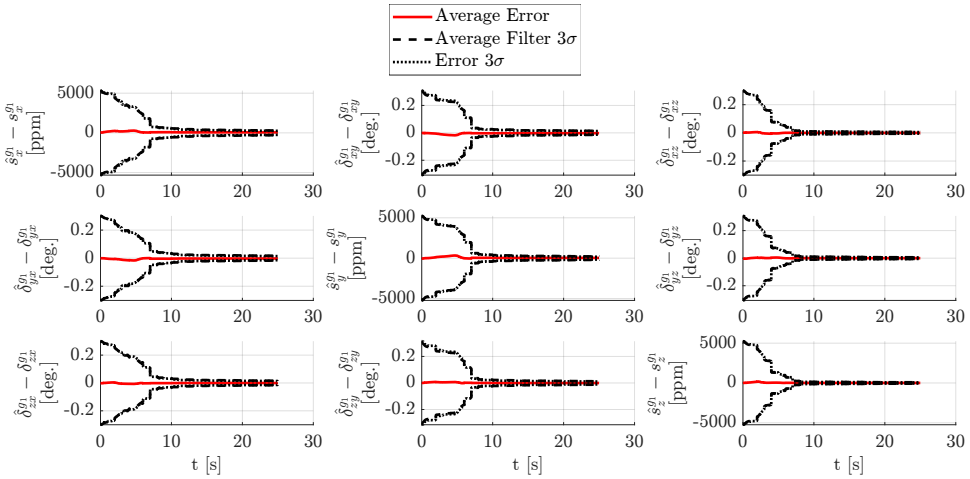

Fig. 3 The multi-gyro filter's distortion matrix estimation performance for the propagation gyro. Units for scale factors and misalignments are ppm **and** deg**, respectively.**

I. Fault Detection and Isolation Testing

FDI perfomance is validated with a 100 run Monte Carlo and the three gyro filter. Two gyro failure modes are presented in Section IV.A are used in two separate tests. Gyro noise characteristics, initial estimate uncertainties, and most vehicle properties are identical to those used in the previous section and are available in Tables 1, 2, and 3. Gyro failure parameters given in Table 4. The only vehicle properties which do not match those in Table 2 are the relative orientations of the gyros.

As discussed previously, the filter's measurement residuals are designed to improve the signal to noise ratio of any given fault signal. This signal to noise ratio can be improved further by orienting the gyros such that no two gyro axes are aligned. To demonstrate this performance FDI is performed for two sets of gyro arrangements. One in which all three gyros are aligned and another in which they are oriented optimally as described by Equations (58) and (59). The

Table 4 Failure parameters for FDI testing.

α	0.01
W^2_{fail}	0.743
T_{fail}	3 s
N_{CVMT}	1000
T_ϵ	25 s
$\boldsymbol{\epsilon}^{g_1}$	$\begin{bmatrix}1\text{e}{-4} & 0 & 0\end{bmatrix}^T$ rad/s^2
T_A	25 s
$\boldsymbol{A}^{g_2}$	blkdiag $\{1,\ 0,\ 0\}$

Table 5 Relative transformations between gyro frames for the FDI demonstration.

$\boldsymbol{q}^{g_1}_{\text{nav}}$	$\begin{bmatrix}0 & 0 & 0 & 1\end{bmatrix}^T$
$\boldsymbol{q}^{g_2}_{\text{nav}}$	$\begin{bmatrix}0.8624 & 0.2500 & -0.2500 & -0.3624\end{bmatrix}^T$
$\boldsymbol{q}^{g_3}_{\text{nav}}$	$\begin{bmatrix}-0.3624 & 0.2500 & -0.2500 & 0.8624\end{bmatrix}^T$

scalar-last quaternions describing relative transformations between gyro frames in this case provided in Table 5.

The first test is a bias drift in the x axis of gyro number one as described in Table 4. Figure 4 shows the $W^2_{g_i}(j)$ history for each component of each gyro when the gyros are aligned for one Monte Carlo run. Figure 5 shows the $W^2_{g_i}(j)$ history when the gyros are positioned as described in Table 5 for another Monte Carlo run using the same process and measurement noise sequence. The test stops once a $W^2_{g_i}(j)$ has exceeded W_{fail} for more than T_{fail} seconds. In both gyro arrangements the fault is detected and correctly identified for all Monte Carlo runs.

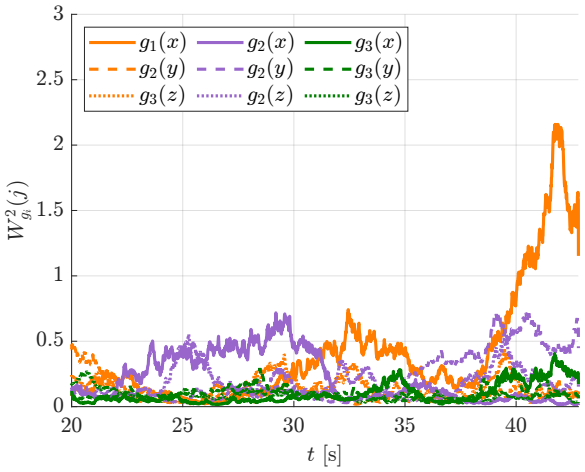

Fig. 4 Cramer-von Mises test statistics under a g_1 bias drift failure at 25 seconds with aligned gyros.

The relative orientation of the gyros in the second arrangement makes fault isolation easier by improving the ratio between $W^2_{g_1}(1)$ and the next highest value. This ratio is recorded and averaged across all Monte Carlo runs. Also

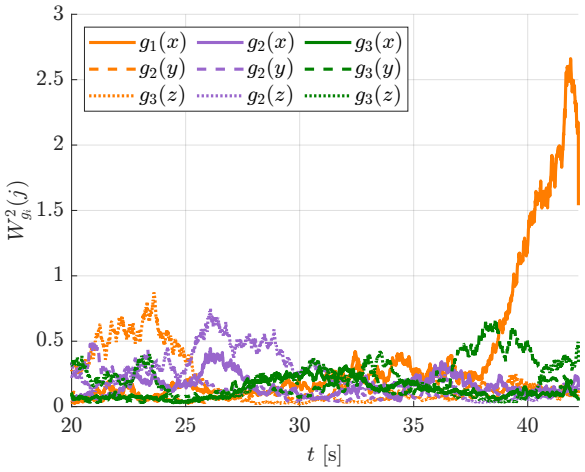

Fig. 5 Cramer-von Mises test statistics under a g_1 bias drift failure at 25 seconds with non-aligned gyros.

Table 6 Monte Carlo averaged fault detection and isolation performance for a gyro bias drift fault.

Item	Aligned Gyros	Optimally Oriented Gyros
Mean Fault Isolation Ratio	5.0	5.4
Mean Fault Detection Time [s]	26.2	26.0

averaged and recorded is the time of fault detection. These results are displayed in Table 6. Aligning the gyros optimally provides a modest improvement in fault isolation ratio. There is no practical improvement in fault detection time.

The second test introduces additional and unmodeled angular random walk in the x axis of the second gyro as described in Table 4. Figure 6 shows the $W^2_{g_i}(j)$ history for each component of each gyro when the gyros are aligned for one Monte Carlo run. Figure 7 shows the $W^2_{g_i}(j)$ history when the gyros are positioned as described in Table 5 for another Monte Carlo run using the same process and measurement noise sequence. Once again the test stops once a $W^2_{g_i}(j)$ has exceeded W_{fail} for more than T_{fail} seconds. In both gyro arrangements the fault is detected and correctly identified in every Monte Carlo run.

As before, the averaged fault isolation ratio and stopping time are displayed in Table 7. In this test optimal gyro alignment provides a significant performance benefit to the fault isolation ratio. Faults are also detected much quicker than the previous case, but once again optimal gyro alignment does not offer much improvment in terms of detection time.

Table 7 Monte Carlo averaged fault detection and isolation performance for a gyro ARW fault.

Item	Aligned Gyros	Optimally Oriented Gyros
Mean Fault Isolation Ratio	8.1	13.9
Mean Fault Detection Time [s]	13.0	12.9

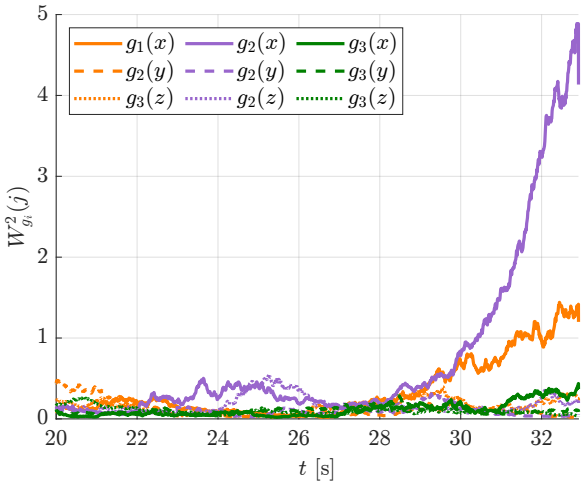

Fig. 6 Cramer-von Mises test statistics under a g_2 angular random walk failure at 25 seconds with aligned gyros.

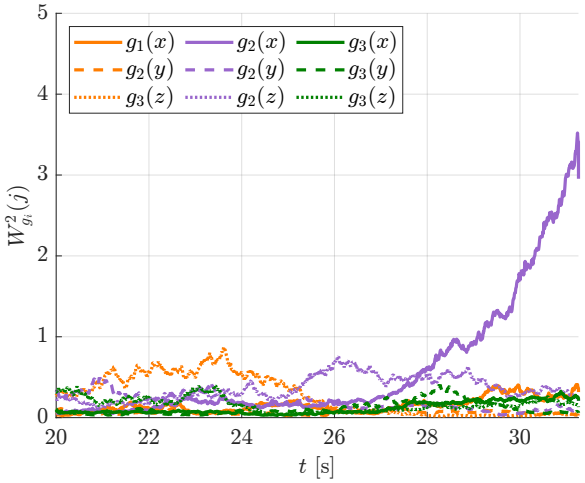

Fig. 7 Cramer-von Mises test statistics under a g_2 angular random walk failure at 25 seconds with non-aligned gyros.

V. Conclusion

This paper proposes a performance index to design optimal relative orientations of multiple three-axis IMUs. It also analytically derives the two-IMU solution and verifies the optimally of a proposed three-IMU solution.

We show that the choice of performance index maximizes the L_1 norm of the matrix mapping the measurement error into the fault vector. While other matrix norms could be maximized, we deem the L_1 norm the most appropriate as it maximizes the projection of each and every error into each and every component of the fault vector. Matrix norms based on spectral decomposition were shown to be inadequate.

Additionally, we derive a new Kalman filter with embedded IMU FDI. In the presence of redundant IMUs, one can be used in a model replacement mode for the filter's propagation phase. The remaining IMU measurements can be fused into the state with the proposed Kalman update formulation. Incorporating additionally external measurements can allow

observability of the vehicle position, velocity, and attitude as well as IMU's biases, scale factors, and misalignments.

Through Monte Carlo methods we've shown that our system provides equivalent performance to an MEKF which averages gyro measurements into a synthetic angular rate measurement with reduced noise characteristics. When a synthetic measurement is used the parameters of each gyro are not independently observable but computational complexity is reduced. Because our system maintains observability of all gyro parameters we are able to develop an efficient method for gyro fault detection based on the Cramer-von Mises goodness of fit test. We show that gyro faults can be quickly and efficiently detected by our system. Furthermore, these faults are more likely to be detected correctly when the gyros are aligned optimally as described by the first part of our work.

Acknowledgments

This work was supported by a NASA Space Technology Graduate Research Opportunity, NASA grant number: 80NSSC20K1195.

References

[1] Hwang, I., Kim, S., Kim, Y., and Seah, C. E., "A Survey of Fault Detection, Isolation, and Reconfiguration Methods," *IEEE Transactions on Control Systems Technology*, Vol. 18, No. 3, 2010, pp. 636–653. doi:10.1109/TCST.2009.2026285.

[2] Savage, P. G., "Strapdown System Algorithms," *Advances in Strapdown Inertial Systems, Lecture Series 133*, 1964.

[3] Markley, F. L., and Crassidis, J. L., *Fundamentals of Spacecraft Attitude Determination and Control*, Space Technology Library, Springer, 2014. doi:10.1109/MCS.2015.2427046.

[4] Parkinson, B. W., and Spilker Jr., J. J., *Global Positioning System: Theory and Applications*, Progress in Astronautics and Aeronautics, Vol. 163, American Istitute of Aeronautics and Astronautics, Washington, D.C., 1996. doi:10.2514/4.866388.

[5] Trawny, N., Mourikis, A. I., Roumeliotis, S. I., Johnson, A. E., and Montgomery, J. F., "Vision-aided inertial navigation for pin-point landing using observations of mapped landmarks," *Journal of Field Robotics*, Vol. 24, No. 5, 2007, pp. 357–378. doi:10.1002/rob.20189.

[6] Mamich, H., and D'Souza, C., "Orion preliminary navigation system design," *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2008, p. 7295. doi:10.2514/6.2008-7295.

[7] Goodman, J. L., "Space shuttle navigation in the GPS era," *Proceedings of the 2001 National Technical Meeting of The Institute of Navigation*, 2001, pp. 709–724.

[8] THIBODEAU, J., and Bauer, S. R., "Redundancy management of the Space Shuttle inertial measurement units," *Guidance and Control Conference*, 1982, p. 1561. doi:10.2514/6.1982-1561.

[9] Sturza, M. A., "Navigation System Integrity Monitoring Using Redundant Measurements," *Journal of the Insitute of Navigation*, Vol. 35, No. 4, 1988–89, pp. 69–87. doi:10.1002/j.2161-4296.1988.tb00975.x.

[10] Gertler, J., "Fault Detection and Isolation Using Parity Relations," *Control Eng. Practice*, Vol. 5, No. 5, 1997, pp. 653–661. doi:10.1016/S0967-0661(97)00047-6.

[11] Pittelkau, M. E., "Calibration and attitude determination with redundant inertial measurement units," *Journal of Guidance, Control, and Dynamics*, Vol. 28, No. 4, 2005, pp. 743–752. doi:10.2514/1.7040.

[12] Zanetti, R., Alaniz, A., Breger, L., Mitchell, I., and Phillips, R., "Fault Detection and Isolation Strategy for Redundant Inertial Measurement Units," *Proceedings of the 2014 Space-Flight Mechanics Meeting*, Advances in the Astronautical Sciences, Vol. 152, Santa Fe, NM, 2014, pp. 671–684. AAS 14-246.

[13] Jafari, M., "Optimal redundant sensor configuration for accuracy increasing in space inertial navigation system," *Aerospace Science and Technology*, Vol. 47, 2015, pp. 467–472. doi:10.1016/j.ast.2015.09.017.

[14] Rahaim, N., and Wood, B., "IMUFDIR Tuning In Response To Structural Deformation Of The Orion Capsule," *44th Annual AAS Guidance, Navigation and Control (GN&C) Conference*, 2022. doi:10.1007/978-3-031-51928-4_31.

[15] Li, Z., Liu, G., Zhang, R., and Zhu, Z., "Fault detection, identification and reconstruction for gyroscope in satellite based on independent component analysis," *Acta Astronautica*, Vol. 68, No. 7-8, 2011, pp. 1015–1023. doi:10.1016/j.actaastro.2010.09.010.

[16] Törnqvist, D., Gustafsson, F., and Klein, I., "GLR tests for fault detection over sliding data windows," *IFAC Proceedings Volumes*, Vol. 38, No. 1, 2005, pp. 743–748. doi:10.3182/20050703-6-CZ-1902.00125.

[17] de Alteriis, G., Conte, C., Moriello, R. S. L., and Accardo, D., "Use of consumer-grade MEMS inertial sensors for accurate attitude determination of drones," *2020 IEEE 7th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, IEEE, 2020, pp. 534–538. doi:10.1109/MetroAeroSpace48742.2020.9160134.

[18] Fontanella, R., de Alteriis, G., Accardo, D., Moriello, R. S. L., and Angrisani, L., "Advanced low-cost integrated inertial systems with multiple consumer grade sensors," *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*, IEEE, 2018, pp. 1–8. doi:10.1109/DASC.2018.8569846.

[19] Rogne, R. H., Bryne, T. H., Fossen, T. I., and Johansen, T. A., "Redundant MEMS-based inertial navigation using nonlinear observers," *Journal of Dynamic Systems, Measurement, and Control*, Vol. 140, No. 7, 2018, p. 071001. doi:10.1115/1.4038647.

[20] Patel, U. N., and Faruque, I. A., "Multi-IMU Based Alternate Navigation Frameworks: Performance & Comparison for UAS," *IEEE Access*, Vol. 10, 2022, pp. 17565–17577. doi:10.1109/ACCESS.2022.3144687.

[21] Beaudoin, Y., Desbiens, A., Gagnon, E., and Landry Jr, R., "Satellite launcher navigation with one versus three IMUs: Sensor positioning and data fusion model analysis," *Sensors*, Vol. 18, No. 6, 2018, p. 1872. doi:10.3390/s18061872.

[22] Tapley, B. D., Schutz, B. E., and Born, G. H., *Statistical Orbit Determination*, Elsevier Academic Press, 2004. doi:10.1016/B978-0-12-683630-1.X5019-X.

[23] Banica, T., Collins, B., and Schlenker, J.-M., "On orthogonal matrices maximizing the 1-norm," *Indiana University mathematics journal*, 2010, pp. 839–856. doi:10.48550/arXiv.0901.2923.

[24] Shuster, M. D., "The nature of the quaternion," *The Journal of the Astronautical Sciences*, Vol. 56, No. 3, 2008, pp. 359–373. doi:10.1007/BF03256558.

[25] Thienel, J. K., "Nonlinear observer/controller designs for spacecraft attitude control systems with uncalibrated gyros," Ph.D. thesis, University of Maryland, College Park, 2004.

[26] Cramér, H., "On the composition of elementary errors: First paper: Mathematical deductions," *Scandinavian Actuarial Journal*, Vol. 1928, No. 1, 1928, pp. 13–74. doi:10.1080/03461238.1928.10416862.

[27] Von Mises, R., "Statistik und wahrheit," *Julius Springer*, Vol. 20, 1928. doi:10.1007/978-3-662-41863-5.

[28] Csörgő, S., and Faraway, J. J., "The exact and asymptotic distributions of Cramér-von Mises statistics," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, Vol. 58, No. 1, 1996, pp. 221–234. doi:10.1111/j.2517-6161.1996.tb02077.x.